UNIVERSITY OF CALIFORNIA SAN DIEGO

Methods for Voice and Swallow Assessment through Laryngeal High-Density Surface Electromyography

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Bioengineering

by

Gladys Ornelas

Committee in charge:

Professor Todd P. Coleman, Chair
Professor Pedro Cabrales
Professor Jacqueline Greene
Professor Tara Javidi
Professor Kevin King
Professor Philip A. Weissbrod

2021

The dissertation of Gladys Ornelas is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

To my parents, for all that they've done to support me and provide me with the opportunities they never had.

To my siblings, for enriching my life through their companionship.

To my partner, for bringing love and joy into my life.

To my non-human children, especially Kero, Effie, and Iglu, for their humorous antics and unwavering emotional support.

EPIGRAPH

Is it so bad, then, to be misunderstood?
Pythagoras was misunderstood, and Socrates, and Jesus,
and Luther, and Copernicus, and Galileo, and Newton,
and every pure and wise spirit that ever took flesh.
To be great is to be misunderstood.
— Ralph Waldo Emerson, Self-Reliance

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Professor Todd Coleman. You've provided me with an invaluable opportunity to delve into research, allowing me to experience multiple aspects ranging from fabrication, to clinical research, to data analysis. I recognize how lucky I have been to experience the diverse and multi-disciplinary field of Bioengineering in its fullest. Most importantly, throughout my academic career, you have always pushed me to accomplish more than I thought I could. Through your empowering mentorship, I've learned what it takes to be a mentor and hope to pass that on to others some day.

Secondly, I would like to thank my co-advisor, Dr. Philip Weissbrod. This project would have never been possible without his guidance, clinical expertise, and support. In addition, I want to thank Dr. David Bracken, who contributed significantly to the organization, development, and data collection of this project with his clinical and musical expertise. I would also like to thank my committee members, Dr. Jacqueline Greene and Professors Kevin King, Pedro Cabrales, and Tara Javidi for our insightful conversations and your guidance throughout this process.

Thank you to my incredible mentors: Dr. Sheila Rosenberg, you shared with me your passion for science and were pivotal in my decision to pursue academic research. Dae, you went above and beyond the role of a grad student mentor. If it weren't for your instrumental guidance in navigating higher education, research, and setting a great example to follow, I would not be a scientist, much less a graduate student. Thank you to my lab-mates, Jonas, Andrew, Gabe, and Phil, for being wonderful colleagues. Alexis, thank you so much for helping me with my coding and technical difficulties, as well as for our enlightening conversations. A special thank you to my office-mate and close friend, Anjulie. You taught me how to work hard and play hard. I will always think fondly of our lab memories, and I will also make sure my next office is stocked with chocolate. Lastly, I am so thankful for my undergraduate research assistants: Hassler Bueno, Ruby Rios-Ramirez, and Xiaoyi Ma. You all have been a joy to work with. Your brilliance, cleverness, and passion for science continues to motivate and inspire me.

I am also grateful to my high school teacher, Mrs. Morales, for introducing me into the world of science and research. Thank you to my parents, Benjamin and Guadalupe, for all that you sacrificed for me, and teaching me the importance of hard work and education. Thank you to my siblings, Lupita and Benji, for keeping my spirits up. Thank you to my wonderful friends: Emma, Asami, Maria, Natalie, and Srihita, you have kept me sane with your companionship, memes, and friendship during life's turbulent times. And Emma, how can I ever thank you for helping me edit this dissertation and the boba runs? Clara, thank you so much for not only being a wonderful friend, but for literally being there for me every day through our productivity-pals work sessions. My thesis would not be where it's at without you. Finally, to my partner, Rogelio, for your unwavering love, humor, and support.

I am thankful for Jack and Judy Elder, American Laryngological Association, San Diego Fellowship, Interfaces Training program, and Frontiers of Innovation Scholars Program for your financial and academic support.

Chapter 2, Section 2.1, in full, is a reprint of the material as it appears in The Laryngoscope 2019. Bracken, David J. ; Ornelas, Gladys; Coleman, Todd P.; Weissbrod, Philip A., The Laryngoscope 2019. The dissertation author was the second author of this paper.

Chapter 2, Section 2.2, in full, is currently in preparation for submission for publication of the material. Ornelas, Gladys; Bracken, David J. ; Weissbrod, Philip A.; Coleman, Todd P., The dissertation author was the primary investigator and author of this material.

Chapter 3, Section 3.1, in full is currently in preparation for submission for publication of the material. Bracken, David J.; Ornelas, Gladys; Weissbrod, Philip A.; Coleman, Todd P., The dissertation author will be the co-author of this paper.

Chapter 3, Section 3.2, in full is currently in preparation for submission for publication of the material. Ornelas, Gladys; Bracken, David J.; Bueno Garcia, Hassler; Weissbrod, Philip A.; Coleman, Todd P., The dissertation author was the primary investigator and author of this material.

Chapter 4, Section 4.1, in part is currently being prepared for submission for publication of the material. Ornelas, Gladys; Ma, Xiaoyi; Rios-Ramirez, Ruby; Weissbrod, Philip A.; Coleman, Todd P., The dissertation author was the primary investigator and first author of this paper.

Chapter 4, Section 4.2, in part is currently being prepared for submission for publication of the material. Ornelas, Gladys; Kurniawan, Jonas F.; Nguyen, Andrew; Pham, Timothy; Sit, Nathan L.J.; Rios-Ramirez, Ruby; Weissbrod, Philip A.; Coleman, Todd P., The dissertation author was the primary investigator and first author of this paper.

VITA

2016	B. S. in Bioengineering: Bioengineering, University of California San Diego
2018	Teaching Assistant University of California, San Diego
2021	Ph. D. in Bioengineering, University of California San Diego

PUBLICATIONS

D. J. Bracken, G. Ornelas, T. P. Coleman, and P. A. Weissbrod, "High-density surface electromyography: A visualization method of laryngeal muscle activity," Laryngoscope, no. 4, pp. 1–7, 2019, doi: 10.1002/lary.27784.

ABSTRACT OF THE DISSERTATION

Methods for Voice and Swallow Assessment through Laryngeal High-Density Surface Electromyography

by

Gladys Ornelas

Doctor of Philosophy in Bioengineering

University of California San Diego, 2021

Professor Todd P. Coleman, Chair

Voice and swallow are complex functions made possible through the coordination of multiple muscles of the throat. Unfortunately, these tasks are adversely impacted by aging, neurologic disorders, nerve injuries, cancer, and stroke—yet lack the tools for objective and non-invasive assessment. For instance, traditional surface electromyography (sEMG) of the throat suffers from drawbacks of "cross-talk" contamination, skin-electrode impedance, and diminished target-muscle specificity, which lead to performance variability and limited clinical utility. This dissertation explores the use of high-density surface electromyography (HDsEMG) coupled with novel implementation of array signal processing techniques to overcome limitations of traditional

sEMG when studying the neck. During phonation in healthy subjects, results yielded power spectrum density energy maps with the capacity to visually distinguish active regions associated with the underlying cricothyroid and anterior strap musculature. Low-pitch and high-pitch differentiation was accomplished using multivariate log likelihood ratio testing with an average Receiver Operating Characteristic area under the curve of 0.97, which exceeds that of traditional sEMG by 0.20. During swallowing in healthy subjects, HDsEMG energy maps confirmed lateral symmetry and dominant activity in the suprahyoid region. Additional studies conducted on human subjects utilizing various swallow textures and complexities demonstrated average EMG duration that increased proportionally with increasing texture complexity. Multivariate analysis improved automated detection of onsets and offsets of swallows and was able to classify one of five distinct textures with an average probability error of 0.16. Preliminary results for validation against the "gold-standard", needle EMG, demonstrate HDsEMG's ability to detect specific localized activity similar to the needle electrode underneath. Lastly, we demonstrate the feasibility of using flexible electronic sensor arrays, in lieu of standard needle and clunky electrodes arrays, to provide greater subject comfort, mobility, and adhesion to the curvature of the neck.

Chapter 1

Introduction

1.1 Prevalence of Swallow and Voice Disorders

Phonation and swallow are basic and necessary functions made possible through complex interactions between the larynx, pharynx, upper aerodigestive tract, and respiratory system. This system can be adversely affected by a variety of disorders, leading to voice or swallow dysfunctions known as dysphonia or dysphagia respectively.

In the elderly, rates of dysphonia have been reported from 20-47% [135]. These disorders can be debilitating and isolating [44, 88, 106], as they can negatively impact an individual's ability to communicate in person or via technology, subsequently reducing quality of life. Hyperfunctional disorders (e.g. overuse or misuse of voice muscles) and neurolaryngeal disorders (e.g. Parkinson's disease or nerve injuries) account for 18% of cases in the general population [22], but increases to 50% for those whose voice is essential for their career such as educators [8].

Like dysphonia, dysphagia can be debilitating and lower the quality of a patient's life. Uncoordinated swallows, or diminished ability to swallow, restricts the patient to certain types of foods and methods of eating, limiting their ability to lead independent and socially active lives. Additionally, swallow disorders may become detrimental or life threatening to patients as they

are at risk for aspiration, pneumonia, and malnutrition [40, 143]. Dysphagia can result from a variety of ailments such as stroke, neurologic disorders, and cancer affecting the head or neck [21]. Aging is another factor, as the elderly rate of dysphagia has been reported to be 14-38% [135].

1.2 Clinical and Therapeutic Assessment

Despite the high prevalence of voice and swallow disorders, their diagnosis and assessment are dependent on subjective interpretation of a patient's history and physical examination, often involving invasive, highly technical, and uncomfortable techniques. Otolaryngologists and Speech-Language Pathologists (SPLs) may use a variety of tools such as MRIs, CAT scans, laryngoscopy, videostroboscopy, esophageal manometry, and needle eletcromyography, to diagnose voice disorders, with the additional tools of esophagograms and videofluoroscopic swallow studies for swallow disorders [12, 13, 100, 129] (Fig. 1.1). As one can imagine, the clinical expertise, time, and costs required to use such methods limit their applicability and setting to the clinic. Management of symptoms is typically done through medication or surgery that tackles the underlying cause of the voice or swallow disorder, though it may not always help with the voice or swallow disorder itself. Additionally, SPLs may utilize physiologic voice and swallow therapy to help patients manage their symptoms through specialized maneuvers, their respiration, resonance, and phonation. However, its effectiveness, especially in relation to swallow disorders, seems to vary with the underlying disorder and types of therapeutic treatments [11, 19, 31, 53, 72, 97, 108].

In terms of understanding neuromuscular health and activity, needle and hookwire electromyography (nEMG) remain the gold-standard. nEMG's capability to identify morphological characteristics of the motor unit signal is a direct indicator of neuromuscular health and its relation to voice and swallow disorders [14, 47, 57, 61, 63, 64, 66, 141, 55, 111, 78, 148]. However its application is limited due to the high technical expertise required to place the needles, restricting

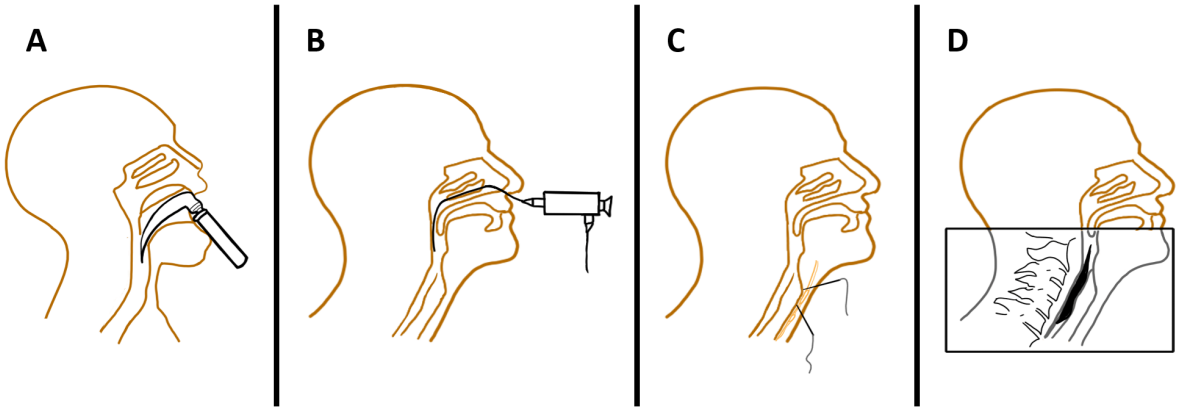

Figure 1.1: An illustration depicting some of the tools used for clinical assessment of voice and swallow disorders. A) Laryngoscopy. B) Videostroboscopy. C) Needle Electromyography. D) Videofluoroscopy.

its application to the clinic, and by patient comfort and pain levels, restricting the number of electrodes that can be inserted, and the length of monitoring.

1.3 The Search for a Noninvasive Assessment Method

Voice and swallow disorders are not limited to the elderly or adult populations. Children are also at risk of developing such disorders but their prevalence is difficult to estimate. In pediatric swallow disorders, estimation is limited due to the lack of standardized assessment procedures and variable definition of dysphagia for children [12]. When considering the vulnerable populations involved with voice and swallow dysfunction, often as a result of other incapacitating disorders, it becomes imperative to develop an objective, scalable, cost-effective, and easy-to-use assessment system that can monitor treatment progression as well.

For voice, attempts have been made to design an objective acoustic assessment system. With the advent of novel voice recognition technologies such as Microsoft Cortana and Amazon Alexa, people increasingly use voice to complete tasks and transactions. However, those with

voice disorders struggle using such technologies, since vocal challenges impact speech patterns and lead to irregularity in pronunciations [95]. Deep learning tools, which power these technologies, struggle to train on such unpredictable patterns, therefore limiting model robustness and fitting interpretability. For example, although deep learning applied during vowel pronunciation has shown promise as an objective classification of Parkinsonian voice [134] and two types of speech dysphonia (laryngeal dystonia and muscle tension dysphonia)[114], its accuracy wavers with different diagnoses and continuous speech [9, 86, 130]. Additionally, within the context of classifying speech disorders, solely using voice is complicated as different disorders can share similarly perceived vocal changes and patients often unknowingly adjust to vocal changes by overcompensating with their muscle use. Therefore it's likely a speech recognition approach that incorporates an understanding of physiologic changes underlying voice disorders is necessary to overcome these performance limitations.

As previously stated, nEMG is the diagnostic tool that can provide objective neurophysiologic data regarding laryngeal dysfunction [140]. Zarzur et al. demonstrated the potential for nEMG in the neck to serve as a diagnostic tool for Parkinson's disease [149]. In their study, nEMG detected hypercontractility during voice rest in 91.5% of their subjects regardless of the PD severity. Despite the promising applications for nEMG, it does not capture voice quality, is invasive and uncomfortable for patients, and requires a high degree of technical skill for application and interpretation. In swallows, nEMG provides an in depth look into the health and pattern activity during a swallow, allowing for both diagnostic and prognostic capabilities [62, 96, 98]. However, swallowing involves a dynamic set of interactions between various muscles of different sizes and locations. The invasive nature of nEMG limits both the number of muscles, and overall space of the neck, that can be monitored.

A non-invasive alternative capable of providing similar neuromuscular information as nEMG, while simultaneously performing voice acoustic analysis or swallow therapeutic maneuvers, will enhance our understanding of voice and swallow disorders and improve assessment.

Such a tool would not only improve assessment of laryngeal disorders, but also our ability to design speech recognition technology that can better serve the voice disorder population.

Studies have examined surface EMG (sEMG) as a tool to diagnose voice and swallow disorders. Its noninvasiveness, ease-of-use, and availability make it an ideal and practical alternative to study neuromuscular health. In laryngology, its application has been investigated for hyperfunctional disorders, dysphagia, and as a biofeedback tool [17, 29, 34, 58, 56, 85, 124, 125], but with limited success as researchers continue to struggle in finding significant difference between healthy subjects and patients diagnosed with voice and swallow disorders [75, 123, 136, 137, 139, 140]. Notably, these studies were limited by a low number of sensors, variability in sensor placement, and lack of voice acoustic analysis, common issues among previous sEMG applications for speech and swallow and lack of standardization [123, 136]. Other studies and even some on-the-market devices such as VitalStim, have investigated the use of sEMG biofeedback during voice and swallow therapy. However, such applications are not standardized, extremely susceptible to bias, and lack an objective assessment of swallow or phonation [20, 30, 105].

Surface EMG has struggled to secure a foothold in voice and swallow assessment due to the lack of standardization, variable results, susceptibility to cross-talk contamination, lack of target-muscle specificity, and loss of intramuscular information. In its traditional setting, sEMG records aggregated muscle action potentials, which makes it difficult to achieve consistent reproducibility and correlate the compounded signals to specific muscles underneath the skin. In essence, its applicability is reduced to detecting binary on-off onset activity and muscle force, features that are not enough to differentiate or assess either voice or swallow disorders due to the overlapping nature and varying sizes of laryngeal muscles. We propose High-density surface electromyography as an alternative and promising solution to traditional sEMG.

1.4 High-Density Surface Electromyography

High Density Surface electromyography (HDsEMG) is a developing technique that places a dense multi-electrode array over a target region. It is a non-invasive alternative that overcomes certain limitations of traditional sEMG. HDsEMG enables acquisition of spatiotemporal patterns allowing interelectrode comparisons to elucidate localized activity correlated to the muscles active underneath. Muscle fiber conduction velocity measurements and single motor unit evaluations are possible through HDsEMG, which enables detection of pathological changes (such as due to neurogenic disorders) at the motor unit level [37, 93]. When using appropriately sized electrodes, HDsEMG provides information regarding innervation zones, fiber length, fiber conduction velocity, and wave activity of muscle contractions [42, 50, 52, 68, 89, 109, 115, 133, 146]. Fig. 1.2 taken from Drost et al. illustrates the scope of information obtained from the three different EMG conventions.

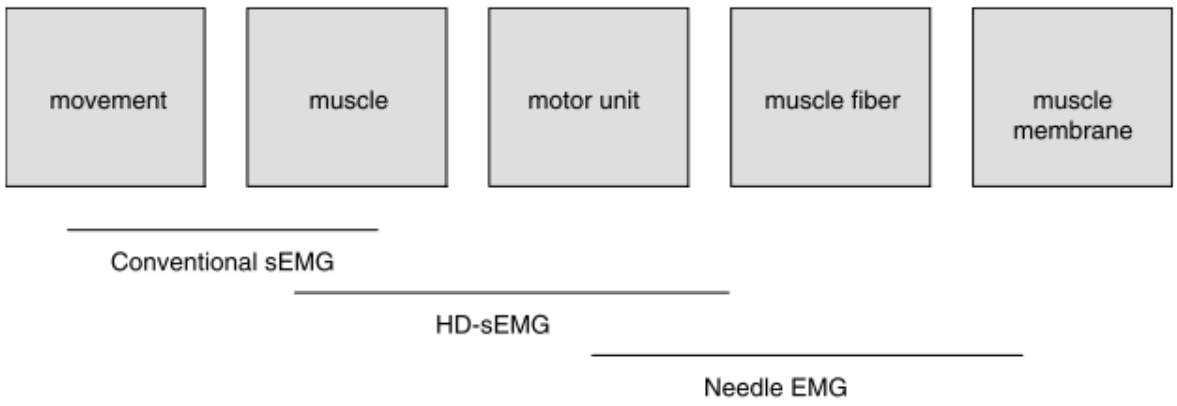

Figure 1.2: This image taken from Drost et al. [37] shows the level of information obtained from EMG configurations.

HDsEMG monitoring provides a large dataset that requires, and makes use of, advanced signal processing techniques such as multivariate analysis, logistic regression, decomposition, weights analysis, amongst others [48, 93]. In using information from all channels, signal quality and robustness is improved while simultaneously providing quantifiable EMG features for an

objective assessment. Though these qualities are exciting for researchers and applications of brain- and human-computer-interfaces, it may in part explain why HDsEMG applications have been limited in the clinic. In clinical applications, many neurophysiologists continue to rely on visual inspection of onscreen signals and auditory evaluation of the discharging units, despite the extensive effort made to quantify nEMG [37]. This is no longer possible with HDsEMG as it requires signal processing and analysis to make use of the spatiotemporal information. Though HDsEMG requires less technical expertise in application and is very easy to use, it does require greater computational ability in order to process and analyze the data.

None-the-less, in research, its applications generate excitement for the possibility of expanding our current day knowledge of neuromuscular disorders and its effects in every day life. For instance, surface electrogastrogram (EGG) was once deemed unusable until studies involving multi-electrode arrays demonstrated the ability to monitor in ambulatory settings, detect peristalsis, and categorize the level of severity in patients with functional Dyspepsia and Gastroparesis, something even clinicians struggle to do [4, 5, 38, 49, 50, 51, 52]. In voice and swallow ongoing studies demonstrate how HDsEMG, coupled with advanced signal processing, can observe the dynamic interplay of laryngeal muscles and differentiation task-specific activities during phonation and swallow [152, 153, 154, 155, 158, 157]. HDsEMG continues to demonstrate its value in monitoring and decoding complex muscular activities without patient discomfort and pain.

The level of information garnered from HDsEMG is paradigm shifting and will change the future of HDsEMG applications in clinical and therapeutics settings. This dissertation will investigate its applications for phonation and swallow. Chapter 2 demonstrates the work done in applying HDsEMG for phonation analysis. In our first publication [25] (Ch 2.1) we demonstrated our ability to differentiate the cricothyroid muscle from the strap muscle during phonatory tasks and detect symmetric activation of the neck. Section 2.2 demonstrates our ability to differentiate low and high pitch phonation, an ability that is diminished with reduced electrodes, illustrating the

importance of a multi-electrode array. The ability to detect low and high phonation is significant as we would expect this distinction to be reduced in unhealthy patients, and therefore can be used as a metric to monitor treatment progression.

Chapter 3 of this dissertation investigates the application of HDsEMG for swallow analysis. Two modalities were used: the first used a 20-channel HDsEMG to study salivary swallows, once again demonstrating symmetry in lateral muscle activity and greater activation energy in the upper rows of the array; the second used a 64-channel to investigate varying swallows using five different swallow complexities: saliva, water, applesauce, fruit cup, and cracker. In this paradigm, we once again find symmetric activity and greater upper-array activity, but we are also able to differentiate the types of swallow. This is significant as we can use this assessment to monitor treatment progression over time, improve therapeutic modalities, and provide an objective assessment of swallow disorders.

Chapter 4 of this dissertation investigates the validation of HDsEMG through concurrent needle and surface electromyography with recordings from the forearm. In efforts to further maximize patient comfort and long-term monitoring, the second half of this chapter will investigate the use of flexible electronic sensors in the neck.

Chapter 2

Phonation

2.1 High-Density Surface Electromyography: A Visualization Method of Laryngeal Muscle Activity

2.1.1 Introduction

Laryngeal electromyography (EMG) is a well-established means to evaluate the neuromuscular activity of both intrinsic and extrinsic laryngeal musculature.[14, 15, 16, 32, 47, 57, 58, 61, 63, 64, 66, 84, 112, 141] Two modalities of EMG have emerged over decades of electrophysiologic inquiry: invasive needle-based electromyography (nEMG) and surface electromyography (sEMG).

nEMG is an invasive test, potentially uncomfortable for patients, and requires technical skill for placement of electrodes and interpretation of tracings. In most cases, the duration of inquiry is also limited due to the necessity of clinical and/or lab settings for needles or hook wires to collect electrophysiologic signal.[32, 65] It is, however, an excellent test for identifying morphological characteristics of the motor unit signal and has applicability in numerous clinical scenarios.[14, 47, 57, 55, 61, 63, 64, 66, 78, 141, 111, 148] The placement of needle electrodes

directly into target tissues by experienced operators affords confidence that the signal obtained is reflective of the interrogated target muscle. Practical concordance between nEMG signals of the same laryngeal muscle have been found to be 95% in experienced hands.[32] Although extremely informative, it is not practical for widespread use due to the aforementioned limitations.

sEMG within the field of laryngology has been used for evaluation of laryngeal hyperfunction and swallow with inconsistent results.[10, 34, 67, 127, 139, 152] Surface EMG 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 interposed or neighboring active muscles resulting in cross talk, and (4) limitations in the ability to describe wave morphology.[37] Its appeal, however, is ease of use, patient comfort, and ability to collect data over extended periods of time.

High-density sEMG (HD sEMG) has the potential to compensate for the previously described spatial selectivity limitations by application of a large number of electrodes within defined area. Each electrode results in a discrete detection volume defined by interelectrode distance and electrode surface 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, as detection volume decreases, the summative contents of the electromyographic signal become more individualized. Although issues of impedance and identification of wave morphology remain, high-density arrays potentially provide a means for differentiation of muscle activation through electrode comparison. High-density arrays allow for evaluation of each point in comparison to the other across phonatory tasks. Additionally, gross visualization of activity can be intuitively displayed as power-density energy maps.[152, 153, 154] Task-specific muscle activation is used to highlight differences between adjacent musculature. In this article, we propose that we can differentiate cricothyroid (CT) muscle activity from rest, and that we can identify different muscle activity patterns between low- and high-pitch phonation. Additionally, we believe there to be a wide variety of diagnostic and therapeutic applications of this technology

should the modality be adequately validated.

2.1.2 Materials and Methods

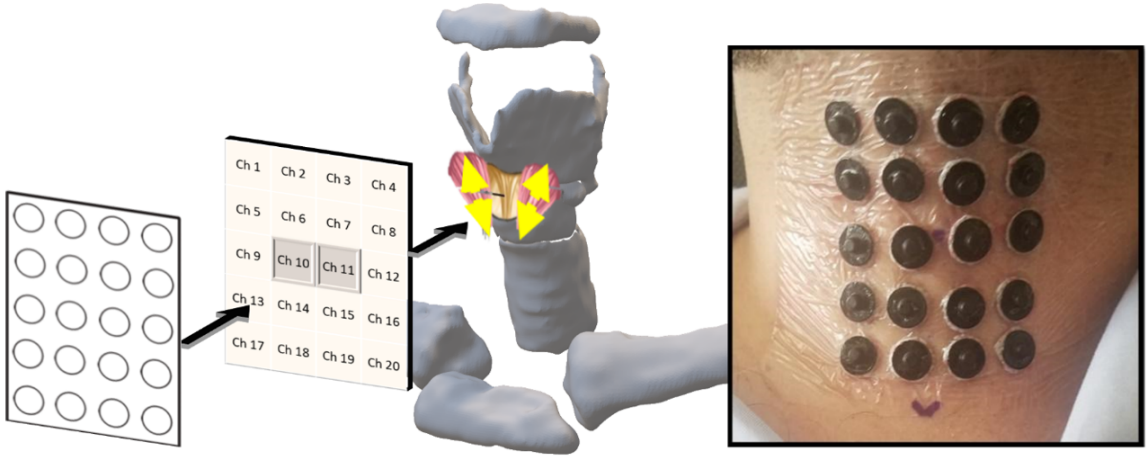

Figure 2.1: High-density surface electromyography array positioned on anterior surface of neck. Central dot (between electrodes 10 and 11) indicates cricothyroid space. Inverted V inferior to array indicates sternal notch. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.

Ten healthy adults (four females, six males, ages 22–51 years, median age = 33.4 years) were enrolled. Inclusion criteria was greater than 18 years of age. Exclusion criteria included a history of laryngeal pathology, a Voice Handicap Index (VHI-10) score of ¿10, subjective dysphonia at the time of recording, previous neck surgery, or neurologic illness.

Ethical approval was granted for this study by the Institutional review board at University of California, San Diego. All participants were seated in a 80 degree slightly reclined position with the neck in slight extension to minimize strap muscle activity at rest. Surface landmarks, specifically the cricothyroid space and sternal notch, were identified with digital palpation and marked for reference. Skin preparation for all electrode sites included standard alcohol wipes and exfoliating impedance-reduction gel (NuPrep Skin Prep Gel; Weaver and Company, Aurora,

CO). Standard electrocardio monitoring electrodes (3 M Red Dot [REF: 2670-5]; 3 M, St. Paul, MN) composed of silver/silver chloride were modified through circumferential removal of the adhesive patch. Electrodes were organized to create a high-density, 20-channel array (five rows by four columns) at 1.5 cm from each electrode center (electrode diameter = 5 mm). An occlusive transparent dressing (Tegaderm; 3 M) was modified with a punch and template. The array was centered on the CT space to ensure electrodes 10 and 11 were overlying the cricothyroid muscles (Fig. 2.1).

Signal Acquisition

A differential amplifier (Brain Vision Device; BrainVision, LLC, Morrisville, NC) was connected to the array, reference and ground electrode were placed overlying the volar surface of the right forearm and left mastoid process respectively, and impedance values were recorded. An audio prompt was played over a 5-minute period for each recorded task. Tasks included the following: (1) rest, (2) low-pitch phonation, (3) high-pitch phonation. Subjects were able to demonstrate adequate difference between low and high pitch as confirmed by the authors with frequency analysis (Audio Frequency Counter; Keuwlsoft, London, United Kingdom). Each task was paced to afford 5 seconds of phonation, on an /i/vowel, followed by 10 seconds of rest. This was repeated for a total of 5 minutes, resulting in 20 recorded intervals of each phonatory task for each participant. EMG data were recorded at a sampling frequency of 500 samples per second using standard EMG recording software (Brain Vision Recorder; BrainVision, LLC). Data were then transferred using EMG analysis software (Brain Vision Analyzer, v. 2.1; BrainVision, LLC) to custom software written in Python programming language (Python Software Foundation, Python Language Reference, v. 3.6.3, https://www. python.org/).

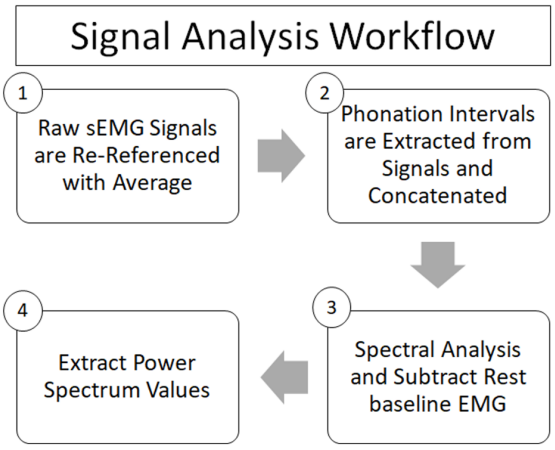

Figure 2.2: Signal analysis method organized by key steps. EMG = electromyography; sEMG = surface electromyography.

Signal Processing

Raw data were processed over a four-step workflow (Fig. 2.2). The average of all the recordings for each dataset was used as a reference and subtracted from the original signals to reduce ambient sources of noise common in all electrode sites.[87] During initial review of data it was noted that in both low- and high-pitch recordings, periodic repetitions of 5-second phonation followed by 10 seconds of rest manifested in oscillatory patterns. Signal analysis sought to characterize EMG contraction only during phonation. Hence, only the segments of each task for the rest, low pitch, and high pitch were considered. In addition, during the onset and offset of phonation, muscles prepare to contract or relax, potentially producing large voltage alterations. Therefore, only the central 3 seconds of each phonatory epoch were extracted. (Fig. 2.3).

Subtraction of rest data from low- and high-pitch data removed the spectral components due to 60 Hz noise (created by the surrounding electrical environment), removed low-frequency drift from electrode motion, and corrected for variability in noise between electrodes. From this, we extracted the average power-density values across all frequency bands. After analysis, separate energy animations of rest, low-pitch, and high-pitch phonatory tasks were created, representing the entire recorded task (Fig. 2.4).

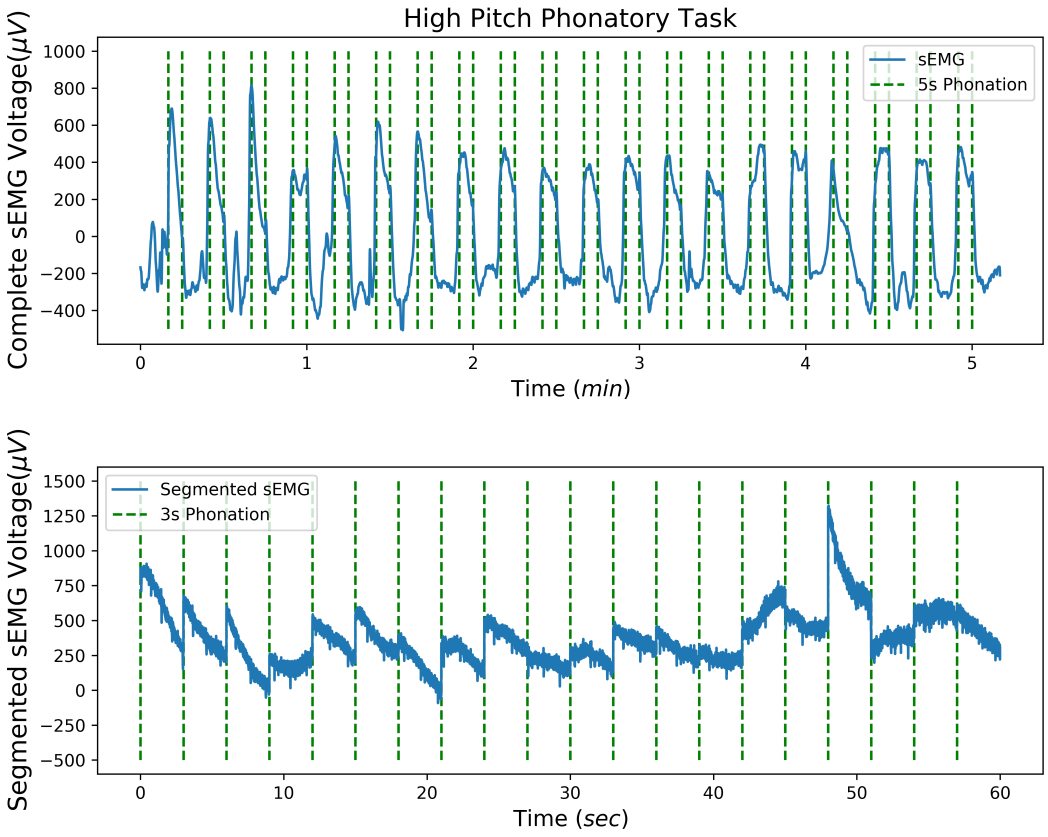

Figure 2.3: Time series plot of signal from electrode 10 of subject 6. (Top) Tracing of a high-pitch task over full 5-minute recording. Vertical dotted lines indicate 5-second phonatory interval. (Bottom) Concatenated segments during high-pitch task. Three-second central intervals over 20 repetitions. sEMG = surface electromyography. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

Statistical analysis of power-density data from each electrode were compared by four methods through one-factor analysis of variance (ANOVA) utilizing Fisher statistics and Pearson coefficients of correlation. Statistical analysis was performed using Microsoft Office Excel (Microsoft Corp., Redmond, WA). Graphical display utilized GraphPad Prism version 7.04 (GraphPad, San Diego, CA,). The first compared each electrode to every other electrode within the same array within the same phonatory task. The second compared the same electrode to itself

between phonatory tasks. The third compared the two left and two right columns by ANOVA analysis. The fourth method compared the two left and two right columns for symmetry by Pearson (r) correlation.

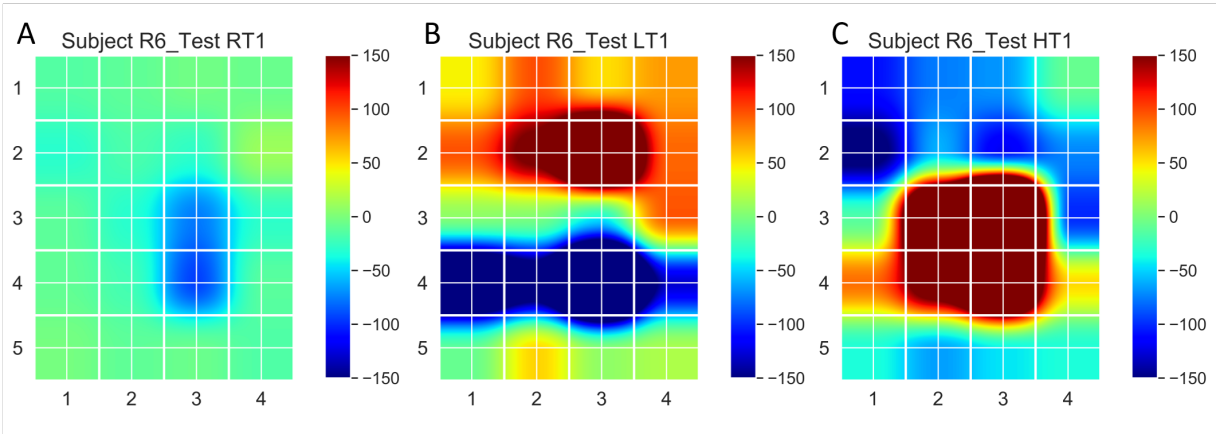

Figure 2.4: Voltage amplitude (μV) energy maps demonstrate the relative intensity of muscle activation within the coronally oriented spatial field. Red indicates increased muscle activation. Blue indicates decreased muscle activation. These maps are representative samples from a single moment in time during each task for subject 6. (A) Rest. (B) Low-pitch phonation. (C) High-pitch phonation. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com.]

2.1.3 Results

Enrollment VHI-10 scores ranged from 0 to 2, and 80% of subjects scored 0 of 40 reflecting healthy phonatory states. Three phonatory tasks (rest, low pitch, high pitch) resulted in spatiotemporal matrices of average spectral power densities across the array. The values for each electrode within the array were subsequently compared to evaluate for significant variance.

Each electrode was numbered and compared to every other electrode within the array for the tasks of rest, low-pitch, and high-pitch tasking per subjects. Each electrode exhibited unique power values despite the same task in 10/10 (100%) subjects ($P < .001$–$.04$), suggesting each electrode is recording a unique underlying signal, despite the uniform task being performed by the subject. Further evaluation compared each electrode to itself between phonatory tasks as follows:

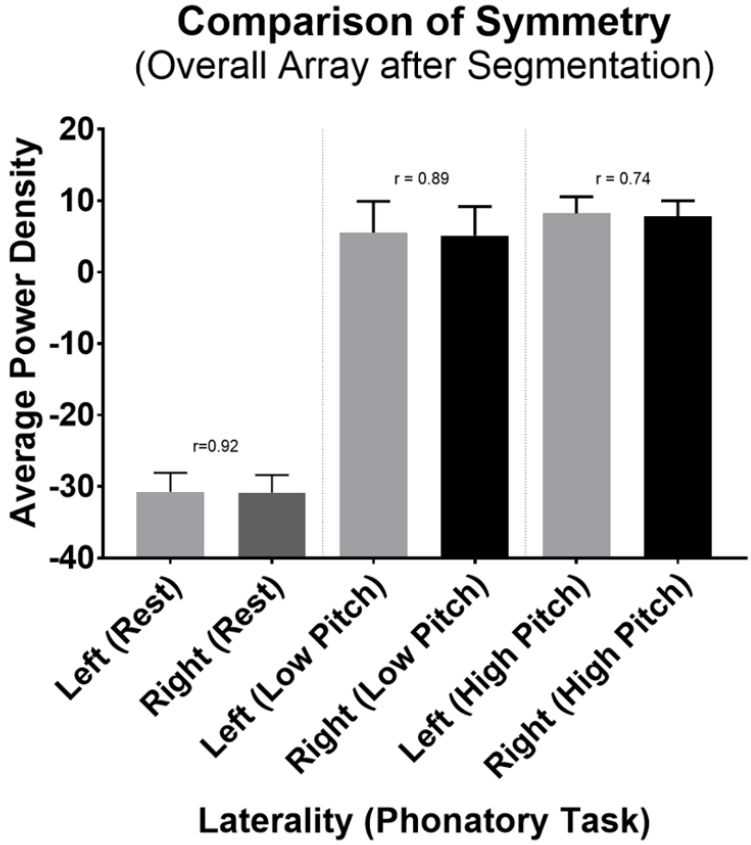

Figure 2.5: Pearson (r) correlation of averaged data of all subjects by task and laterality demonstrating symmetry across the overall array for all subjects. An r approaching 1.0 indicates strong correlation.

rest to low, rest to high, and low to high, yielding rest versus low pitch (10/10, 100%, P $<$ .001), rest versus high pitch (10/10, 100%, P $<$ .001), and low pitch versus high pitch (9/10, 90%, P $<$ .001–.085). Laterality of each array was then analyzed, anticipating symmetry in the context of healthy participants. When the two left and two right columns of the array were considered, there were no statistically significant variances in 9/10 (90%, P = .02–.94) of subjects during rest, 8/10 (80%, P = .03–.64)) during low-pitch phonation, and 10/10 (100%, P = .07–.91) during high-pitch phonation. Pearson (r test) correlation additionally confirmed symmetry between the left and right side at rest (r = 0.92, P $<$ .001), low pitch (r = 0.89, P $<$ .001), and at high pitch (r = 0.74,

P = .015) (Fig. 2.5). Finally, most subjects, 8/10 (80%), demonstrated a pattern of high-pitch phonation as the highest measured average power spectra compared to both rest and low-pitch phonation (Fig.2.6).

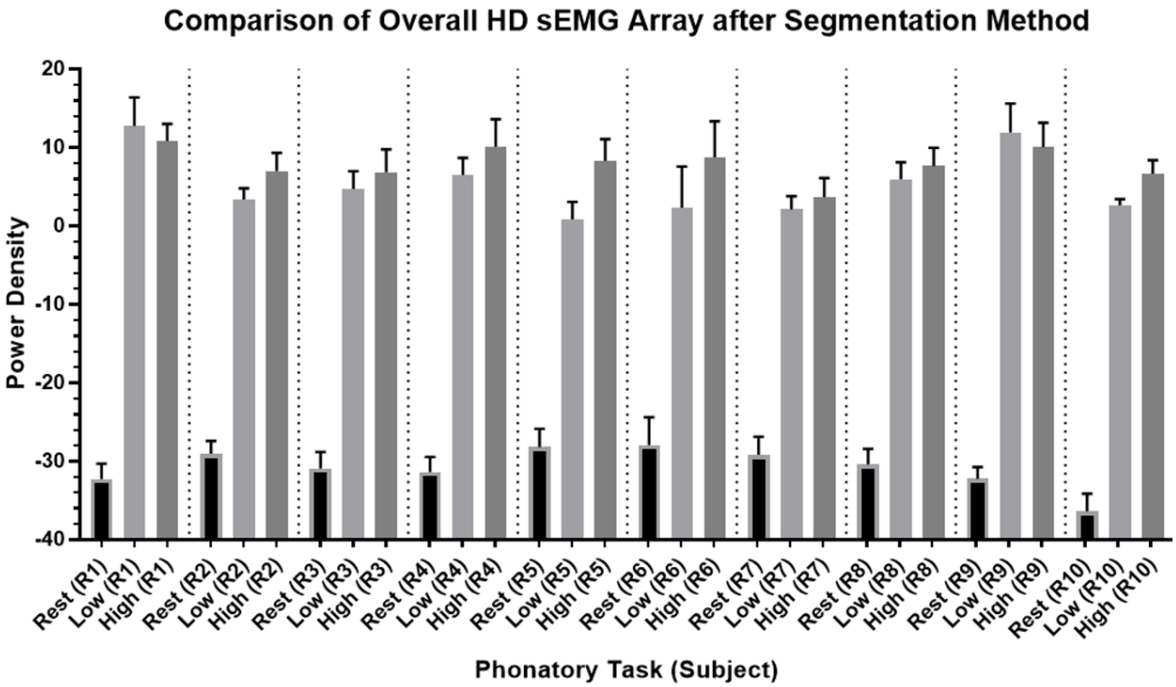

Figure 2.6: Average power spectra for all subjects after isolation of phonatory data. Eight of 10 (overall) demonstrate a pattern in which high-pitchphonation resulted in the highest measured amplitudes. HD sEMG = high-density surface electromyography. [Color figure can be viewed in the online issue, which is available at www.laryngoscope.com]

2.1.4 Discussion

Laryngeal muscle function within the anterior neck is a complex and dynamic process, summative of the interplay between both intrinsic and extrinsic laryngeal muscle activity.[63] During phonation, these muscles may contract in isolation or in concert, depending on the task and extent of recruitment in habitual, hyperfunctional, or hypofunctional states.[61, 67, 84, 139]

sEMG has been explored in numerous medical applications (e.g., electrocardiography, electroencephalography). Applications in laryngology have been varied and previously focused on hyperfunctional disorders, dysphagia, and use as a biofeedback tool. [56, 58, 125] Prior studies for laryngeal or anterior neck application have seen limited success with conflicting data. [56, 58, 123, 124, 125, 126, 128, 159] One of the challenges of sEMG interpretation, especially when only one or a few electrodes are used across a discrete surface area, is the dynamic relationship of skin to underlying structures that occur with movement. This limits interpretation due to uncertainties of muscle position. In other words, is the signal an accurate reflection of activity or has the muscle of interest moved relative to the skin? An HD sEMG array enables comparison of multiple sites and allows for improved spatial resolution. Further visual representation of muscle activity within the anterior neck is possible through power-density energy maps. [131, 152, 153, 159]

The CT muscle has been highlighted in this project due to its anterior position in relation to the laryngotracheal apparatus and absence of overlying cartilage. Additionally, its increased recruitment during high-pitch phonation allows for isolation during specific phonatory tasks. As such, in this study, subjects were asked to maintain a target pitch above 300 Hz for males, and 400 Hz females to ensure reliable CT activation. [69, 80] Contamination of signal by muscles associated with swallowing was minimized, because concurrent swallowing would not be expected within each central 3-second phonatory segment. Gross neck movement was limited by volitional control and cervical support provided by the examination chair. Signal-to-noise ratio (SNR) was reduced by subtracting the power spectral density of the rest task from that of the corresponding low and high phonatory tasks. It is well established in signal detection theory that that SNRs can be improved through evaluation of a repetition paradigm. Comparison of repeated signal allows for noise variance reduction [46].

When electrodes were initially compared to one another within the array during the same task, each electrode demonstrated unique power output. This indicated that each site

was different from its counterparts within each task when comparing for difference in location alone during the specific task. When each electrode was then compared to itself between tasks, controlling for spatial location while changing the activity of underlying musculature, a significant difference in power output was noted between rest versus low and rest versus high in all subjects. When comparing low versus high, the majority still met statistical significance for variance, demonstrating some loss in differentiation between low and high alone. This would be expected compared to the stark contrast of rest and phonation, but demonstrated an ability to begin to discriminate a difference between low and high phonation in comparison of power alone.

In this study, HD sEMG and power maps were used to identify CT muscle activation preferentially during high-pitch phonation as compared to low pitch in nine of 10 subjects. It is possible that in the remaining patient, the CT muscle was not adequately activated due to subject compliance, or that the CT muscle in some individuals is just no more active in high-pitch versus low-pitch voicing. We were also unable to account for potential contribution from the intralaryngeal musculature during phonatory tasks as a result of the shielding provided by of the thyroid cartilage.

Prior HD sEMG use on the anterior neck has been limited. The first publication of its use focused on assessment of pharyngeal function during swallow of substances of different viscosity [152, 154]. More recently, the same group, in a series of four patients, utilized high-density arrays to evaluate energy distribution during vocal tasks. That study demonstrated identification of muscle activity through evaluation of phonation at contrasting levels of loudness and pitch glides [153] In all of these studies, there remain several limitations of HD sEMG that are inherent to all forms of sEMG.

Impedance of the skin-electrode interface is a known limitation that negatively impacts all forms of sEMG. Close adherence of the electrode grid to the contours of the anterior neck can reduce interference. Even with high-level skin prep and careful electrode placement, this can be a persistent problem. Specificity of muscle location and associated signal acquisition is also

of concern. In this study, this issue is overcome by the broad acquisition area of the array and preferential CT activation during high-pitched phonation. Strap muscle activation was reduced as much as possible, with neutral supported head position during testing paradigms.

The next iteration of this study will include concurrent, fine-wire signal acquisition to further support proof of concept. We also expect that acquiring data on individuals with vagal lesions or high cervical plexus injury may allow for "knockout" conditions to allow for visualization of discrete muscle function and dysfunction. We would emphasize that with the current technology, specifics about wave morphology is not the goal of HD-sEMG. Rather, its use is intended to be complementary to the finite morphological inquiry of nEMG. The primary goal of this study's utilization of HD sEMG is gross identification of muscle activity. Analysis was focused on less-granular aspects of signal such as average power density and energy map generation, but did not include a more detailed analysis of whether signalmorphology is identifiable within the tracings. The authors believe that morphological description may be possible through future iterations of array design and enhanced signal processing, but is beyond current technologic capability of this device version.

Additional limitations of this study include a small sample size (n = 10). Due to technical limitations, concurrent acoustic data were not captured with this subject series. Also, in their current form, the electrode array profile and recording equipment are somewhat cumbersome and would not be practical for extended durations. Future work could exploit recent developments in flexible and stretchable skin-mounted electronics [76] along with advances in scalable fabrication procedures [73, 77], to produce high-density, high-fidelity, and minimally obtrusive electrophysiologic monitoring systems.

Finally, we are excited by the ever-advancing field of machine learning. EMG tracings have prominent potential as a medium of input given their characteristics [46]. We believe it likely that subtle nuances that may be imperceptible to gross visualization within energy maps and/or tracings could be used as a means to train a computer to recognize clinically meaningful

data such as difference between recurrent laryngeal nerve, superior laryngeal nerve, or proximal vagal nerve palsies.

Despite these limitations, HD sEMG represents an exciting new variant of surface electromyography. Its promise lies in the improvement on existing sEMG modalities to compensate for spatial selectivity loss and to globally monitor the anterior neck function through a dense electrode configuration. Potential application includes diagnostic utility to identify laryngeal nerve injury in the absence of endoscopic capabilities, to compare hyperfunctional states, or as a visual biofeedback tool during rehabilitative efforts.

2.1.5 Conclusion

We demonstrate an ability to identify differences of power spectra within an HD sEMG array during rest, lowpitch phonation, and high-pitch phonation across all subjects. Regions of increased power density during high-pitch phonation correspond to those electrodes most likely to be overlying or adjacent to the cricothyroid muscle. Furthermore, review of energy maps generated for each subject afford gross recognition of the cricothyroid activity within the anterior neck. HD sEMG and derived 2D coronal energy map interpretation demonstrates exciting potential for clinical application as a diagnostic tool, therapeutic monitoring device, and visual biofeedback medium of the activity of infrahyoid and cricothyroid muscles.

2.2 Laryngeal High-Density Surface Electromyography for Pitch Differentiation during Phonation

2.2.1 Introduction

Voice is produced by complex interactions between the larynx, pharynx, upper aerodigestive tract, and respiratory system. This system can be adversely affected by a variety of disorders,

leading to voice changes known as dysphonia. In the elderly, rates of dysphonia have been reported from 20-47% [135]. These disorders can be debilitating and isolating [44, 88, 106] as they can negatively impact an individual's ability to communicate in person or via technology, subsequently reducing quality of life. Hyperfunctional disorders (e.g. overuse or misuse of voice muscles) and neurolaryngeal disorders (e.g. Parkinson's disease or nerve injuries) account for 18% of cases in the general population [22], but increases to 50% for those whose voice is essential for their career such as educators [8].

With the advent of novel voice recognition technologies such as Microsoft Cortana and Amazon Alexa, people increasingly use voice to complete tasks and transactions. However, those with voice disorders struggle using such technologies as vocal challenges impact speech patterns and lead to irregularity in pronunciations [95]. Deep learning tools, which power these technologies, struggle to train on such unpredictable patterns limiting model robustness and fitting interpretability. For example, although deep learning applied during vowel pronunciation has shown promise as an objective classification of Parkinsonian voice [134] and two types of speech dysphonia (laryngeal dystonia and muscle tension dysphonia)[114], its accuracy wavers with different diagnoses and continuous speech [9, 86, 130]. Fundamental frequency (F_0) is necessary for acoustic analysis of voice but many dysphonic voices ($\sim$80% of patients) lack a clear F_0 [81]. This reduces the success rate and range of application for dysphonic acoustic models [82, 120]. Another key component missing in voice disorders is the "periodicity" of sound. This, coupled with the lack of a F_0, are setbacks for traditional acoustic analysis.

Similarly, in the clinic, evaluation of treatment outcomes and symptom progression rely on invasive tools to visualize the muscles of the larynx, or through auditory assessment of voice quality by the clinician [12, 13, 100, 129]. The invasive tools such as laryngoscopy, videostroboscopy, and needle electromyography are uncomfortable and require a high degree of technical expertise to use. This limits their application and are therefore not typically used. For assessing improvements in voice quality during treatment, the "trained clinical ear" is considered

the “gold-standard”.

Within the context of classifying speech disorders, solely using voice is complicated as different disorders can share similarly perceived vocal changes and patients often unknowingly overcome changes by activating other muscles. Needle electromyography (nEMG) is the diagnostic tool that can provide objective neurophysiologic data regarding laryngeal dysfunction [14, 47, 55, 57, 61, 63, 64, 66, 78, 111, 141, 148]. However, it does not capture voice quality, is invasive, uncomfortable for patients, and requires a high degree of technical skill for application and interpretation. Therefore it’s likely a speech recognition approach that incorporates an understanding of physiologic changes underlying voice disorders is necessary to overcome these performance limitations.

A non-invasive alternative capable of providing similar neuromuscular information, while simultaneously performing voice acoustic analysis, will enhance our understanding of voice disorders and their assessment. Such a tool would also improve our ability to design speech-recognition technology that can better serve the voice disorder population. Studies have examined surface EMG (sEMG) as a tool to diagnose hyperfunctional voice disorders, but with limited success: none have found a significant difference between normal and pathological patients [75, 123, 139]. Notably, these studies were limited by a low number of sensors, variability in sensor placement, and lack of voice acoustic analysis, common issues among previous sEMG applications for speech [123]. During speech, deep structures move in reference to the surface, potentially confounding the specificity of signal collection. Additionally, when working with the elderly population, skin naturally degrades and becomes fairly loose, leading to something called a “goose-neck”. Therefore, even when a limited number of electrodes are placed correctly over the muscles, over time and with physical activities, the skin will move around. High Density Surface electromyography (HDsEMG), which places a dense array of sEMG electrodes over a target muscle region, combined with voice acoustic analysis, is a non-invasive alternative that may overcome this barrier.

HDsEMG enables acquisition of spatiotemporal patterns allowing interelectrode comparisons to elucidate localized activity correlated to the muscles hypothesized to be active during vocal tasks. In Section 2.1 [25], (Fig. 2.4) we demonstrated our ability to differentiate the cricothyroid (CT) muscle from the strap muscle during phonatory tasks. This section aims to build from Section 2.1 and perform pitch differentiation through HDsEMG.

In literature, it is typically believed that the CT muscle is more active during high pitch phonation and during pitch modulation, whereas the laryngeal adductor muscles (Thyroarytenoid (TA) and Lateral Cricoarytenoid (CA)) are believed to be more active during low pitch phonation [18, 39, 60, 90, 119, 132]. However, due to the placement of the adductors muscles being encapsulated by the larynx, the ability to record them from the surface of the skin is reduced. Nonetheless, we hypothesized low-pitch and high-pitch activity are distinguishable through HDsEMG. Section 2.2 explores how the spatio-temporal data from HDsEMG will allow for pitch differentiation.

2.2.2 Materials and Methods

Eleven healthy adults (four females, seven males, ages 22–51 years, median age 33.4 years) were enrolled. Inclusion criteria was greater than 18 years of age. Exclusion criteria included a history of laryngeal pathology, a Voice Handicap Index (VHI-10) score of >10, subjective dysphonia at the time of recording, previous neck surgery, or neurologic illness.

Ethical approval was granted for this study by the Institutional review board at University of California, San Diego. All participants were seated in an $\sim$80 degree slightly reclined position with the neck in slight extension to minimize strap muscle activity at rest. Surface landmarks, specifically the cricothyroid space and sternal notch, were identified with digital palpation and marked for reference. Skin preparation for all electrode sites included standard alcohol wipes and exfoliating impedance-reduction gel (NuPrep Skin Prep Gel; Weaver and Company, Aurora, CO). Standard electrocardio monitoring electrodes (3 M Red Dot [REF: 2670-5]; 3 M, St. Paul,

MN) composed of silver/silver chloride were modified through circumferential removal of the adhesive patch. Electrodes were organized to create a high-density, 20-channel array (five rows by four columns) at 1.5 cm from each electrode center (electrode diameter = 5 mm). An occlusive transparent dressing (Tegaderm; 3M) was modified with a punch and template. The array was centered on the CT space to ensure electrodes 10 and 11 were overlying the cricothyroid muscles (Fig. 2.1).

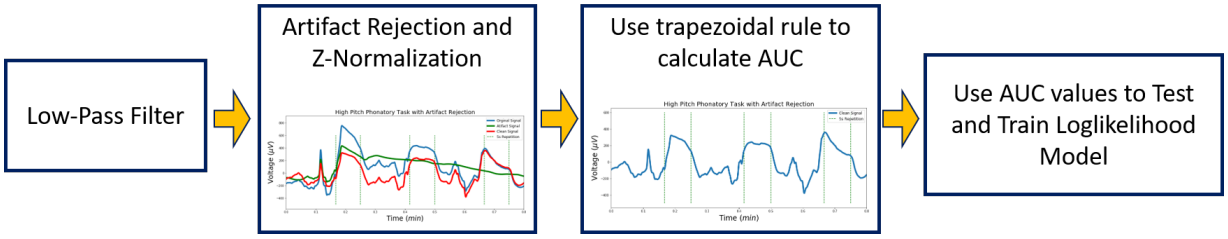

Figure 2.7: Block diagram depicting the signal analysis workflow for pitch classification.

Signal Acquisition

The signal acquisition system, a differential amplifier (Brain Vision Device; BrainVision, LLC, Morrisville, NC) was connected to the electrodes. The reference and ground electrode were placed over the volar surface of the right forearm and left mastoid bone respectively. Impedance values were recorded to be within acceptable values ($<$10 kΩ). For the experiment, an audio prompt was played over a 5-minute period for each recorded task. Tasks included the following: (1) rest, (2) low-pitch phonation, (3) high-pitch phonation. Subjects were able to demonstrate adequate difference between low- and high-pitch as confirmed by the authors with frequency analysis (Audio Frequency Counter; Keuwlsoft, London, United Kingdom). Each task was paced to afford 5 seconds of phonation, on an /i/vowel, followed by 10 seconds of rest. This was repeated for a total of 5 minutes, resulting in 20 recorded intervals of each phonatory task for each participant. EMG data were recorded at a sampling frequency of 500 samples per second using standard EMG recording software (Brain Vision Recorder; BrainVision, LLC). Data was then

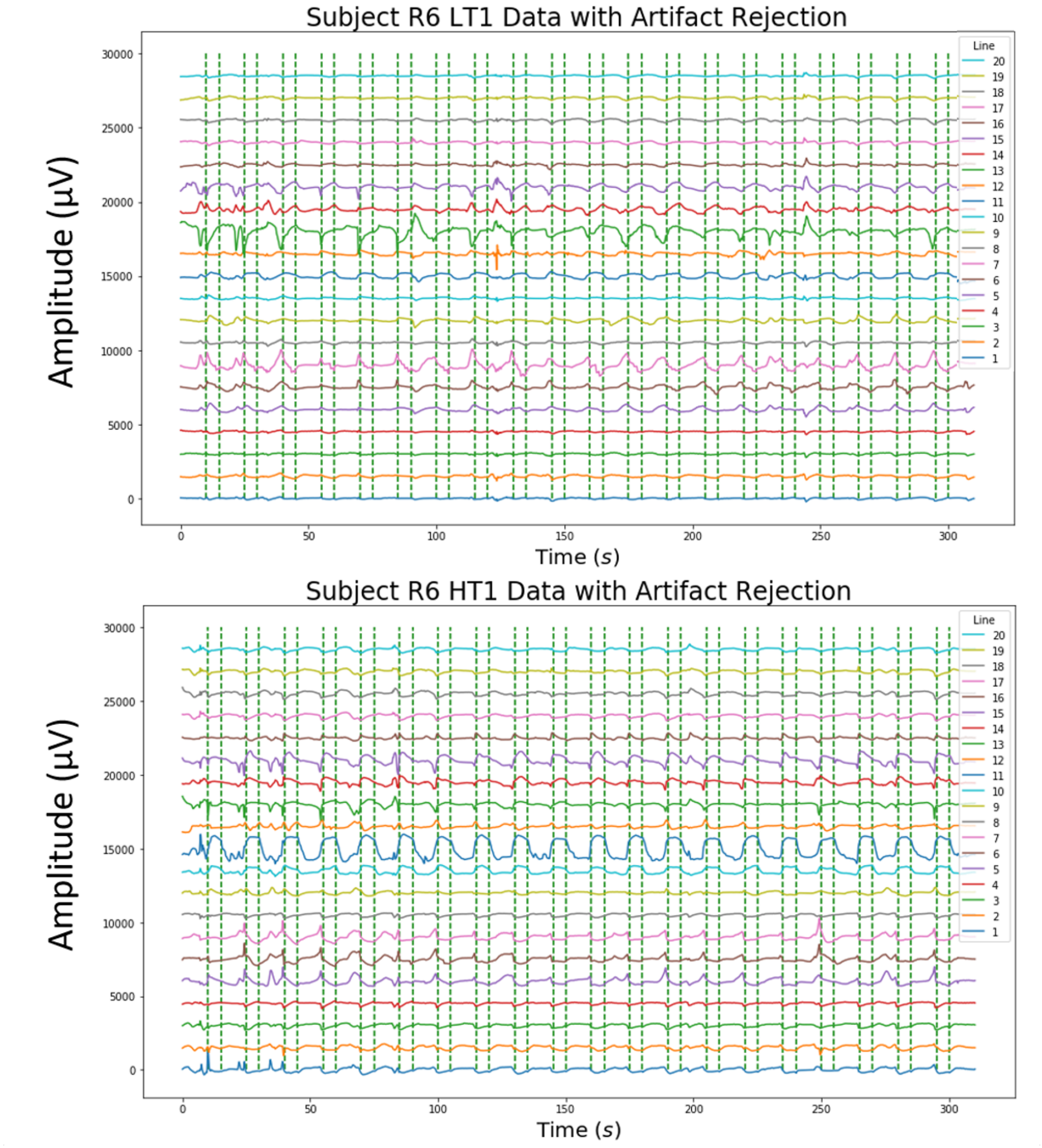

Figure 2.8: Top: Signal tracings depicting all 20 channels during low-pitch phonation. Note the smaller amplitudes of oscillation. Bottom: Signal tracings depicting 20 channels during high-pitch phonation. At this point the oscillatory activity resembles a square wave like quality and has greater amplitudes in certain channels (e.g. Channel 11).

transferred using EMG analysis software (Brain Vision Analyzer, v. 2.1; BrainVision, LLC) to custom software written in Python programming language (Python Software Foundation, Python Language Reference, v. 3.6.3, https://www. python.org/).

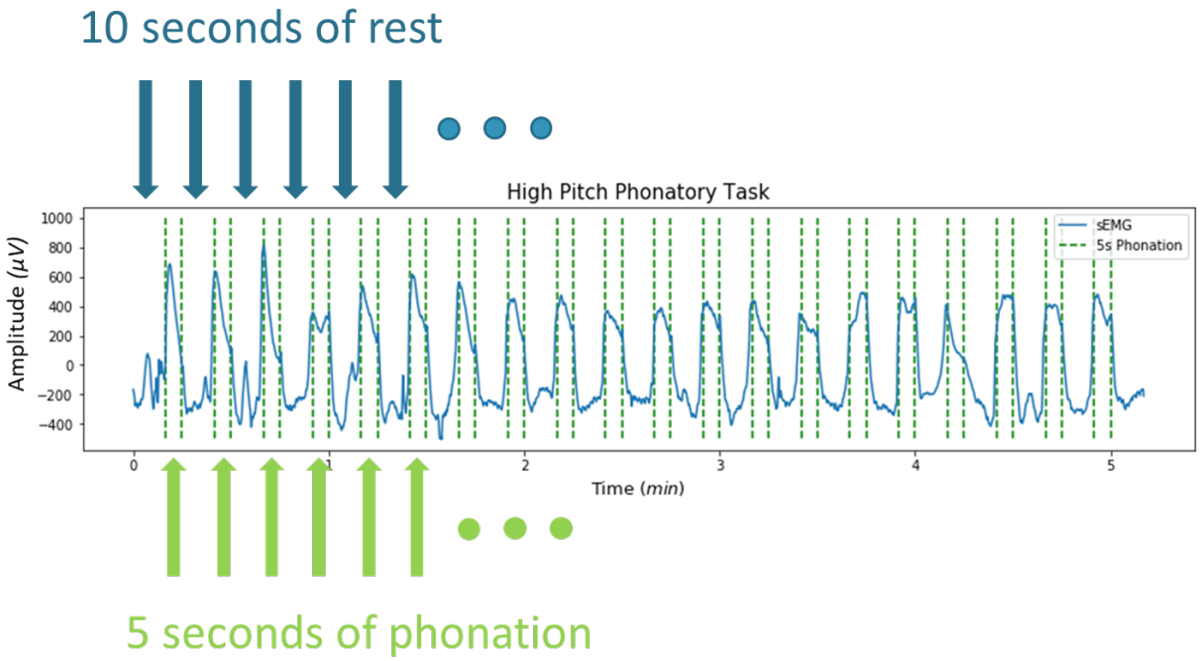

Figure 2.9: A high pitch phonation from a single subject demonstrating the 10 seconds of rest followed by 5 seconds of phonation. This cycle of 15 seconds appeared with greater amplitudes in channels believed to sit over the laryngeal muscles.

Signal Processing

During initial review of the data it was observed that in both low- and high-pitch recordings, periodic repetitions of the 5-second phonation followed by 10-seconds of rest manifested in oscillatory patterns. These oscillatory patterns had greater amplitudes in electrodes nearest the region of the laryngeal muscles (Fig. 2.8). This oscillation had a cycle of 15 s, or an engineered frequency of 0.067 Hz, due to the 10 s of rest followed by 5 s of phonation (Fig. 2.9). Making use of this engineered frequency, a novel method of artifact rejection was implemented [51] after a low-pass butterworth filter with a cut-off frequency of 3 Hz. Following artifact rejection, the

data was standardized using Z-normalization. The trapezoidal rule was used to calculate the area under the curve only during phonation segments. This provided a total of 20 different areas, one per repetition, for each of the 20 electrodes. Fig. 2.7 illustrates a block-diagram of the processing used.

Log Likelihood Ratio Test

This data set contains 11 subjects, 3 phonation tasks (rest, low-pitch, and high-pitch), 20 channels, and 20 repetitions (or trials) of each tasks. Considering how the spatial patterns of activation change during different phonations, we deem it sensible to estimate the distribution of activity in different phonations and use them to perform hypothesis testing. In doing so, this provides an approach to classify the phonation from the spatial patterns extracted from the HDsEMG.

Our logic is as follows: the adductors are active during phonation, while the CT muscles are active during high-pitch phonation and pitch changes. Therefore, during rest, the array should capture minimal activity from both muscle groups. During low-pitch phonation, the array will only capture activity in the adductor region. During high-pitch phonation, we expect both both muscle sets should be active. This is further explained in Fig. 2.10 and Table 2.1.

We can now solve a statistical distribution for each of the tasks: P_{rest} is our distribution for the data at rest, P_{low} is the distribution for the data during low-pitch, and P_{high} is the distribution for the data at high-pitch. Our hypothesis is that in healthy subjects, the distribution of Rest should not equal that of Low nor high, nor should the distribution of low equal that of high. Therefore our null hypothesis is essentially: $P_{rest} = P_{low} = P_{high}$. This is detailed in Table 2.2.

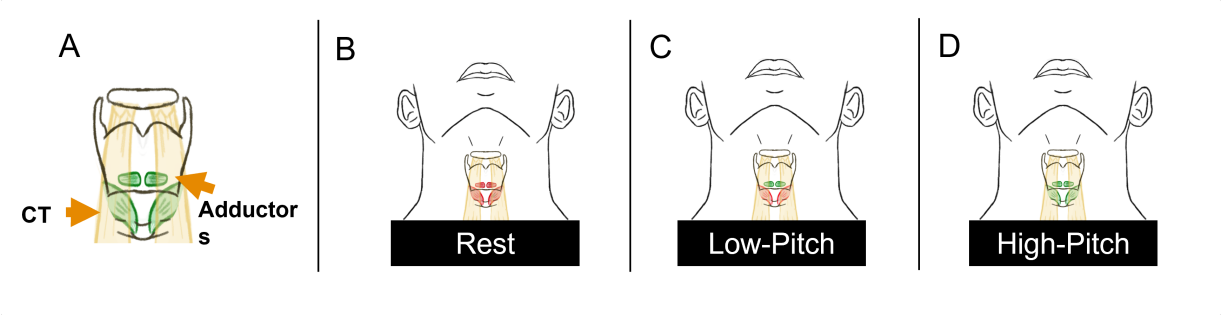

Figure 2.10: A) Illustration of the CT and adductor muscle sets in their respective region. B) During rest, both muscle sets are expected to produce little to no activity, hence the red color. C) During low-pitch activity, the CT muscles (red) are expected to be minimally active, while the adductors (green) are producing high activity. D) During high-pitch activity, both muscle sets are expected to be active (green).

Table 2.1: During rest, energy from both muscle sets should be low. During low-pitch phonation, energy detected from CT region should be low, while adductor region should be high. During high-pitch phonation, energy from both regions should be high.

Phonation:	Rest		Low-Pitch		High-Pitch	
Region:	CT	Adductors	CT	Adductors	CT	Adductors
Energy:	low	low	low	high	high	high

Table 2.2: Binary hypothesis testing for phonation. We have three statistical tests to perform: 1) a test pertaining to $P_{rest} = P_{low}$, 2) a test pertaining to $P_{low} = P_{high}$, and 3) a test pertaining $P_{rest} = P_{high}$.

Test:	$P_{rest} = P_{low}$	$P_{rest} = P_{high}$	$P_{low} = P_{high}$
Expected:	No	No	No

We assume each of the m trials are length 20 vectors and are independent and identically distributed. Therefore we define the channels across the 20 m trials as: $\underline{v}^{(1)}, \underline{v}^{(2)}, \ldots, \underline{v}^{(m)}$, where each $\underline{v}^{(i)}$ is length 20.

After mean-centering, the covariance structure $\hat{K}$ and mean $\underline{\mu}$ is estimated as:

$$\hat{K} = \frac{1}{m} \sum_{i=1}^{m} \underline{v}^{(i)} \underline{v}^{(i)^T}$$

$$\underline{\mu} = \frac{1}{m} \sum_{i=1}^{m} \underline{v}^{(i)}$$

Assume each $\underline{v}^{(i)}$ is a sample from a $\mathcal{N}(\underline{\mu}, \hat{K})$, given by the Gaussian random vector distribution $\mathcal{N}(\underline{\mu}, K)$, given by

$$f(\underline{u}) = \left(\frac{1}{\sqrt{2\pi}}\right)^{20} \frac{1}{\sqrt{\det(K)}} \exp\left(-\frac{1}{2}(\underline{u}-\underline{\mu})^T K^{-1}(\underline{u}-\underline{\mu})\right).$$

We then use the above procedure to compute $\underline{\mu}_{\text{high}}, \hat{K}_{\text{high}}$ and similarly $\underline{\mu}_{\text{low}}, \hat{K}_{\text{low}}$.
Using the Leave-One-Out Cross-Validation method, we train with 19 phonation trials, and leave one out to test with. We then shift through these combinations so that each trials is left out once to test. Using the vectors that were left out, we now test the log likelihood ratio test (*LLR*):

$$\begin{aligned}
LLR(u) &= \log f_{\text{high}}(\underline{u}) / f_{\text{low}}(\underline{u}) \\
&= \log\left(\frac{\left(\frac{1}{\sqrt{2\pi}}\right)^{20} \frac{1}{\sqrt{\det(\hat{K}_{\text{high}})}} \exp\left(-\frac{1}{2}(\underline{u}-\underline{\mu}_{\text{high}})^T \hat{K}_{\text{high}}^{-1}(\underline{u}-\underline{\mu}_{\text{high}})\right)}{\left(\frac{1}{\sqrt{2\pi}}\right)^{20} \frac{1}{\sqrt{\det(\hat{K}_{\text{low}})}} \exp\left(-\frac{1}{2}(\underline{u}-\underline{\mu}_{\text{low}})^T \hat{K}_{\text{low}}^{-1}(\underline{u}-\underline{\mu}_{\text{low}})\right)}\right) \\
&= \frac{1}{2}\log\det(\hat{K}_{\text{low}}) - \frac{1}{2}\log\det(\hat{K}_{\text{high}}) - \frac{1}{2}(\underline{u}-\underline{\mu}_{\text{high}})^T \hat{K}_{\text{high}}^{-1}(\underline{u}-\underline{\mu}_{\text{high}}) \\
&+ \frac{1}{2}(\underline{u}-\underline{\mu}_{\text{low}})^T \hat{K}_{\text{low}}^{-1}(\underline{u}-\underline{\mu}_{\text{low}})
\end{aligned}$$

The LLR values were then put through a Receiver Operating Character curve.

2.2.3 Results

Qualitative inspection of the data demonstrates activity specific to channels, during certain tasks. Additionally the amplitudes increased when nearest the laryngeal region. Though channels 10 and 11 were placed over the symmetric CT muscles, these channels did not always record the greatest amplitudes. These shifts varied between subjects with no discernible pattern between tasks.

Regarding pitch differentiation, two different ROC curves were obtained. The first used

the features from the entirety of the multi-electrode array as shown by Fig.2.11, Top. This first ROC demonstrated a distinction between low- and high-pitch phonation in all subjects with an area under the curve (AUC) of 0.97. The second ROC (Fig.2.11, Bottom) was calculated using only two electrodes over the general CT location to replicate what one might do in traditional sEMG [123, 126]. This ROC curve demonstrated greater variation between subjects with a lowered AUC of 0.77.

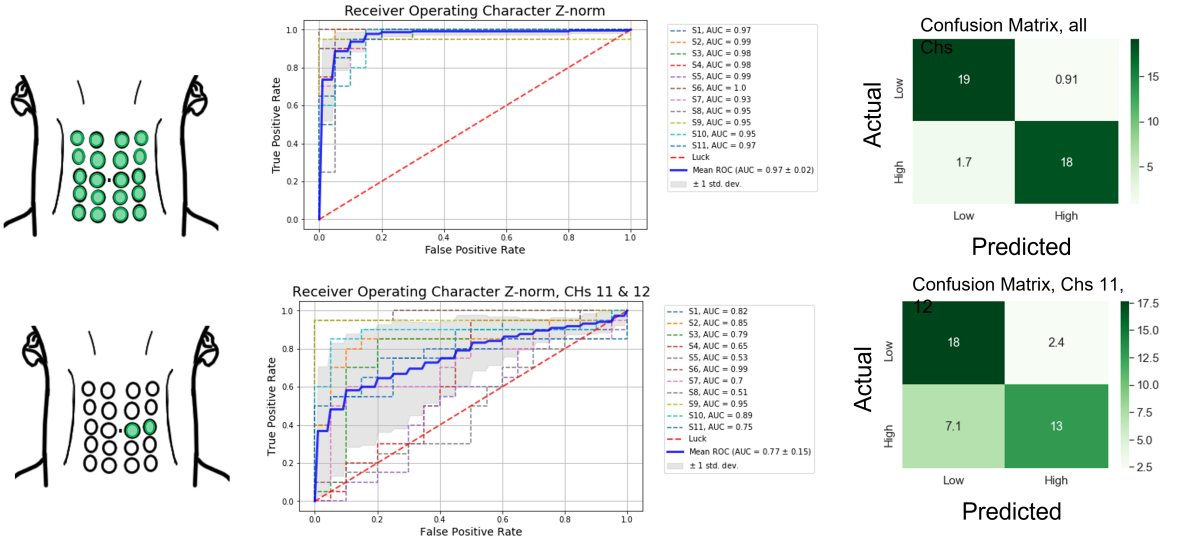

Figure 2.11: Top: ROC curve with an AUC = 0.94, obtained using all the channels. Bottom: ROC curve obtained with only two channels, replicating one of the traditional configurations of laryngeal sEMG.

2.2.4 Discussion

There is no doubt regarding the complexity of voice disorders which has limited our ability to implement non-invasive tools for assessment and monitoring of treatment progression over time. When studying such a dynamic region, neuromusuclar information is needed to distinguish between disorders and phonation activity. Though nEMG provides the needed muscular information, it is uncomfortable and requires a high-level of expertise to administer.

Acoustic analysis of dysphonic voice has been promising, but its applications are limited to specific disorders and does not provide muscular information. A non-invasive alternative capable of providing similar neuromuscular information, while simultaneously performing voice acoustic analysis, will enhance our understanding of voice disorders and improve their assessment. Such a tool would not only improve assessment of voice disorders, but also our ability to design speech recognition technology that can better serve the voice disorder population.

This study sought to investigate the ability of HDsEMG to classify low vs high-pitch phonation. The CT and adductor muscles have been highlighted in this project due to their roles in high and low-pitch phonation. Additionally, the CT muscles' relative closeness to the skin, and increased recruitment during high-pitch phonation, allows its isolation during high-pitch phonation. Subjects were asked to maintain a target pitch above 300 Hz for males, and 400 Hz females to ensure reliable CT activation [69, 80]. Crosstalk contamination by muscles associated with swallowing was minimized since concurrent swallowing would not be expected within each central 5-second phonatory segment. Large neck movements were limited by volitional control and cervical support provided by the examination chair.

The multi-electrode array (MEA) was centered such that electrodes 10 and 11 rested over the CT muscle. Despite this placement, these electrodes did not always record the greatest amplitudes during phonation. The largest recorded amplitude shifted without discernible pattern in both low and high-pitch phonation. This confirms inconsistencies associated with traditional laryngeal sEMG. Without a MEA, the traditional set-up comprised of a few electrodes runs the risk of missing the phonation activty due to natural movements of the larynx, neck, and crosstalk contamination. HDsEMG enables the use of spatio-temporal data to locate the muscles of interest during tasks in which only those muscles are active.

This is further exemplified in the ROC results. The HDsEMG ROC curve had an AUC of 0.97. Across subjects, the AUC range was 0.93 - 1.0. The ROC from two channels, mimicking traditional sEMG, had an AUC of 0.77. There was greater subject variability in AUCs, ranging

from 0.51 - 0.99. An interesting observation is that Subjects 4, 7, 8 had low SNR and low amplitudes during phonation across tasks. Nonetheless, through HDsEMG, their ROC-AUCs were 0.99, 0.93, and 0.95 respectively. However their AUC dropped to 0.65, 0.7. and 0.51 respectively in the 2-ch ROC. This leads us to believe signal quality has a greater impact in traditional sEMG. It would be interesting to observe how SNR impacts classification in a future study.

Limitations of this study include a small sample size (n = 11) and lack of concurrent acoustic data. Additionally, in its current form, the electrode array profile and recording equipment are somewhat cumbersome and would not be practical for extended durations or ambulatory settings. Future work could exploit recent developments in flexible and stretchable skin-mounted electronics [76] along with advances in scalable fabrication procedures [73, 77], to produce high-density, high-fidelity, and minimally obtrusive electrophysiologic monitoring systems. This is explored in its infancy during Section 3.2.

Despite these limitations, HDsEMG represents an exciting new variant of surface electromyography. Its promise lies in the improvement on existing sEMG modalities to compensate for spatial selectivity loss, signal quality, and to globally monitor the anterior neck function through a MEA. Potential application includes assessment utility to identify laryngeal nerve injury in the absence of endoscopic capabilities, to compare hyperfunctional states, or as a visual biofeedback tool during rehabilitative efforts.

2.2.5 Conclusion

Spatiotemporal data from HDsEMG allows us to classify low- versus high-pitch phonation, across all subjects, an ability that is reduced without the MEA. In the future, combining these space and time-varying HDsEMG features with voice recording waveforms will allow binary or multi-class characterization of vocal dysfunction across voice disorders. In addition, features learned in pathology characterization may be utilized, through transfer learning, to improve

existing voice recognition paradigms (e.g. Apple Siri, Google Alexa). Transfer learning involves the use of information learned in one classification task for a separate but related task, typically enabling breakthrough performance [6, 70] and in this case addressing the vulnerabilities in existing voice recognition technology. Furthermore, state space modeling can characterize relationships between features of voice recordings and the underlying latent state, HDsEMG [79, 113]. With these relationships, voice classification tasks can be performed even when only voice recording data is available, thus widely extending the reach of our applied methodology.

2.3 Acknowledgments

The authors thank Dr. Armen Gharibans, Dr. Dae Y. Kang and Dr. Avram Hecht for their initial work in this project regarding surface EMG of the anterior neck.

Section 2.1, in full, is a reprint of the material as it appears in The Laryngoscope 2019. Bracken, David J. ; Ornelas, Gladys; Coleman, Todd P.; Weissbrod, Philip A., The Laryngoscope 2019. The dissertation author was the second author of this paper.

Section 2.2, in full, is unpublished material and is in preparation for submission. Ornelas, Gladys; Bracken, David J.; Weissbrod, Philip A.; Coleman, Todd P., The dissertation author was the first author of this paper.

Chapter 3

Swallow

3.1 High Density Surface Electromyography: Introductory Applications to Dry Swallow Analysis

3.1.1 Introduction

Swallowing is a complex neuromuscular process critical to successful oral intake. Impaired swallow function has a physiologic and psychologic impact on health and quality of life with a lifetime incidence of nearly 38% [107]. In the U.S.A. alone, nearly 10 million adults reported symptoms of dysphagia, of whom only one-fifth sought care and only 37% were given a diagnosis [99, 107]. The most common causes of dysphagia occur as a consequence of stroke (11.2%), neurologic disorders (7.2%), and head and neck cancer (4.9%) [21].

Current oropharyngeal swallow function assessment involves behavioral and instrumented examination. Common modalities of objective swallow assessments include endoscopic analysis and contrast fluoroscopy, which can be considered a standard method due their prevalence in clinical care and diagnosis. Additional procedures that provide a quantitative approach to physiological inquiry during swallowing include pharyngeal manometry, electromyography,

electroglottography, and mechanomyography [12, 29, 34, 116, 138] While these distinct methods provide a broad collection of tools for diagnostic inquiry, only electromyography provides a direct measurement of neuromuscular activity and ensures immediate access to muscular health data. This is vital to move forward with the diagnosis and treatment of disorders such as dysphagia.

Electromyography applications towards swallowing assessment has an increasing body of literature investigating both invasive and noninvasive methods of measurement [10, 35, 41, 107, 136]. Needle electromyography (nEMG), a method commonly used to analyze laryngeal function remains the gold-standard for quantitative EMG due to its ability to identify discrete motor unit action potentials and waveform morphology. This specificity of signal is achieved through small detection volumes and directed anatomic placement. Therefore the recorded electromyographic measurement is thought to specifically correlate to the intended muscle. Despite this specificity, it is not a practical method for widespread clinical use in swallow evaluation due to the challenges of analyzing concurrent target muscles required for a comprehensive study, the technical challenges with accurate needle placement, the discomfort experienced by patients on insertion, and the inherent limited duration of inquiry secondary to its indwelling nature. Surface electromyography (sEMG) of the neck during swallow introduces a non-invasive method and has been typically performed via bipolar electrode configurations. This modality yields an output of myoelectric signal intensity, which is the sum of individual source voltages captured beneath the electrode on the area of interest. The non- specific summative nature of traditional bipolar sEMG, compounded by the neuromuscular density of the anterior neck, is a concern. Overlapping of relatively small muscle bodies coupled with the dynamic relationship of muscles-to-skin surface change during swallow can increase the likelihood of inadvertently recording adjacent muscle activity. This phenomena known as “cross talk” can confound the interpretation of individual signal sources. A more comprehensive consideration of limitations include: (1) impedance of the skin-electrode interface, (2) distance between the neuromuscular junction (NMJ) and electrodes, (3) lack of specificity due to interposed or neighboring active muscles, (4) dynamic changes

between superficial and deep structures during activation and rest, and (5) limitations in the ability to describe wave morphology [37].

Here we utilize 20-channel HDsEMG arrays to evaluate the anterior neck during dry swallow tasks. This study aims to show proof-of-concept that HDsEMG can be utilized to capture electromyographic swallow data and provide additional spatiotemporal information not available with current bipolar configurations or non-EMG approaches (i.e. behavioral evaluation, endoscopy, fluoroscopy, or high resolution manometry). Along with diagnostic applications and further development of the proposed tool to measure anterior neck muscular function during swallowing, we may be able to use this tool for rehabilitative and biofeedback applications. This manuscript represents the first step in evaluating HDsEMG as a new method in diagnosis and management of dysphagia.

3.1.2 Materials and Methods

Ethical approval for this study was granted by the Institutional Review Board at the University of California, San Diego (IRB # 161477). Ten healthy adults (3 females, 7 males, ages 22 -51yrs, median 33.4 yrs) were enrolled. Inclusion criteria was greater than 18 years of age. Exclusion criteria included a history of dysphagia, dysphonia, previous neck surgery, or neurological illness.

To minimize strap muscle activity at rest, all participants were seated in a semi-reclined position ($\sim$ 20 degrees from vertical) with the neck in slight extension and supported. The cricothyroid (CT) space was identified with digital palpation and marked as a reference point for grid orientation. Skin preparation with standard alcohol wipes and exfoliating impedance-reduction gel (NuPrep Skin Prep Gel, Weaver and Company, Aurora, Colorado, USA) was performed across the entire anterior neck.

A high density 20- channel array was fabricated from standard electro-cardiac monitoring electrodes (3M Red Dot (REF: 2670-5), St. Paul, Minnesota, USA) composed of silver/silver

chloride. Electrodes were organized into 5 rows and 4 columns at 1.5 cm from each electrode center (electrode diameter 1 cm). An occlusive transparent dressing (Tegaderm™, 3M Medical, Global) was modified with a punch and template to allow for exact confirmation to each electrode. The array was centered on the CT space, with electrodes 10 and 11 overlying the cricothyroid space. Electrodes 1-4 approximated the hyoid bone (Fig. 3.1).

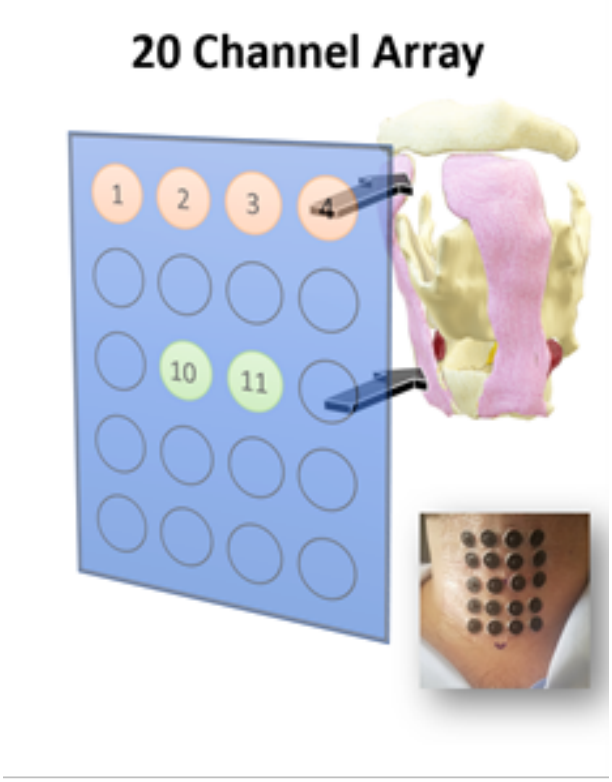

Figure 3.1: 20-channel array oriented such that its center overlayed the cricothyroid space. Row 1 (electrodes 1-4) cranially oriented and juxtaposed to the hyoid bone.

Signal Acquisition

The HDsEMG array used a differential amplifier (Brain Vision Device, BrainVision, LLC, Morrisville, North Carolina, USA) sampling at 500 Hz with a reference and ground electrode placed over the volar surface of the right forearm and left mastoid process respectively. Subjects were asked to perform self-paced dry swallows over an interval of 1 minute. Subjects were

careful to limit mouth opening, maintain neutral neck positioning, and refrain from phonation during recording. The number of swallow events and time of occurrence was noted grossly by observation of the subject. Recorded data was transferred to EMG analysis software (Brain Vision Analyzer, v. 2.1, BrainVision, LLC, Morrisville, North Carolina, USA) and subsequently to custom software written in Python programming language (Python Software Foundation, Python Language Reference, v. 3.7.1).

Signal Processing

Raw data was processed with a workflow composed of four key steps. First, a high-pass filter (20 Hz) and notch-filter (60 Hz) were applied to remove cardiac myoelectric activity and powerline noise respectively. The root mean square (RMS) of recorded voltages was obtained and a linear envelope was generated using a window size of 1 ms and a cutoff frequency of 3 Hz [144]. A threshold of EMG voltage deflection of greater than 3 standard deviations above a rest baseline was used to define onset. Offset was defined as a greater than 3 standard deviation return to rest baseline within each channel [10, 29, 30]. Onset and offset points were used to separate RMS voltage data into rest and swallowing epochs. Data analysis was only performed on the resulting swallowing epochs for each channel.

Data output included average and maximum voltage amplitudes (microvolts) and signal duration (milliseconds) at each electrode within the array. Statistical analysis of voltage data included one and two factor Analysis of Variance (ANOVA) and Pearson's coefficients of correlation. Statistical analysis was performed using Microsoft Office Excel (Microsoft Corporation, Redmond, WA, USA) and GraphPad Prism (GraphPad, v7.04, San Diego, CA, USA). Statistical analysis was performed in 3 groups: (1) Comparison of each electrode to all other single electrodes within the array for each subject. This was performed across successive swallow events for the purpose of assessing signal heterogeneity. (2) Comparison of row 1 (electrodes 1 – 4) and row 3 (electrodes 9-12) for the purpose of cranial – caudal axis comparison. (3) Comparison of

columns 1 and 2 against columns 3 and 4 as left and right columns respectively for the purpose of evaluation of signal symmetry in normal subjects.

For the purpose of generating activation "energy maps", power density was calculated through analysis of RMS change over the cross sectional area with respect to time. A blue-to-red color gradient was assigned to power density values during animation of signal data for ease of muscle activity visualization.

3.1.3 Results

A 20-Channel HDsEMG array provided EMG data of swallowing events in 3 categories of inquiry: (1) signal heterogeneity within subjects across the same electrodes; (2) cranial-caudal axis variation; (3) and signal symmetry within the swallows of the same subject. These conditions were chosen for the purpose of: (1) monitoring instances of EMG change or variation within the same swallow task with respect to the "perspective" of each electrode's detection volume over consecutive swallows; (2) to monitor possible association of suprahyoid musculature to pharyngeal swallow; and (3) to evaluate the presence of signal similarity with respect to laterality in normal individuals.

Voltage outputs were not statistically significant for variance within subjects over consecutive swallows. The same recorded electrode for the same subject was similar and relatively homogeneous ($p=0.25$). Evaluation of voltage difference with respect to cranial and caudal signal acquisition perspective amidst the array used row 1 (electrodes 1-4) and row 3 (electrodes 9-12). In 70% (7/10) of subject swallows, cranially oriented row 1 demonstrated greater voltage average amplitude for each swallow compared to the caudally oriented row 3 during each swallow event (Fig. 3.2). The correlation of left and right laterality, or symmetry of each array, was analyzed comparing 20-channel electrode columns 1-2 vs 3-4. Pearson's correlation (values of $r=0.88$-0.97) demonstrated symmetry in 90% of subjects across all swallows (Fig. 3.3). While not a primary output, the comparison of the onset/offset time periods as defined electromyographically for each

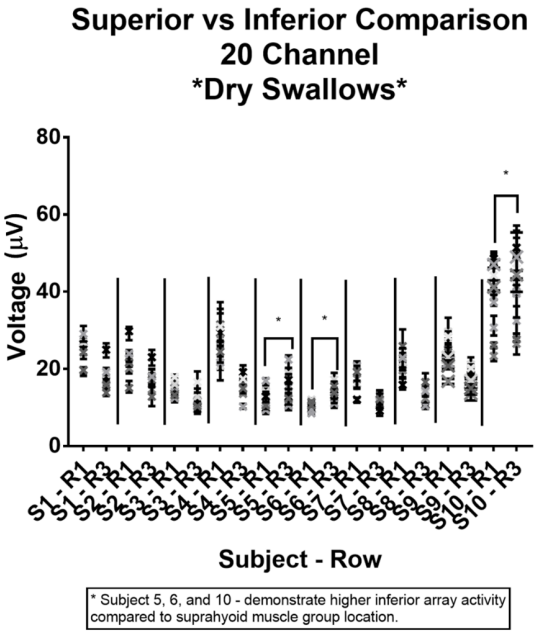

Figure 3.2: Cranio-caudal comparison of average voltage output for each electrode across swallows for each subject. Comparison of row 1 (electrodes 1-4) and row 3 (electrodes 9-12).

dry swallow showed a mean duration of 2.1 - 5.2 seconds (average 3.27 seconds). Statistically significant outliers of single event durations were noted in Subjects 4 and 8 (standard deviation of 3.7 and 2.6) with maximum durations of 13 and 8 seconds respectively. The average swallow duration for subjects 4 and 8 though were 4.9 and 3.5 seconds respectively.

Finally, voltage data over time was used to generate dynamic RMS power energy maps to allow for gross visualization of muscle activity within the array (Fig. 3.4). These energy maps were used to grossly visualize the area of maximal muscle activity and transit across the array during single swallow events.

3.1.4 Discussion

Swallowing is a highly coordinated neuromuscular process involving the interplay of multiple muscle groups concurrently and sequentially. The extent and order of neuromuscular

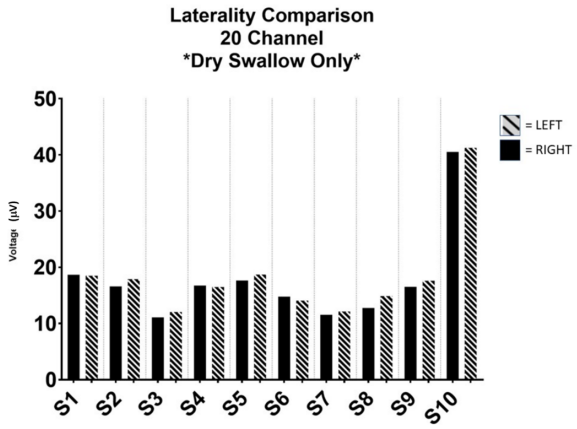

Figure 3.3: Comparison of laterality, or symmetry of muscle activity, for all swallows by subject. Right represents columns 1-2 and Left represents columns 3-4.

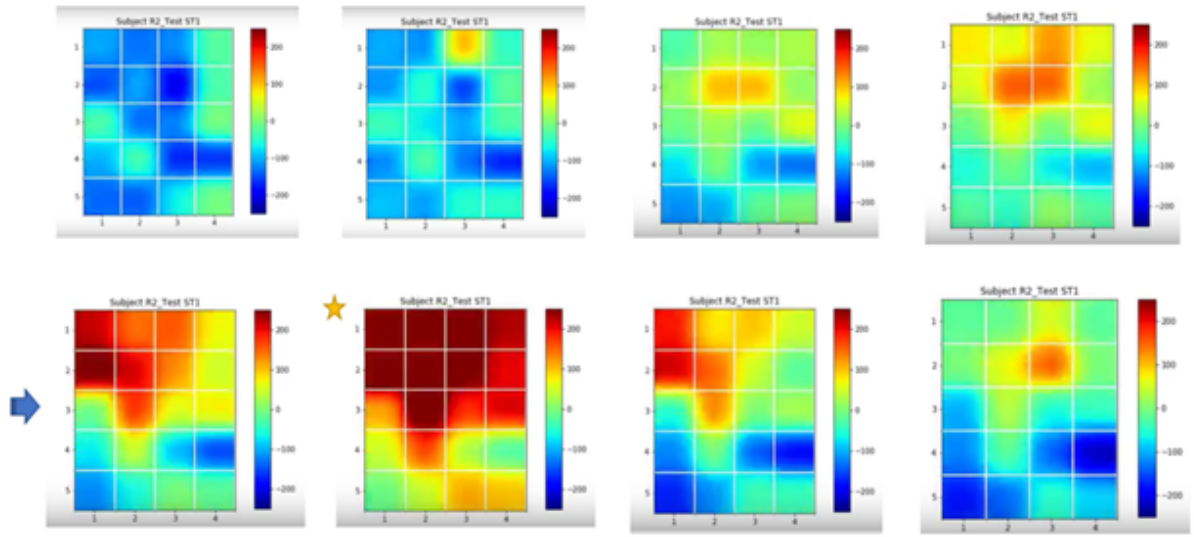

Figure 3.4: Power density, or energy map, for subject 2 over a single swallow events. Note cranial dominant array activity. Signal originates cranially and centrally with lateral spread and the resolution back towards cranially oriented center.

activation is commonly altered by neurologic illness, stroke, iatrogenic manipulation or injury during surgery. sEMG offers a convenient, non-invasive, and time-efficient manner of studying such changes in neuromuscular activation. To date, it has been used to study hyper-functional

laryngeal disorders, dysphagia, applied as a biofeedback tool [56, 59, 124, 125] and allowed comparisons of laterality, symmetry, and pattern of neuromuscular activation. Despite such advancements, sEMG has struggled to secure a foothold as a robust method for swallow analysis [41, 71, 117, 140, 142] due to a distinct set of limitations involving skin-to-electrode impedance, risk of signal cross-talk, and diminished electrode-to-target muscle specificity.

In an effort to address some of the limitations of bipolar sEMG, high-density surface electromyography (HDsEMG) uses an array of electrodes, each with their own small detection volumes that provide simultaneous recording of muscle activity in a broad region of interest. Each electrode's discrete detection volume is defined by interelectrode distance and electrode surface area. Consider that as detection volume decreases, the summative contents of each electromyographic signal become more individualized, allowing for interelectrode comparison. This ability to compare in a spatial and temporally oriented means across fixed points with known orientation to surface anatomy demonstrates potential benefit of high-density configurations. Given the dynamic nature and muscle recruitment patterns during normal physiologic swallow, increased spatial resolution is critical towards accurate interpretation of muscle activation during swallow events. While limitations of skin impedance, NMJ depth, and reliable wave morphology detection remain, HDsEMG provides a means for differentiating muscle activation in a dense neuromuscular region through increased spatial resolution of detection [25, 37, 152, 153, 154, 156].

In this study, ten patients performed sequential self-paced dry swallows with 20-channel HDsEMG data revealing consistent signal activity within subjects, greater cranial vs caudal array activity, and symmetric laterality of activity, and similar swallow durations. Each dry salivary swallow resulted in similar voltage outputs from the perspective of each discrete detection volume for the same subject. Greater degrees of variation were noted when comparing the same electrode across subjects. As each individual is hypothesized to have variable NMJ density and anatomical correlation, inter-subject variation would be expected.

Greater voltage activity in the upper rows was seen across the majority of swallows. We hypothesize this to correlate to suprahyoid muscle activity. This muscle group, known to be prominent during hyolaryngeal elevation, is located above row 1 and more superficially oriented compared to other muscles groups activated during swallow. It may be a dominant contributing EMG source within its summative detection volume.

Symmetry analysis, or the ability to discriminate sEMG signal laterality, is critical to clinical applications given the common occurrence and functional effects of pharyngeal or strap muscle paralysis, or paresis, in disorders of swallowing. All subject trials demonstrated symmetry in physiologic activity grossly. A confluence of factors may lead to isolated asymmetry over a series of repetitions. Asymmetry in isolated occurrences may represent a normal variant within the healthy adult population and should be considered in further studies. Pathologic conditions, such as unilateral high vagal injury and hemispheric stroke, would provide unique opportunities to enhance the system's ability to identify clinically meaningful laterality differences. These findings are consistent with prior work noting high concordance of muscle activation with regards to laterality [154].

Previous authors have defined an electrophysiologic single swallow event by a myoelectric onset and offset. Swallow onset was defined as an increase in voltage greater than a predetermined standard deviation (SD) above rest. Offset was defined by return to resting voltage [30, 34, 138]. The durations between onset and offset during swallow, defined by the method stated above, are longer than standard accepted physiologic norms of approximately less than 1 second. The physiologic duration of the pharyngeal phase of swallow is readily visualized by current modalities of radiographic or endoscopic inquiry. In our study the onset and offset of swallow was defined by deflection in an electromyographic baseline, likely to occur both in preparation and after the event itself, and may augment the duration defined through gross visualization alone. Two subjects demonstrated a greater variability in duration between their own swallow events, indicating that this may indeed vary to a greater degree within a larger patient cohort or even for an individual

from swallow to swallow.

HDsEMG also lends itself towards intuitive visualization via power density energy maps. Points of maximum muscle activity were indicated by a "red-out" of the spatial grid (Fig. 3.4). Central tendency of myoelectric source and gross symmetry were noted. Further analysis techniques are needed for evaluation of power density maps for this purpose. This would include determination of ideal sampling frequency, definition of critical thresholds and correlation to that currently reported in bipolar systems. It is likely that these "maps" may be the most clinically relevant components when considering HDsEMG as a tool for visual biofeedback and therapeutic patient driven application.

Limitations of this study include its sample size, variable swallow bolus volume, and variable subject self-pacing of swallows. Additionally, the array may have presented conformational challenges in approximating the neck surface contour, which may yield outliers or artifactually decreased amplitudes secondary to variable electrode-to-skin surface interface. While HDsEMG lends itself to collection of signals over a defined region, it needs to be compared to needle EMG in order to properly understand the source of recorded myoelectric data and its representation of the muscles of interest.

HDsEMG array profiles and recording equipment remain cumbersome and not practical for extended recording periods in its present form. Future work will exploit recent developments in flexible and stretchable skin-mounted electronics along with advances in scalable fabrication procedures to produce high-density, high-fidelity, and minimally obtrusive electrophysiologic monitoring systems [73, 76]. In addition, further investigation towards the application of statistical learning techniques such as EMG signal decomposition, pattern recognition, or wavelet analysis for higher degrees of quantitative comparison and attempts to approximate the strengths of needle EMG are critical [36, 48, 118, 147]. Signal quantity and pattern characteristics of HDsEMG output lend themselves well to concepts in machine learning and are a focus of additional ongoing work.

3.1.5 Conclusion

HDsEMG represents an exciting new variant of surface electromyography. Its promise lies in improving existing sEMG modalities to compensate for spatial selectivity loss and to globally monitor the anterior neck for function through a dense electrode configuration over broader surface areas. Here, we demonstrate an ability to identify differences in voltage within an HDsEMG array during swallow events. Cranial aspects of the array demonstrated higher voltages consistent with the expected region of suprahyoid muscle activation during pharyngeal swallow. Symmetry is commonly seen between the left and right of the array in healthy subjects. Looking forward, there are numerous potential applications ranging from neuro-diagnostics to biofeedback monitoring for rehabilitative tasks. With the addition of machine learning paradigms, this technology holds significant promise to develop into a powerful clinical and diagnostic tool [36, 48, 118].

3.2 From Thin Liquids to Dry Solids: Classification of Complex Swallows using Laryngeal High-Density Surface Electromyography

3.2.1 Introduction

Swallow dysfunction may be caused by a variety of issues such as aging, neurologic disorders, stroke, and cancer [21]. Issues in swallowing have both a physiologic and psychologic impact on the patient's health and quality of life [99, 107]. Dysphagia, or the diminished ability to swallow, restricts the patient to certain types of food and methods of eating, limiting their ability to lead independent and socially active lives. Additionally, dysphagia may become detrimental or life threatening to patients as they are at risk for aspiration, pneumonia, and malnutrition

[40, 143]. In the elderly population, the rate of dysphagia has been reported from 14-38% [135]. In children, prevalence of dysphagia is difficult to estimate due to the lack of standardized assessment procedures and variable definition of dysphagia for children [12].

Management of dysphagia relies on medication, surgery, or Botox injections to treat the underlying cause of uncoordinated muscles. In a subset of the population, speech-language pathologists often use swallow therapy as a common treatment option for dysphagia. Patients are often asked to perform certain maneuvers or exercises to assist their swallowing. Others may swallow substances of various consistencies that range from thin liquids to extremely thick liquids. However, it lacks a non-invasive and easy-to-use objective measurement of treatment success or patient compliance. Additionally, its effectiveness seems to vary with the underlying disorder and types of therapeutic treatments [11, 19, 20, 31, 53, 72, 97, 108].

Biofeedback may be a recommended addition for swallow therapy to aid with error-based learning [20]. A typical biofeedback tool is surface electromyography in which the subject is asked to reach a set amplitude threshold during their swallow. The idea is that the signal's amplitude will increase with increased muscle contraction force [20, 33]. However, its efficacy remains an ongoing question due to limitations in existing objective and non-invasive monitoring modalities. Additionally, traditional sEMG is not capable of distinguishing individual muscles [34, 37] and is faced with additional limitations. There is a lack of standardization in terms of the number electrodes used, the location where electrodes are placed, and the mechanism of data acquisition and processing [23, 26, 30, 91, 122, 136, 151]. These issues prevent sEMG from addressing the unmet need for a non-invasive system capable of objectively monitoring patient compliance and treatment progression over time. Such information would allow patients and clinicians to better understand the swallow disorder and better treatment options. A solution may lie with high-density surface electromyography (HDsEMG) through spatiotemporal data analysis.

HDsEMG has been used extensively in other physiological regions for effective muscle disambiguation [37]. It is important to note that with the neck, we are dealing a multiple

overlapping muscles, of various sizes, which are dynamically contracting at different points in time to allow the food bolus to travel from the mouth to the esophagus. In voice and swallow, ongoing studies demonstrate how HDsEMG, coupled with advance signal processing, can observe the dynamic interplay of laryngeal muscles and differentiate task-specific activities during phonation and swallow [25, 152, 153, 154, 155, 158, 157]. HDsEMG continues to demonstrate its value in monitoring and decoding complex muscular activities without patient discomfort and pain. Advances in signal processing and statistical learning of physiological signals have improved their health care applications [43, 1, 101, 2, 3, 104, 94, 7, 145] leading us to believe HDsEMG will be an effective monitoring option for swallow disorders in the clinic and therapy.

This paper explores the use of HDsEMG to classify swallows of various textures and complexities ranging from thin liquids to dry solids. If successful, HDsEMG holds the capacity to serve as an objective measurement of patient compliance. Through HDsEMG, we will be able to monitor changes over long periods of time and understand a patient's progress with physical therapy. HDsEMG may pave the way for future diagnostic applications such as a secondary evaluation in emergency rooms following a stroke for unilateral nerve paralysis and vocal fold paralysis.

3.2.2 Materials and Methods

Ethical approval for this study was granted by the Institutional Review Board at the University of California, San Diego (IRB # 161477). Three healthy adults (1 female, 2 males, ages 20 - 40 yrs, average 25.7 yrs) were enrolled. Inclusion criteria was greater than 18 years of age. Exclusion criteria included a history of dysphagia, dysphonia, previous neck surgery, or neurological illness.

To minimize strap muscle activity at rest, all participants were seated in a semi-reclined position ($\sim$10 degrees from vertical) with the neck in slight extension and supported. The cricothyroid (CT) space was identified with digital palpation and marked as a reference point for grid

orientation. Skin preparation with standard alcohol wipes and exfoliating impedance-reduction gel (NuPrep Skin Prep Gel, Weaver and Company, Aurora, Colorado, USA) was performed across the entire anterior neck. A prefabricated array (TMSi, HD sEMG Grid, International, the Netherlands) with 64 channels (8 rows by 8 columns) and an interelectrode distance of 8.5 mm (electrode diameter 4mm) was centered on the CT space, with electrodes 36 and 37 overlying the cricothyroid space, and electrodes 1-8 overlying the hyoid bone. Fig. 3.5 demonstrates the array on a subject.

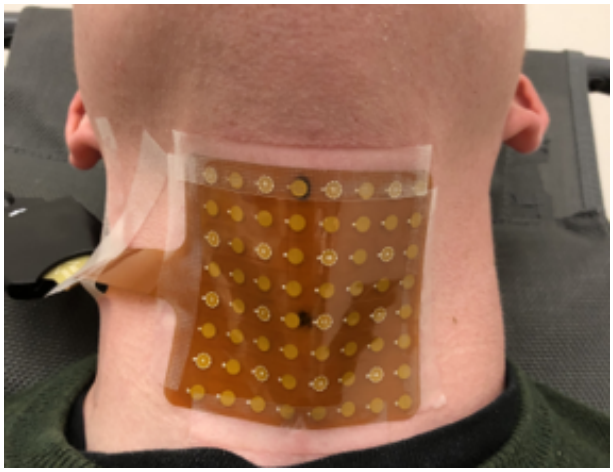

Figure 3.5: A 64-channel array oriented such that its center overlayed the cricothyroid space. Row 1 (electrodes 1-8) overlayed the hyoid bone.

Signal Acquisition

The HDsEMG array utilized the REFA 128-model system (REFA, TMSi, International, the Netherlands) sampled at 1024 Hz and used the data average as a reference, while the ground electrode was placed on the wrist. Impedance values were recorded for both arrays to evaluate for technical flaw; all channels were within acceptable range (< 10 kΩ).

The subjects were asked to swallow substances of various textures each through a 1-minute period of electromyographic signal recording. The swallow textures were as follows: (1) dry/salivary (2) thin liquid (water), (3) puree (apple sauce), (4) mixed consistency (fruit cup),

(5) dry solids (cracker). Subjects were careful to limit mouth opening, maintain neutral neck positioning, and refrain from phonation during recording. The number of swallow events and time of occurrence was noted by observation of the subject from the research coordinator.

The recorded data was transferred for EMG analysis through the system's respective software (Polybench, TMSi, International, the Netherlands) and subsequently to custom software written in Python programming language (Python Software Foundation, Python Language Reference, v.3.7.1.python.org/).

Signal Processing

Initially, the raw data was processed with a low-pass filter using a cut-off frequency of 3 Hz. Then it was rectified with a moving root-mean square using a window size of 1 s. A linear envelope was generated using a window size of 1 ms and a cutoff frequency of 3Hz (Winter D.A. 2009).

The on and off time-points of a swallow were detected using two methods. The first looked at the traditional method of setting a threshold and obtaining points that surpass the threshold [30, 34, 138]. The threshold was calculated as 2 standard deviations above the average resting baseline. The second method utilized the Mahalanobis distance through the following rationale:

Our HDsEMG multivariate data is made of 64 channels therefore giving us 64 possible baselines. Defining the baseline as noise, we can now compute the mean and covariance matrix of the noise as μ_{noise} and $\hat{K}$ respectively. The $\hat{K}$ covariance estimate is obtained with an L1-regularized model fitting procedure.

To determine if a data point of the signal, x, is a swallow or noise, we can take the difference between the mean and the data sample, divided by standard deviation to normalize, which is formally known as Mahalanobis Distance metric.

$$d(x, \underline{\mu}_{noise}) = \sqrt{(x - \underline{\mu}_{noise})^T \hat{K}^{-1} (x - \underline{\mu}_{noise})}$$

The greater the difference a data point x is from $\underline{\mu}_{noise}$, the more likely that point is a swallow and not noise.

Thirdly, the mahalanobis distance metric results, were implemented a binary classification using a One-Versus-All method.

This allows us to use a binary logistic regression model for multi-class classification.

We train M logistic regression models, where:

- M is the number of classes, 5, one for each type of swallow.
- Y is the yes or no that distinguishes one swallow type from the rest.
- R, regressors, is the information from swallow, in this case our amplitude after the RMS (V_{RMS})
- x, weights
- N is our gaussian noise with a distribution of $N(0, \sigma^2)$

Therefore given this information, where $j = 1...M$ classes, we ask ourselves:

$$p(\underline{y}|R^{(j)}, \underline{x}^{(j)})$$

$$y = \begin{cases} 0 & \text{if sample is the same as the rest} \\ 1 & \text{if sample is different from the rest} \end{cases}$$

Each swallow class is passed through M models and the largest value produced is the best class estimate.

3.2.3 Results

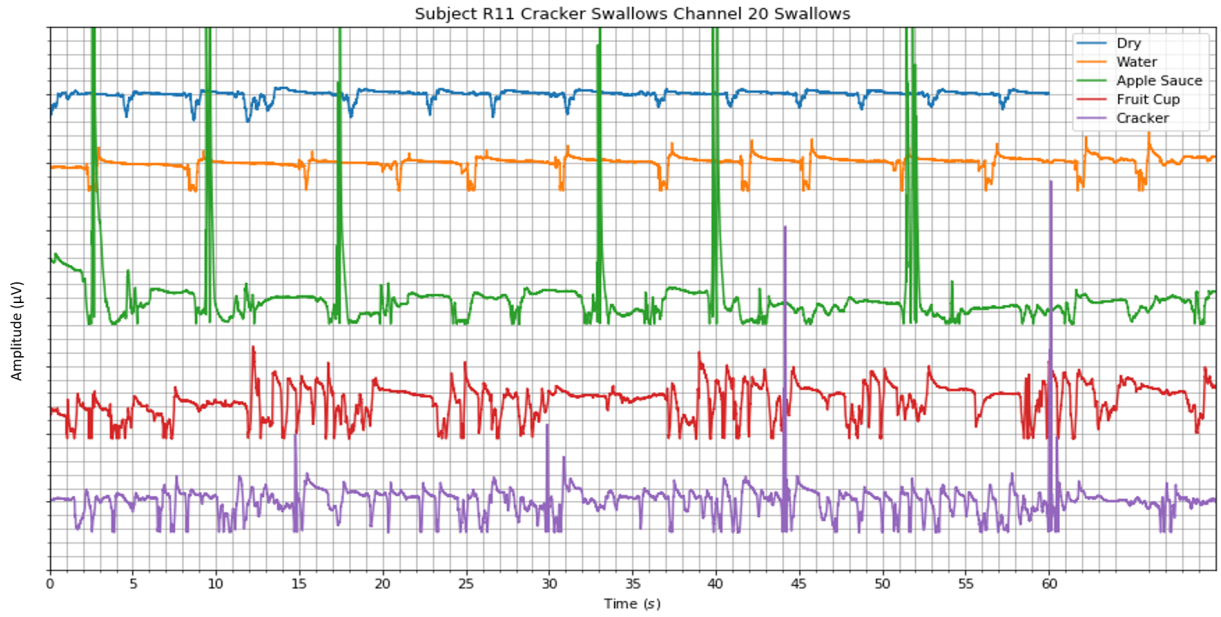

Figure 3.6: Time-series plot of the raw data of 5 swallows for a single subject and channel.

Visualization of the raw data illustrated distinct wave morphologies that appear to be specific to each swallow task (Fig. 3.6). In all of the subjects, the dry/saliva swallows and water swallows appeared to have the simplest forms of EMG activity, only exhibiting between one to three peaks throughout the swallow epoch. As the complexity of food bolus increased, the number of peaks present began to increase as well. In one subject, the swallow peaks during a cracker swallow appeared continuous, with little "rest" in between the swallow epochs (Fig. 3.6).

After processing, a linear envelope of the data smoothed the waveforms and better illustrated the varying number of EMG peaks during their respective tasks. Fig. 3.7 demonstrates a single epoch of swallow for each of the tasks in one subject and one channel. The overall duration of these swallow epochs demonstrate a monotonically increasing duration proportional with swallow complexity (Fig. 3.8).

Swallow detection performed via the traditional method showed performance varied between channels when compared to the ground truth. Using Malahanobs distance as a metric

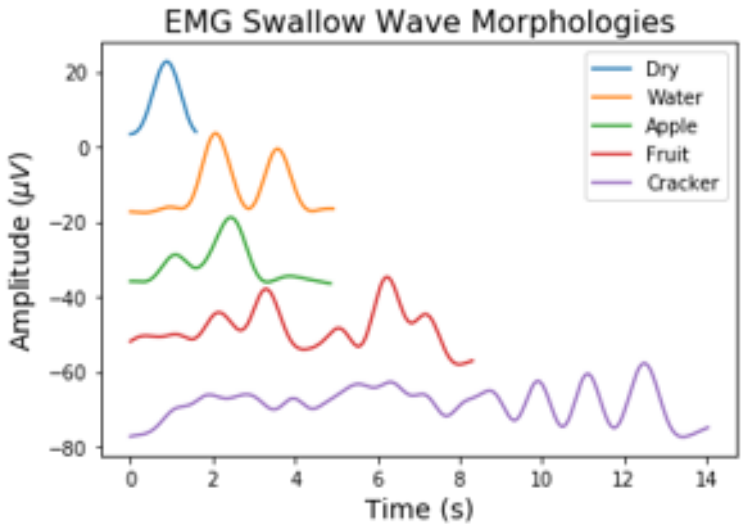

Figure 3.7: A linear envelope of the 5 different swallows from a single channel and subject using a low-pass filter with a cut-off of 3 Hz.

for swallow detection improved the ability to distinguish rest from activity. Moreover, it did not require any input from the researcher. Classification of tasks using the one-versus-all method provided promising results. Classification was significantly better in the raw data than with the processed data.

3.2.4 Discussion

The process of swallowing requires the coordinated contraction of various muscles in different locations and at different times. Neurologic disorders, stroke, nerve injury, and aging can affect the level and order of the muscles' activation. This impairs a patient's ability ability to swallow, therefore lowering their quality of life while increasing their risk of malnutrition, aspiration, and pneumonia. A standard treatment option for patients can be swallow therapy through speech-language pathologists. There, patients can learn about various positions and exercises to minimize their risk of aspiration and encourage muscle coordination.

Biofeedback through surface EMG has been a potential method of increasing patient engagement in the clinic [56, 59, 124, 125]. However it lacks standardization, is susceptible

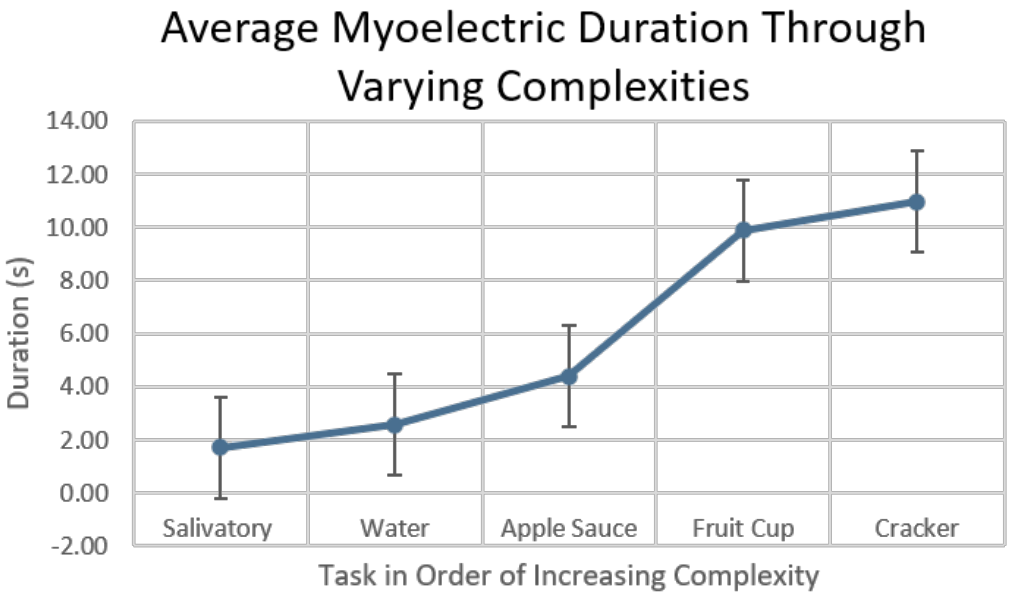

Figure 3.8: Averaged EMG activity across subjects per task.

to bias, and does not measure compliance. Moreover, in the clinic, sEMG has struggled to secure a foothold due to its inability to robustly record specific muscle activity in such a dense neuromusuclar area [41, 71, 117, 138, 142]. Notable limitations involving risk of signal cross-talk and diminished electrode-to-target muscle specificity make its clinical applicability doubtful. In research, flexible strain sensors were able to distinguish between types of swallow bolus and dysphagic versus nondysphagic swallows [103]. When paired with sEMG, Ramirez et al. found that it was possible to differentiate swallow activity from non-swallow activity.

In this study, three patients performed self-paced swallows using 5 different bolus texture complexities ranging from dry (saliva), water, apple sauce, fruit cup, and cracker. Results indicated a monotonic increase in EMG activity duration proportional to the increasing complexity of the swallow bolus. In terms of duration, similarities were seen between swallows of similar complexity. At times, the duration during a water swallow was similar to that of an apple sauce swallow. This was similarly seen between fruit cup and cracker swallows. During classification the average error probability was 0.16 in the raw data. The error though increased with complex

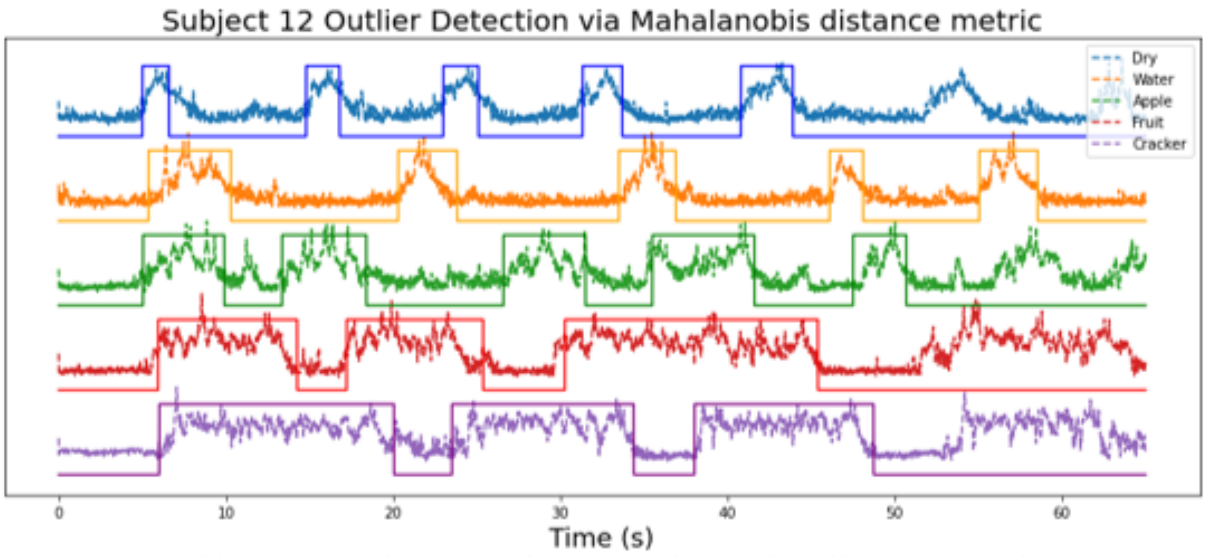

Figure 3.9: Mahalanobis distance for swallow detection. The solid lines represent the "ground-truth" epochs of swallow. The dashes lines represent the mahalanobis distance.

swallows such as fruit cup and crackers. This may be due to the nature of the swallow itself. A swallow epoch was recorded through visual observation by the research coordinator. The start and stop of a swallow was determined by the initiation of swallow until the full completion. However, during fruit cup and cracker swallows, the subject does not swallow in one movement, but is chewing and swallowing simultaneously. As a result, it could be argued that during complex swallows, a swallow epoch contains many small swallow EMG peaks. Future studies could mediate this discrepancy by having subjects perform swallows using a controlled set of increasing volumes. If the number of peaks during a complex swallow increases with the volume, then it may be possible to address this discrepancy.

Its interesting to note that typically in the clinic, the EMG data received is pre-processed by the signal acquisition system. When we demonstrated our distinct morphologies to our collaborating clinician, he stated he had never seen such signals before. Furthermore, classification appeared to work best with raw data. Perhaps there is a unique aspect to the raw data which makes it ideal for classification that is lost during processing. This could be a limitation that affected previous research attempting to do the same. Limitations in this study include a low-number

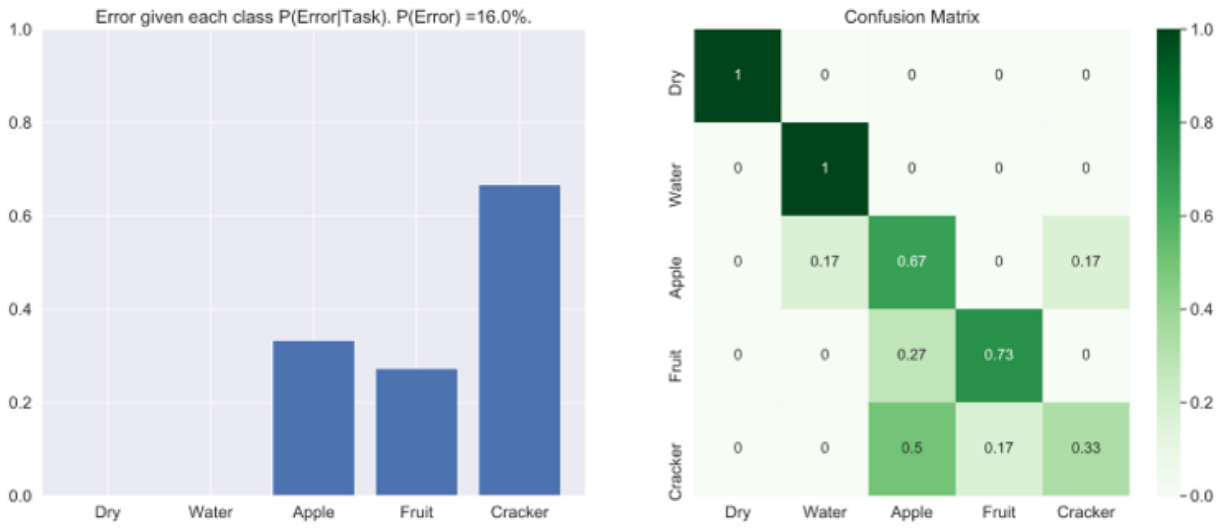

Figure 3.10: Probablity of error in classfictain using the one-versus-all method for the five different types of swallows.

of subjects due to COVID-19, lack of volume control for the swallow bolus, and the lack of time-controlled swallows. We chose for subjects to perform self-paced swallows as we found that in the patient population, they struggled with swallow initiation. We desired a dataset that would most fit an ambulatory setting, one in which a patient could go home, place a flexible electronic sensor array, and eat at their own pace.

The use of different swallow textures was intended to mimic the different tasks they may be asked to perform during swallow therapy. To develop en effective and objective biofeedback and ambulatory monitoring system, it is critical to understand the extent of information that can be obtained noninvasively. This will allow us to determine what therapists and patients will find more pertinent towards managing their dysphagic symptoms. Though current swallow sEMG devices in the market are sufficient to encourage increased contraction, this could be falsely reported at home via an incorrect movement or other laryngeal activation. In developing a system capable of distinguishing swallows, we can better determine a subject's progress. For instance, perhaps at the start of treatment, a subject is unable to initiate swallow. Therefore there is an increase of preparatory movement across all types of swallows. We would then expect the swallow types are

indistinguishable with a high error probability. Over time with HDsEMG, we can use energy maps to monitor symmetric activity, duration of swallows, and classification of swallows to determine improve treatment decision-making. Coupled with other non-invasive devices such as flexible strain sensors, we anticipate a greater understanding of dysphagic swallows and their response to medical and therapeutic treatment. Future studies should investigate the use of controlled swallow volumes and improved feature detection for classification. It may also be useful to study dimension reduction and investigate the contributions from individual electrodes to a swallow's classification.

3.2.5 Conclusion

Spatiotemporal data from HDsEMG improves the application of sEMG for clinical and therapeutic purposes in swallow disorders. Section 3.2 demonstrates the ability to once again detect lateral array symmetry of activity, and greater EMG activity in the upper rows of the array. In using different textures of swallow, we found that the duration of EMG activity increases proportionally with the complexity of the swallow texture. Using the RMS voltage and one-versus-all classification, we were able to distinguish the five different classes of swallows with a probability error of 0.16. These results demonstrate a promising application in the therapeutic field for dysphagia. Currently the features used are common in the realm of clinical EMG applications, but with improved feature detection, and dimensionality reduction, this technology holds significant promise to develop into a powerful clinical and diagnostic tool with machine learning paradigms.

3.3 Acknowledgments

Chapter 3, Section 3.1, in full is currently in preparation for submission for publication of the material. Bracken, David J.; Ornelas, Gladys; Weissbrod, Philip A.; Coleman, Todd P., The

dissertation author will be the co-author of this paper.

Chapter 3, Section 3.2, in full is currently in preparation for submission for publication of the material. Ornelas, Gladys; Bracken, David J.; Bueno Garcia, Hassler; Weissbrod, Philip A.; Coleman, Todd P., The dissertation author was the primary investigator and author of this material.

Chapter 4

Clinical Applications

4.1 Validation against the Gold Standard: A Study of the Forearm

4.1.1 Introduction

Chapters 2 and 3 demonstrated the value of HDsEMG for voice and swallow healthcare applications. However we are still dealing with an intricate muscular layout and significant overlap of neighboring muscles. Many will undoubtedly ask, how can we prove our results are consistent with the clinical "gold-standard". Though the eventual goal is to produce a portable ambulatory system, its necessary to first validate the protocol for both voice and swallow. To do so, a sensible approach is to perform concurrent cutaneous and subcutaneous EMG recording of the laryngeal muscles. Unfortunately, the initiation of this project occurred simultaneously with the COVID-19 pandemic. In the clinical field of laryngology, clinical research involving human subjects was restricted to only studies deemed necessary or related to COVID-19. Additionally, the social-distancing and stay-at-home mandate restricted any recruitment of healthy subjects or collaboration with research colleagues. As a result, the project was unable to recruit subjects for

validation experiments in the neck. Though I was unable to conduct laryngeal related experiments, I initiated a plan to design an analytical workflow using an alternative experiment: concurrent nEMG and HDsEMG on the forearm.

Electromyography from the forearm has been studied extensively in the realm of human-computer interface and prosthetic design [24, 102, 110]. Therefore, in the era of work-from-home, the forearm was chosen for many reasons: 1) there is extensive literature to compare results with, meaning that a self-taught researcher could easily validate the protocol; 2) I would be able to insert the needles into one forearm using my other hand. This was especially important since I could not rely on the presence of my research assistants or colleagues to help with needle insertion; 3) I could study individual digits, allowing me to mimic task-specific activities such as the phonation tasks used to separate the adductor and CT muscles; and 4) the forearm is also a muscularly dense area, which once again, could approach a similar paradigm that may be experienced in the throat, yet the muscles of the hand are relatively superficial and easy to find for an nEMG novice.

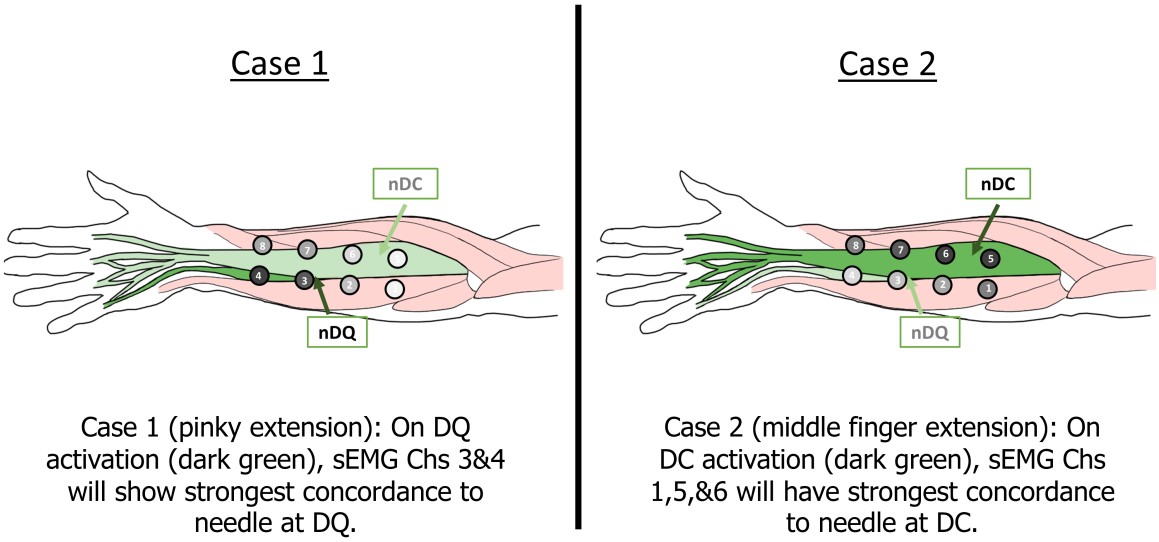

Figure 4.1: Illustration depicting the experimental paradigm A and hypothesis using concurrent needle and surface channels for validation.

The goal of Section 4.1 is to produce an analytical workflow that would be applicable for validation in the neck. The hypothesis behind this project is that data from the surface electrodes

should have the greatest concordance to their nearest needle electrode. This concordance will increase when the muscle of interest, in which the needle is inserted, becomes active. For instance, the Extensor Digiti Quinti (DQ) is located laterally in the forearm, in the posterior compartment, and can be activated through fifth-digit (pinky) extension. The Extensor Digitorum Communis (DC) is a superficial muscle that runs near the middle of the forearm, in the posterior compartment. It extends into the middle four digits, though extension of the third digit (middle-finger) is an effective way of isolating it. My hypothesis is illustrated by fig. 4.1 in which eight surface channels are placed along the forearm and two needle electrodes are inserted into the DQ (nDQ) and DC (nDC) muscle. Specifically, during pinky-finger contraction, case 1, surface channels 3 and 4 are expected to have greatest similarity with nDQ. In case 2, during middle-finger contraction, surface electrodes 5 and 6 will be most similar to the needle at DC (nDC).

Should this workflow be extrapolated to applications of the neck, a similar hypothesis will apply. During phonation of a high-pitch tone, electrodes nearest the CT region should exhibit greater concordance. During low-pitch phonation, we expect greater concordance in the adductor region. These results could similarly be tested through the LLR hypothesis testing mentioned in section 2.2.

4.1.2 Materials and Methods

The pilot experiment was conducted on a single subject, age 27. Two paradigms were used. Paradigm A followed the paradigm depicted in Fig. 4.1. First the left-forearm skin was prepared with standard alcohol wipes and exfoliating impedance-reduction gel (NuPrep Skin Prep Gel, Weaver and Company, Aurora, Colorado, USA). Next the insertion sites were determined using standard anatomical points from literature [92] as reference. By extending the pinky and middle-fingers respectively, we searched for the most prominent area of contraction along the forearm and marked the spots using a skin-safe medical marker. Paradigm B marked an additional insertion point for a third needle electrode, this time inserted into the Extensor Pollicis

Brevis which was activated by a thumb extension. The needles used were 0.38mm stainless steel electrodes (Neuroguard Disposable Subdermal EEG Needle electrode (REF: S41-638), Neuro Supply, Loveland, Ohio, USA).

The surface electrodes used were composed of standard silver/silver chloride electrocardiac monitoring electrodes (3M Red Dot (REF: 2670-5), St. Paul, Minnesota, USA). For paradigm A, the electrodes were organized into 2 rows and 4 columns with an interelectrode distance of 2 cm along the rows and 3 cm between the columns, center-to-center (electrode diameter 1 cm). Paradigm B used an array of 3 rows and 6 columns, with the same spacing mentioned for paradigm A. Fig. 4.2 is a picture depicting both electrode configurations. To minimize movement artifact, the arm was rested over a table throughout the experiment.

Signal Acquisition

Signal acquisition for paradigm A used two machines. The needle data was obtained through the Biopac MP150 using a gain of 1000 and with a 5 kHz low-pass set on DC (BIOPAC Systems, Inc, Goleta, California, USA). The sampling frequency was 4000 Hz. The surface data was collected through the OpenBCI Cyton device (http://openbci.com) with a sampling rate of 250 Hz and a common ground electrode placed on the elbow. Paradigm B used a research device called the WANDmini (Muller Lab, Department of Electrical Engineering and Computer Sciences, UC Berkeley, California, USA) for both needle and surface channels. Sampling rate was 1024 Hz.

A total of four tasks were performed by the subject: pinky extension, middle finger extension, index finger extension, and thumb extension. The extensions were sustained for 5 s in between 10 s of rest. This was repeated three times for each finger. The raw data was then processed through through custom software written in Python programming language (Python Software Foundation, Python Language Reference, v. 3.7.1).

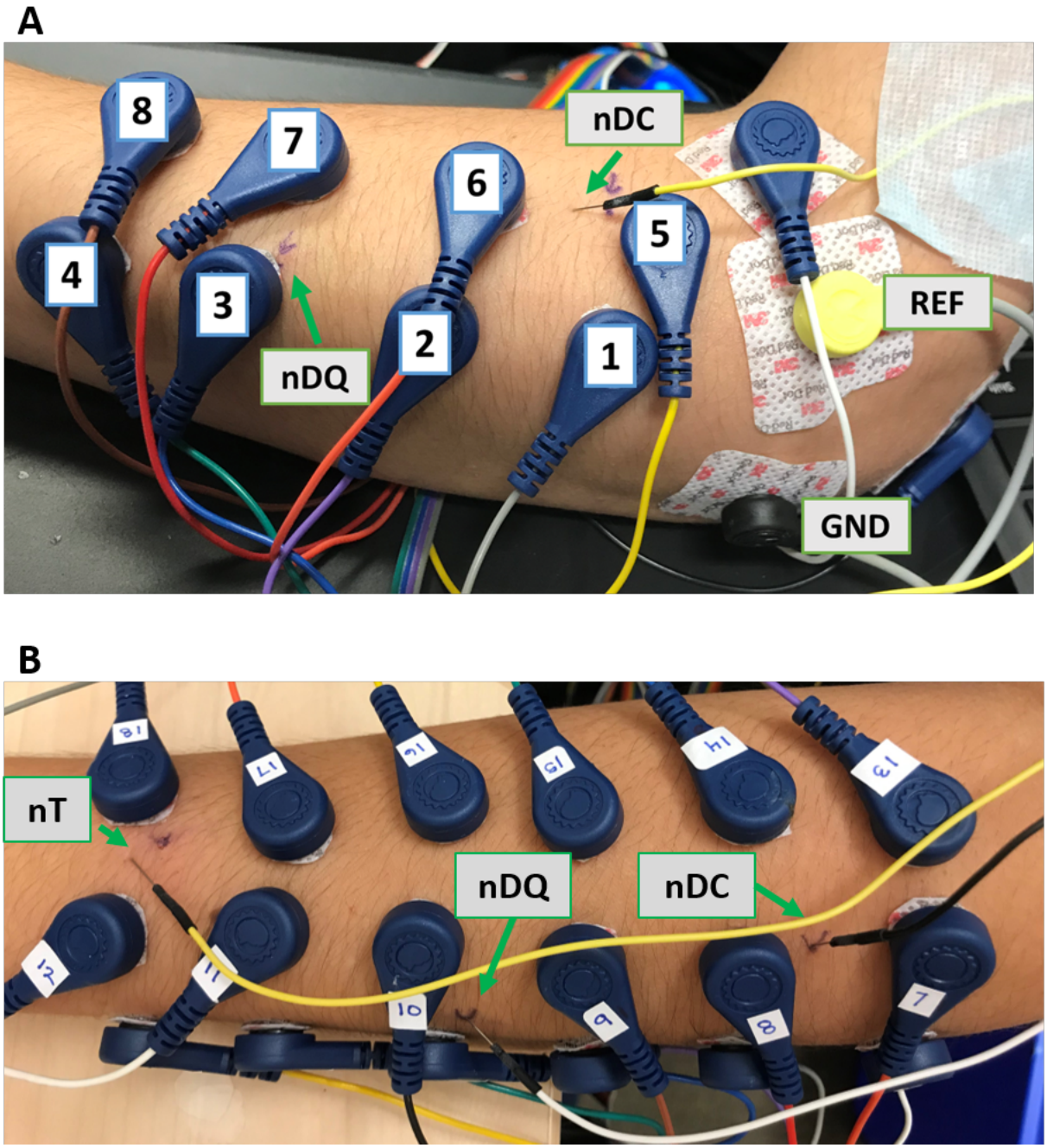

Figure 4.2: This figure depicts the two paradigms used for this validation experiment. Top picture demonstrates the 8-channel paradigm with two needle channels, referred to as paradigm A in this chapter. The bottom picture depicts the 18-channel HDsEMG configuration with 3 needle electrodes, referred to as paradigm B.

Signal Processing

Pearson's Correlation

Both surface and needle data were first rectified with a moving root mean square (RMS) using a window of 1 s. Then a linear envelope was generated using a low-pass Butterworth filter with a cutoff frequency of 3Hz. In paradigm A, the needle data was then down-sampled to match the length of the surface data. This was simply done by taking every 16th sample-point of the needle data. An additional step used a threshold calculated from the average of the data. If a point from the enveloped data surpassed the threshold, it was labeled 1. If the point was below the threshold, it was labeled 0. This was done to determine on a binary scale, how often did surface recording of muscle activity match that the needle-recorded EMG activity.

Two Pearson's correlation (PC) statistics were calculated finally calculated after processing. The first was called a "global" PC. This global PC was calculated by comparing the entire nDQ data with each of the sEMG channels, then repeating this with the nDC data. This was done to compare the two datasets across the duration of the experiment. The second PC, referred to as a "local" PC, involved comparing the needle data to the surface channels only during specific tasks. First the data was segmented by tasks. Then during each finger extension task, PC was calculated between the surface channels and the needle inserted into the activating muscle.

Onset Detection

This method sought to count the number of onsets of EMG burst activity using the python Biosppy EMG signal analysis module [27]. This module essentially processes the data with a Butterworth high-pass filter using a cut-off frequency of 100 Hz and an order of 4. The filtered signal is then rectified using the absolute value and smoothed with a moving average filter [121]. Finally, a threshold is calculated as two standard deviations from the average of the processed data. Processed data points above the threshold for over 100 ms are labeled as onsets. It was hypothesized that the surface channels will have a number of counts similar to those of the needle near it. Fig. 4.3 is a block diagram depicting this process as well as a plot of the results for the

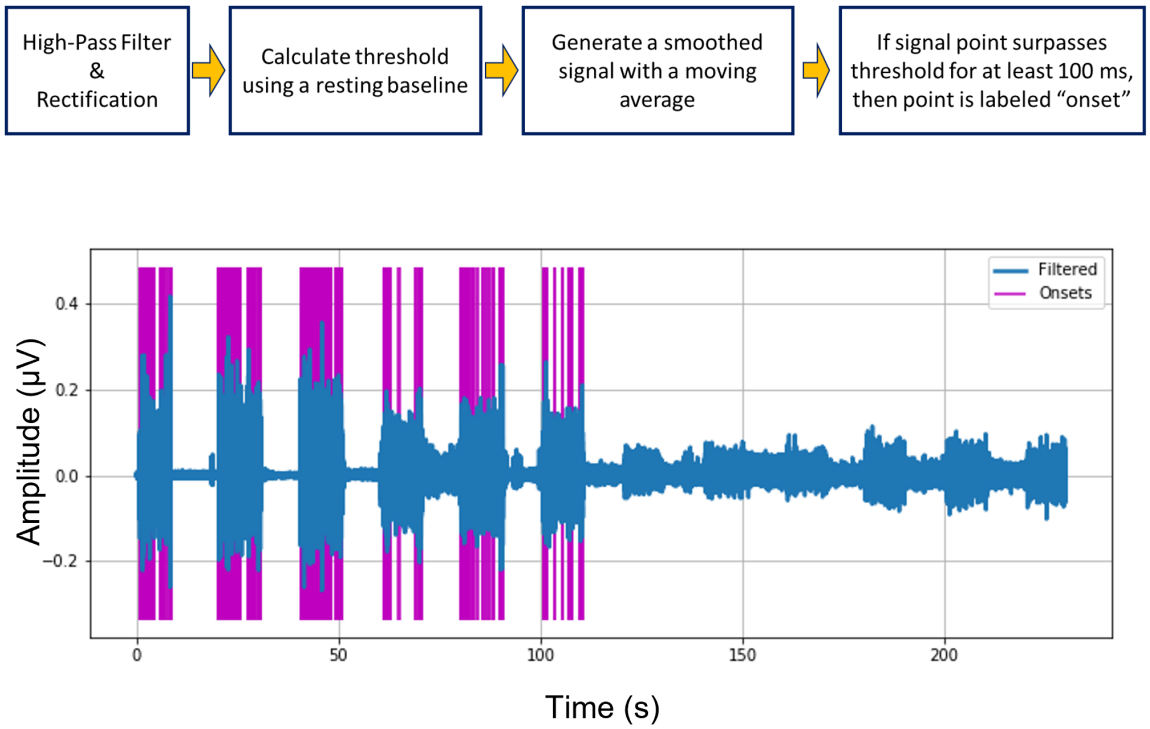

Figure 4.3: The block diagram depicts the processing of the data to count EMG activity onsets. The plot below is of nDQ during all extension tasks. The magenta lines are the onsets detected with the baseline threshold.

nDQ data of paradigm A.

4.1.3 Results

For the Pearson’s correlation results, the global PC demonstrated some correlation between a needle and the nearest surface channels. This is most evident in looking at the nDC vs HDsEMG PC Fig 4.4. For pinky extension, concordance between needle and surface channels increased with the local PC. However, during middle-finger extension, correlation was high throughout the channels.

Fig. 4.5 depicts six different energy maps for the onset counts. The top three are from paradigm A which used 8 surface channels and 2 needle electrodes. Here we see a strong similarity

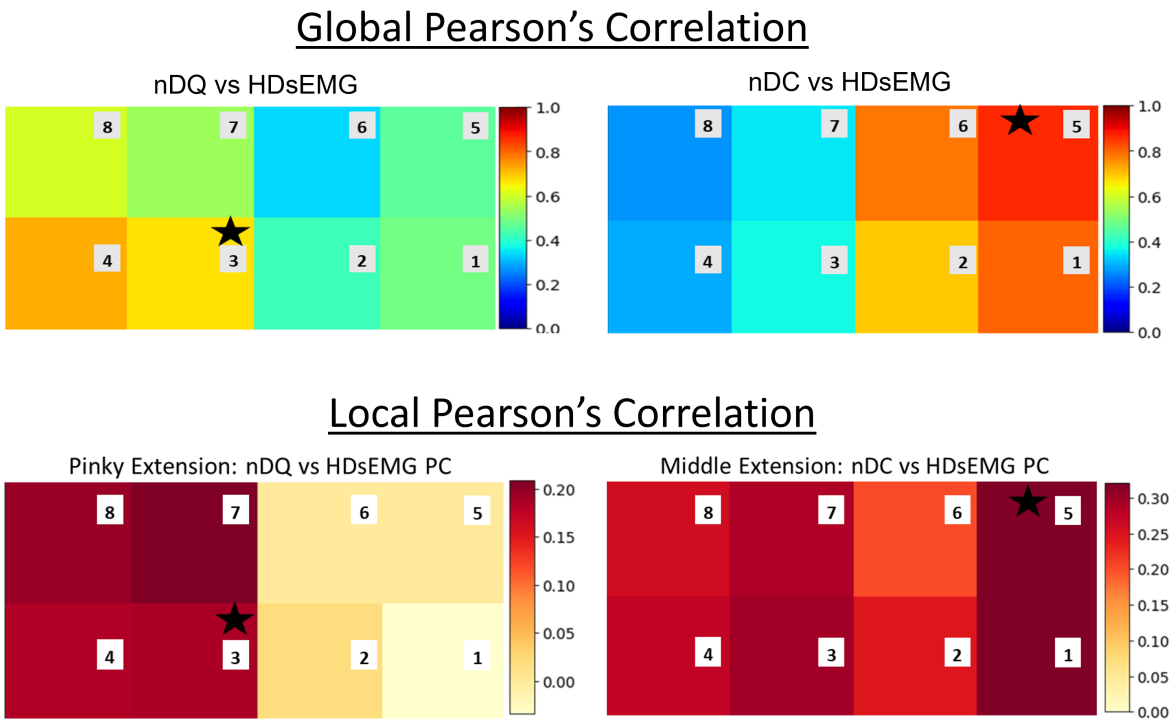

Figure 4.4: The global PC is was calculated across the entirety of the experiment between the surface channels and needles. The local PC was calculated only with segmented tasks.

in onset counts between the needle channel and the surface channels closest to it. Performing extensions in fingers without needles, such as the thumb and the index finger, served as great negative controls. During the thumb extension, there is very little onset activity recorded by either nDQ or nDC needles (Fig 4.5), as well as most of the surrounding electrodes, except for channels 7 and 8. Surface channels 7 and 8 recorded low onset counts (<40).

The promising onset-count results from paradigm A encouraged me to increase the number of surface channels and add an additional needle, leading to paradigm B. Onset counts from paradigm B are depicted as the bottom three images in Fig 4.5. Here we once again see similar numbers between the surface and needle data that validate the results from paradigm A. Additionally, adding the third needle (nT) allowed the thumb extension task to go from being a negative control to serving as a positive control and further validating our hypothesis.

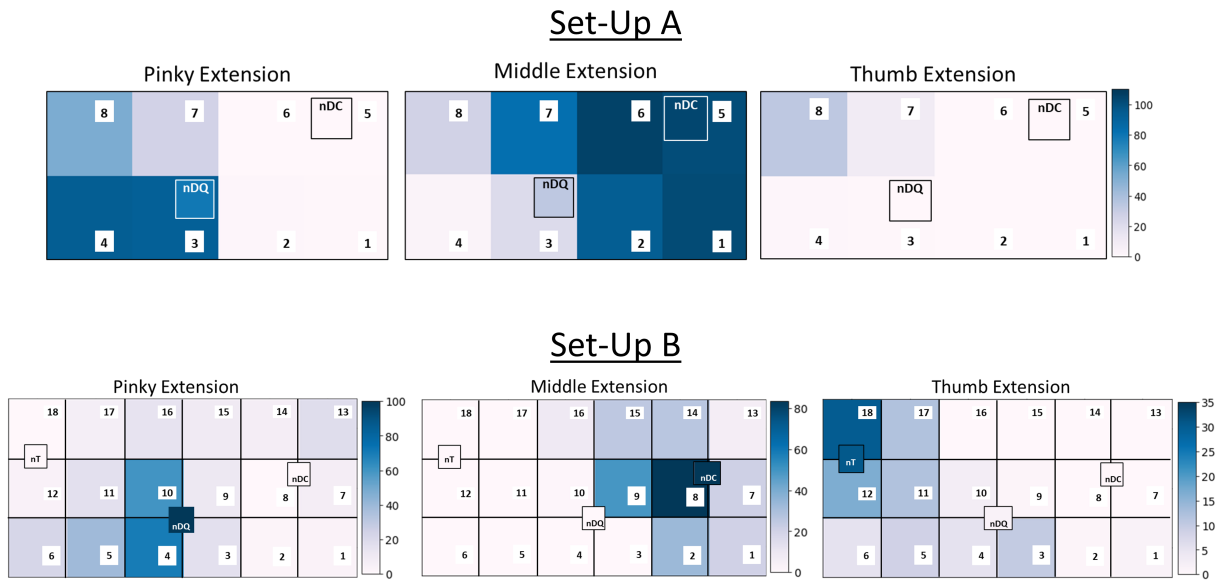

Figure 4.5: The top three figures represent results from 3 tasks (pink, middle, thumb extension) using paradigm A. Paradigm A consists of 8 surface channels and two needles. The bottom three depict results from paradigm B, now using 18 surface channels and 3 needle electrodes. In paradigm B, the thumb extension task is now a positive control.

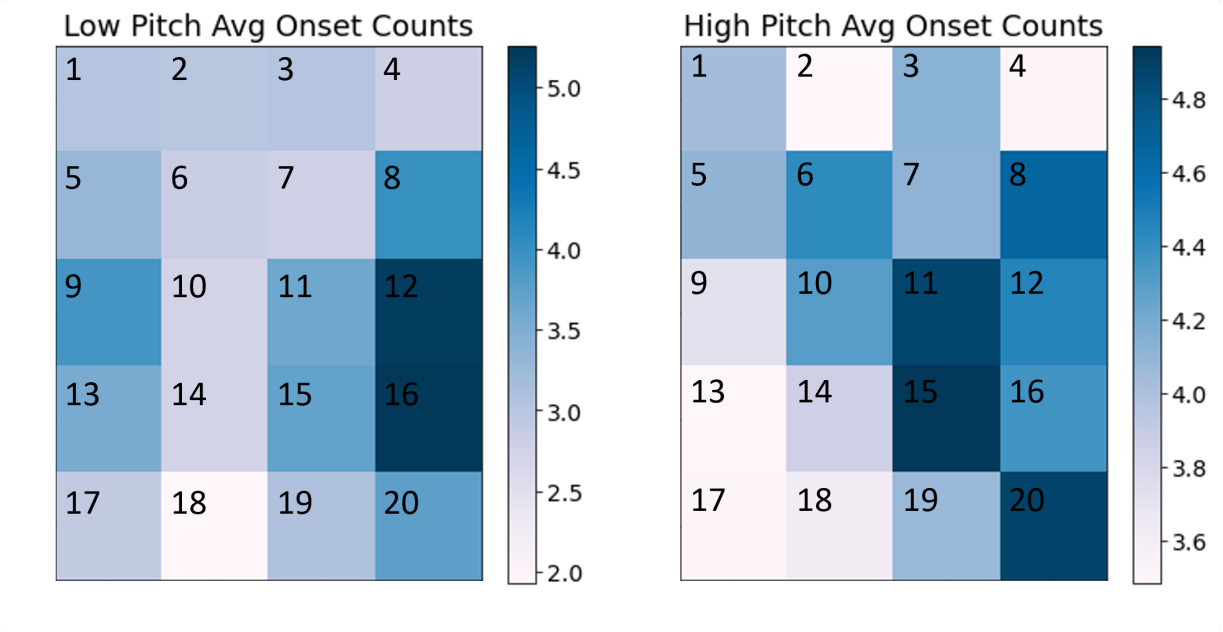

Figure 4.6: The two images above represent the averaged onset counts for low and high-pitch phonation from the previous dataset (chapter 2). Note the significant difference in amplitudes between forearm and the neck.

Lastly, though we were unable to perform the experiment on the throat, we applied the onset-count workflow to the phonation data from chapter 2. This dataset was put through the same signal analysis using the Biosppy python module. Using the same rationale from the log likelihood ratio (LLR) test from Ch. 2.2, the onset counts from the low vs high pitch differentiation were used to evaluate their LLR and ROC curve. The pulse counts averaged across trials and subjects are shown in Fig 4.6 while Fig 4.7 is the resulting ROC curve. Though the variance between subjects has increased, the average AUC was calculated to be 0.84.

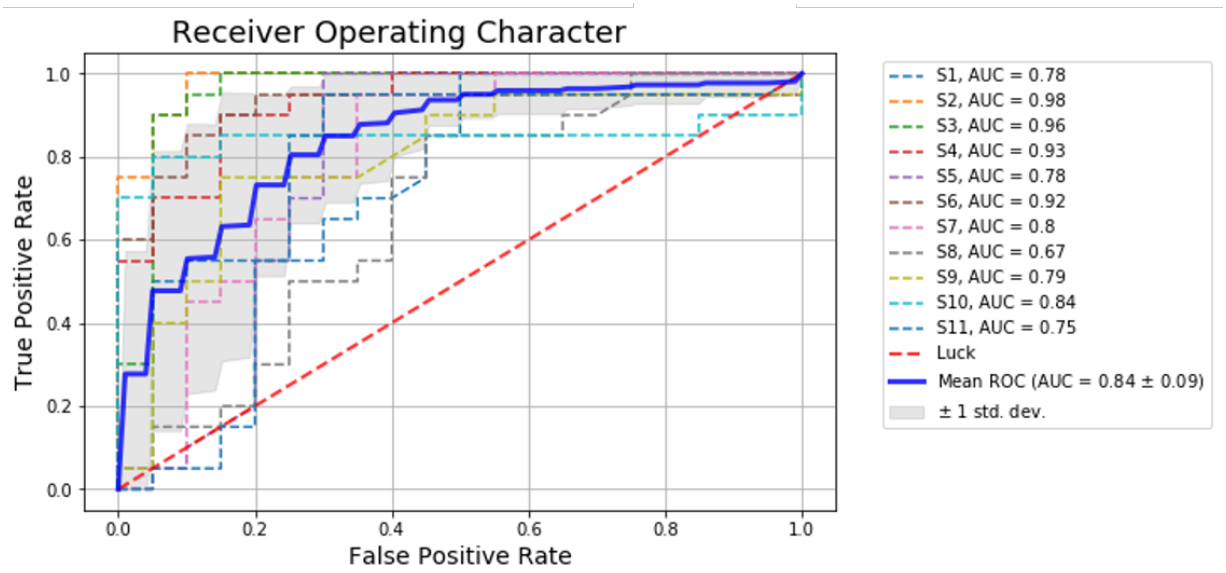

Figure 4.7: Though visually there doesn't appear to be a significant difference between the counts for low and high-pitch phonation when averaged across subject and task, the ROC curve shows a positive distinction between the two with an averaged AUC of 0.84.

4.1.4 Discussion

The throat is a complex region with many muscles working together to initiate various activities: swallowing, speaking, maintaining stability, and physical movement. The muscles are of different sizes, located at various depths from the surface of the neck, and operate at different times. With the number of nerves, muscles, tendons, and bones, there is no doubt regarding

the complexity of the neck and throat, and the difficulties in trying to monitor it non-invasively. Therefore historically only invasive methods and subjective qualitative methods have been used to diagnose, assess, and monitor disorders of the throat. Nonetheless, HDsEMG may provide an objective non-invasive assessment tool for voice and swallow disorders of the neck.

Though laryngeal HDsEMG research is promising, it is necessary to confirm results against the clinical "gold-standard", needle or hookwire EMG, before its implementation in the clinic or therapeutic settings. Unfortunately the COVID-19 pandemic hampered the initiation of this study. Not only was human-subject research restricted to critical or COVID-related projects, but the stay-at-home mandate made it difficult to perform invasive experiments single-handedly. A compromise was found with using the forearm. Performing simultaneous HDsEMG and nEMG on the forearm was possibly while alone. I was able to still use one hand freely while the other was readied for the experiment. Additionally, the forearm is a highly studied area. In the absence of expert training or supervision, this allowed me to compare my results with existing studies to ensure my needle placement and monitoring was correct.

The hypothesis of this project posited there would be greater concordance between the spatially-related needle and surface channels, especially when the target muscle (in which the needle was inserted) became active. For nDQ, we expected greater similarities to surface channels 3, 4, and 7 (paradigm A) and 3, 4, 9, and 10 in paradigm B. For nDC, we expected surface channels 1, 5, 6, and 7 (paradigm A) and 1, 7, 8, and 9 (paradigm B). The nT needle was only present in paradigm B, and we expected it to be most similar to surface channels 17 and 18.

Pearson's correlation was first used to study the concordance between nEMG and HD-sEMG due to its simplicity and widespread use. A global PC demonstrated a stronger correlation between HDsEMG and the nDC channel. This correlation significantly increased in the local PC (calculated only during the middle-finger extension) exhibited by a "reddening" of the entire map. This may be a due to the fact that the DC muscle is large and extends into three of the five digits. As a result, many of the sEMG channels laid over, or near the muscle. Therefore, when the

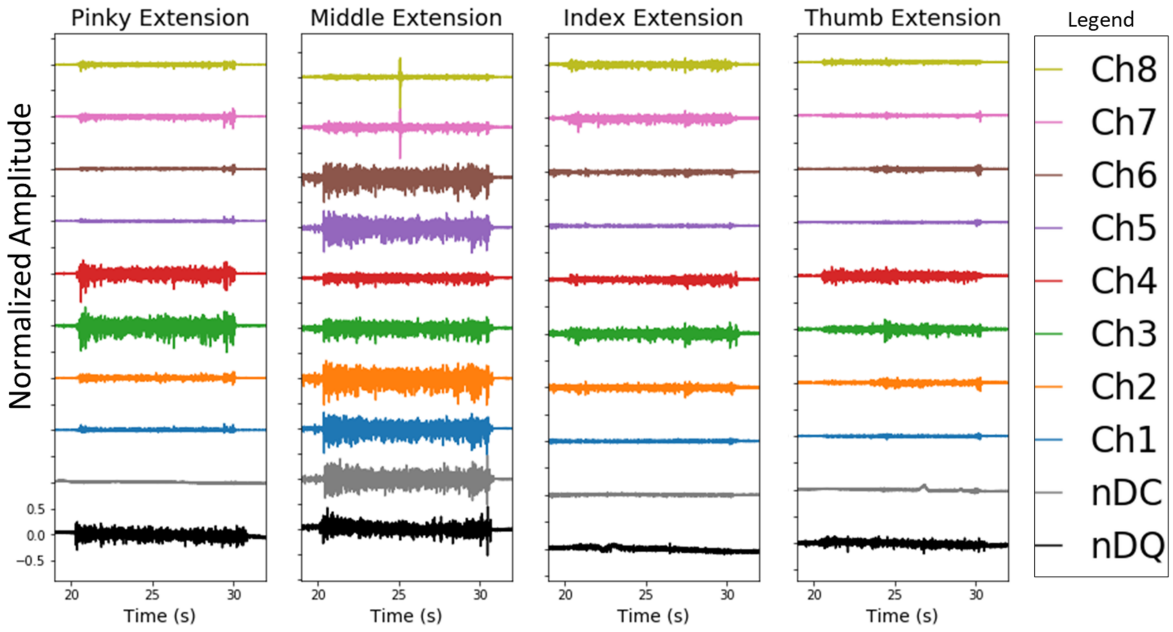

Figure 4.8: Processed time-series data for paradigm A. Each plot contains a single finger-extension repetition across all channels. The voltages have been normalized and centered at 0.

DC muscle is activated, it is recorded by many of the channels, hence the high correlation (see Middle Extension plot Fig. 4.8). For nDQ, there is a low correlation in the global PC comparison. However this correlation increases when using the local PC. Unlike the DC muscle, the DQ muscle is small and located deeper in the forearm. This may explain why the correlation is only evident during the pinky-finger extension. The time-series plots in Fig. 4.8 demonstrate how the normalized amplitudes compare to the needles during specific finger extension tasks.

Though the PC results supported our hypothesis, the concordance was weaker then expected. This could potentially be a problem in translating this workflow to laryngeal EMG. The neck is much more complex than the forearm and the target muscles are smaller and deal with significant overlap of neighboring muscles. Therefore, the PC method may be weak when applied to the neck and a better analysis method may be needed. This led to the use of onset detection, which proved to be fruitful. In paradigms A and B, both nDC and nDQ recorded similar counts of

onset activity as those of their neighboring surface electrodes (Fig. 4.5). Through onset detection, a greater spatial specificity was achieved across all tasks and supported our hypothesis. Pinky extension sees similar counts between nDQ and surface channels 3 and 4 (paradigm A) and 4 and 10 (paradigm B). Though in paradigm B we expected channels 3 and 9 to also detect the same number of onsets as nDQ, the lack of detection could be due to the direction of muscle fiber activation (down the forearm towards the fingers). Similarly for nDC, in paradigm A we see the expected results in the predicted surface channels, however in paradigm B, it is restricted to surface channels 8 and 9. This is an interesting result as it demonstrates that volume conduction at the surface of the skin is not as big of an issue as typically depicted by literature. In essence, detecting similar onset counts at the surface, in less of the expected channels, tells us that sEMG is much more specific than we believed, but requires an HDsEMG to detect across space. This is promising for applications in complex surface regions such as the neck. Moreover, these results were further verified through the addition of a third needle, nT, in paradigm B. Though nT was inserted into the Extensor Pollicis Brevis which extends into the middle of the forearm, only the surface channels closest to the edge, and near the needle, detected similar counts as nT.

When applying the onset-activity detection workflow to the voice data, we continued to see greater counts in the central regions hypothesized to be where the laryngeal muscles are located. However its important to note that the exact channels that detected the greatest counts weren't consistent between subjects. This can be seen with the averaged counts depicted in Fig. 4.6. The averaged results illustrate a grid array skewed towards the reader's right. This could be a result of human error when placing the electrode grid, noisy electrodes due to improper skin-preparation, or a natural tendency of the subject to activate one side more than the other. Nonetheless, there were quantitative differences in counts detected between low- and high-pitch detection that produced an ROC curve with an AUC of 0.84 (Fig. 4.7). This implies that in these two types of pitches, there is unique activity at the neuromusuclar level that is still captured non-invasively. Perhaps in unhealthy subjects, this wouldn't be the case. Another important

characteristic is the significant difference in amplitudes between forearm and the neck onset-counts. On the forearm, onsets were detected in the quantity of 100 in a single subject, whereas in voice, the number of counts detected were less than 20, and around six when averaged across all 10 subjects. This may be due to the loss of signal as it transverses the many layers between the muscle and skin in the neck. This may be a set-back in future studies, but could also be used as another metric for healthy vs unhealthy subject comparison.

4.1.5 Conclusion

HDsEMG improves existing sEMG modalities by compensating for spatial selectivity loss and globally monitoring the anterior neck through a dense electrode configuration. The results from this study were a positive direction towards validating our hypothesis and results from chapter 2. Through onset-activity detection, we demonstrate the spatial correlation between needle and HDsEMG. Though the results come from a different anatomical region, the project itself helped design an experimental workflow for laryngeal EMG. Despite the lower number of counts detected from the neck, there were sufficient distributional differences to distinguish low from high-pitch phonation using the log likelihood ratio test. Moreover, the spatial relevance for surface monitoring of laryngeal muscles is evident from this study.

4.2 Flexible Sensors Electronics for Laryngeal Ambulatory Monitoring

4.2.1 Introduction

Though HDsEMG provides the necessary spatiotemporal data to non-invasively monitor complex muscular regions, its applications are limited by the amount of time needed to prepare the multi-electrode array (MEA), large and costly signal acquisition systems, and complex data

analysis [37, 157]. In rehabilitory and clinical settings, the limitations may be compensated by the benefits of HDsEMG. However, in emergency and ambulatory settings, clinicians and patients are limited by time, and therefore require a ready-to-use MEA. A solution may involve having assistants prepare the arrays before hand, or having the arrays be mass-produced by a medical company. Unfortunately, this solution falls short in ambulatory monitoring. A system capable of at-home-monitoring should be easy for the patient to use. Additionally, if we want to ensure compliance, the monitoring process should be as comfortable and unobtrusive as possible. In its current stage, both the machine and arrays are bulky and complicated to use. Moreover, the cables connected to the arrays are heavy and pull at the skin, making them obtrusive for everyday use. Though HDsEMG is relatively more comfortable and easy-to-use than nEMG, the system can be significantly improved with Flexible Electronic Sensors (FES) and smaller signal acquisition devices.

Figure 4.9: The BrainVision shown to the left is biopotential system made especially for electroencephalography monitoring. This system was used to record the data in Chapters 2 and Section 3.1. To the right is the TMSi REFA system made to record various types of biopotential signals. This device was used to record data from Section 3.2.

FES are thin electrodes that are developed over Polydimethylsiloxane (PDMS) through a process of metal deposition and photolithography. This microfabrication process creates thin and

stretchable electrodes which adhere to the curvature of the skin and provider greater user-comfort. In recent years, fabrication of flexible sensors has improved, making them a feasible and scalable option for the near future [73, 77]. Through custom-made layouts using micro-fabrication, we can generate various designs to conform to an individual's anatomy. This allows for a customizable electrode interface at a relatively low cost. Fig. 4.10 is of two types of FES made in Dr. Todd Coleman's Neural Interaction Lab (Department of Bioengineering, UC San Diego, La Jolla, California).

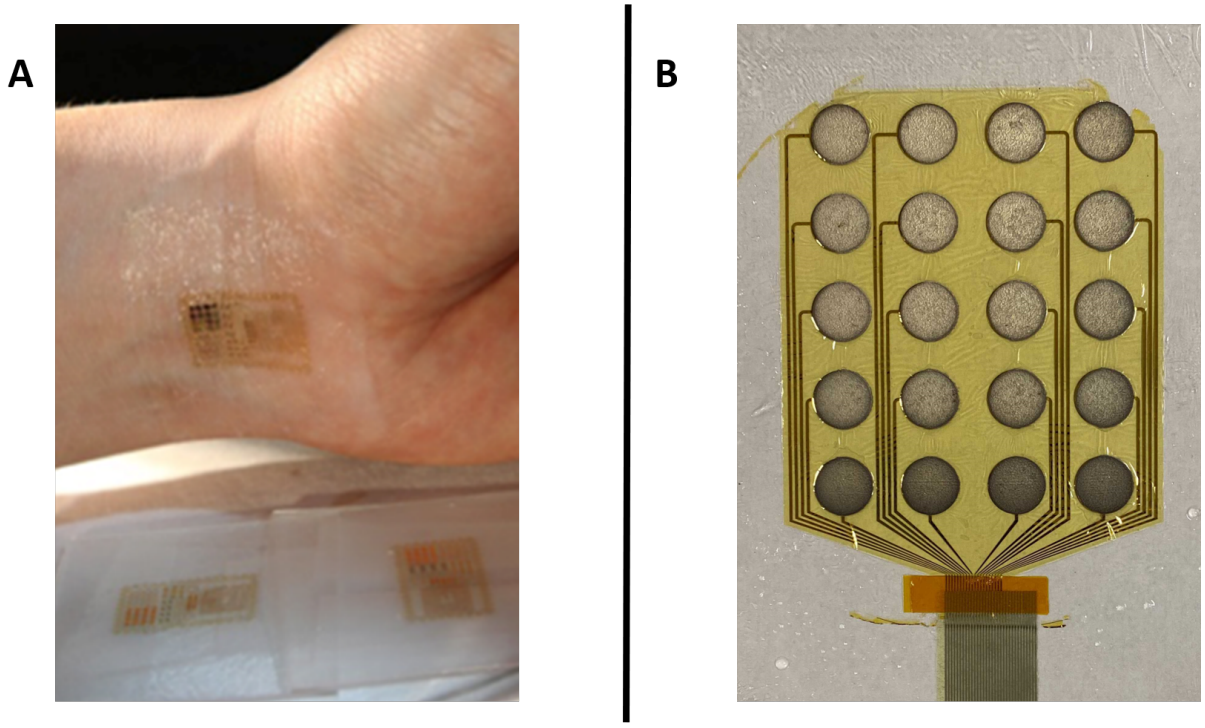

Figure 4.10: Picture A depicts a flexible electronic sensor from Kang et al. [73]. Picture B is of the FES customized to replicate the MEA used in Chapter 2 for the neck.

This dissertation has used five different types of biopotential signal acquisition systems: BrainVision (Chapters 2 and Section 3.1), TMSi REFA (Section 3.2), BIOPAC MP150 (Section 4.1), OpenBCI Cyton (Section 4.1), and WANDmini (Section 4.1). In fig. 4.9, the BrainVision and TMSi are shown. These two systems have a wide sampling range and are exceptional for monitoring a variety of biopotential signals with a high number of channels (>20-channels). A

setback though is in their high-cost (>$ 30,000.00), bulkiness, and complicated user-interface. The BIOPAC device is similar in terms of cost, but is further limited by the low number of channels it can record from (<10 channels), and even bigger size– requiring a cart to move the many chained channel amplifiers. The OpenBCI Cyton is an ideal size, however its sampling rate is limited to 250 Hz using eight channels. The sampling rate drops to 125 Hz if recording from eight to 16 channels. Though its relatively affordable at <$1,500.00, the low sampling rate and number of channels does not make it an ideal solution for laryngeal monitoring. The WANDmini, a system in development by Dr. Rikky Muller's research lab (Muller Lab, Department of Electrical Engineering and Computer Sciences, UC Berkeley, California, USA) boasts a sampling rate of 1024 Hz and a 64-channel capacity. This system is derived from their previous design [74, 150] and was a suitable fit for this project.

This project investigates the use of FES coupled with a miniaturized signal acquisition system for HDsEMG in the neck. Unfortunately, COVID-19 prevented the enrollment of research subjects and therefore this pilot study was only performed on a single subject.

4.2.2 Materials and Methods

Ethical approval for this study was granted by the Institutional Review Board at the University of California, San Diego (IRB # 161477). The pilot experiment was conducted on a single subject, age 27. Inclusion criteria was greater than 18 years of age. Exclusion criteria included a history of dysphagia, dysphonia, previous neck surgery, or neurological illness.

To minimize strap muscle activity at rest, all participants were seated in a semi-reclined position ($\sim$ 20 degrees from vertical) with the neck in slight extension and supported. The cricothyroid (CT) space was identified with digital palpation and marked as a reference point for grid orientation. Skin preparation with standard alcohol wipes and exfoliating impedance-reduction gel (NuPrep Skin Prep Gel, Weaver and Company, Aurora, Colorado, USA) was performed across the entire anterior neck.

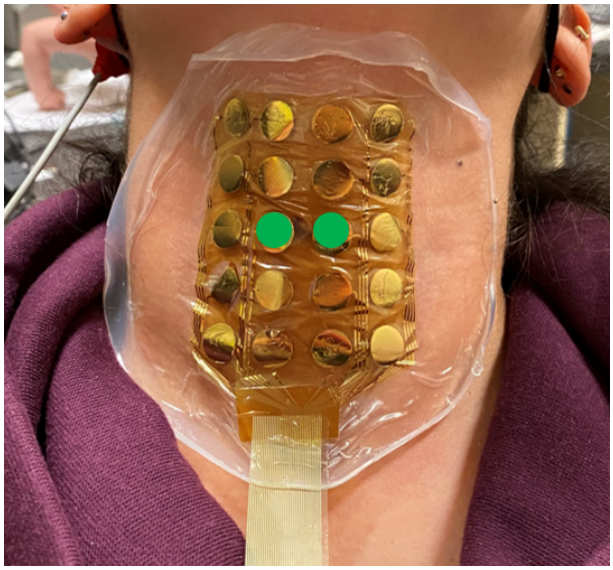

Figure 4.11: This figure shows the customized FES HDsEMG array over the neck. Channels 10 and 11 were placed over the CT muscle and are marked by the green dots.

The high density 20-channel FES array was composed of silver/silver chloride and was fabricated using the same process depicted in a recently published method to fabricate stretchable silver/silver-chloride multi-electrode arrays by our research group [83]. For this application, the FES were organized into 5 rows and 4 columns with an electrode spacing of 1.5 cm between each electrode center (electrode diameter 1 cm). The array was centered on the cricothyroid (CT) space, with electrodes 10 and 11 overlying the CT space, marked by green circles in fig. 4.11.

Signal Acquisition

The HDsEMG array used a research device called the WANDmini (Muller Lab, Department of Electrical Engineering and Computer Sciences, UC Berkeley, California, USA). The sampling rate was 1024 Hz and used an average reference from all the channels when recording. A ground electrode was placed on the left mastoid bone behind the ear (3M Red Dot (REF: 2670-5), St. Paul, Minnesota, USA). An audio prompt was played over a 5-minute period for each recorded task. Tasks included the following: (1) rest, (2) low-pitch phonation, (3) high-pitch phonation. The subject was able to demonstrate adequate difference between low and high pitch as confirmed

by the authors with frequency analysis (Audio Frequency Counter; Keuwlsoft, London, United Kingdom). Each task was paced to afford 5 seconds of phonation, on an /i/vowel, followed by 10 seconds of rest. This was repeated for a total of 5 minutes, resulting in 20 recorded intervals of each phonatory task for each participant. The data was analyzed via custom software written in Python programming language (Python Software Foundation, Python Language Reference, v. 3.7.1).

Signal Processing

The data was processed using the same workflow described in Chapter 2. In particular, the goal was to compare the ROC results using the log likelihood ratio test on the FES data, with the results from Section 2.1. Though there is a significant difference in the number of subjects, there is a merit in understanding the potential of FES when compared to a traditional HDsEMG set-up.

4.2.3 Results

Typical preparation of the traditional HDsEMG array from Chapter 2 took about 45 minutes per array. This included cutting the electrodes, placing them into their orientation, and placing the 3M Tegaderm over it then cutting to allow a connection with the DIN snap electrode leads. This did not include the time spent connecting the electrodes to their individual cables. This time was not recorded as they were connected while the subject's skin was being prepared. Preparing the FES took about 20 minutes. This involved connecting the electrodes to the system, spraying the FES array with a skin-safe medical adhesive, and then lifting it from its substrate to place on the skin. Placement of the FES on the skin was much easier since in the old array, the DIN snap leads connected to the electrodes were a visual obstruction in placing the electrodes over their designated landmarks.

The 20-Channel FES HDsEMG array and the miniaturized signal acquisition system provided EMG data during two types of phonation: low- and high-pitch. Figure 4.12 presents the

two types of phonation and their processed time-series data. Channels 1, 14, and 20 appeared to have abnormally high peaks. During low-pitch activity, channels 9 and 13 recorded the clearest oscillatory pattern we expected during phonation. In high-pitch phonation, this pattern was detected in channels 9, 13, and 19.

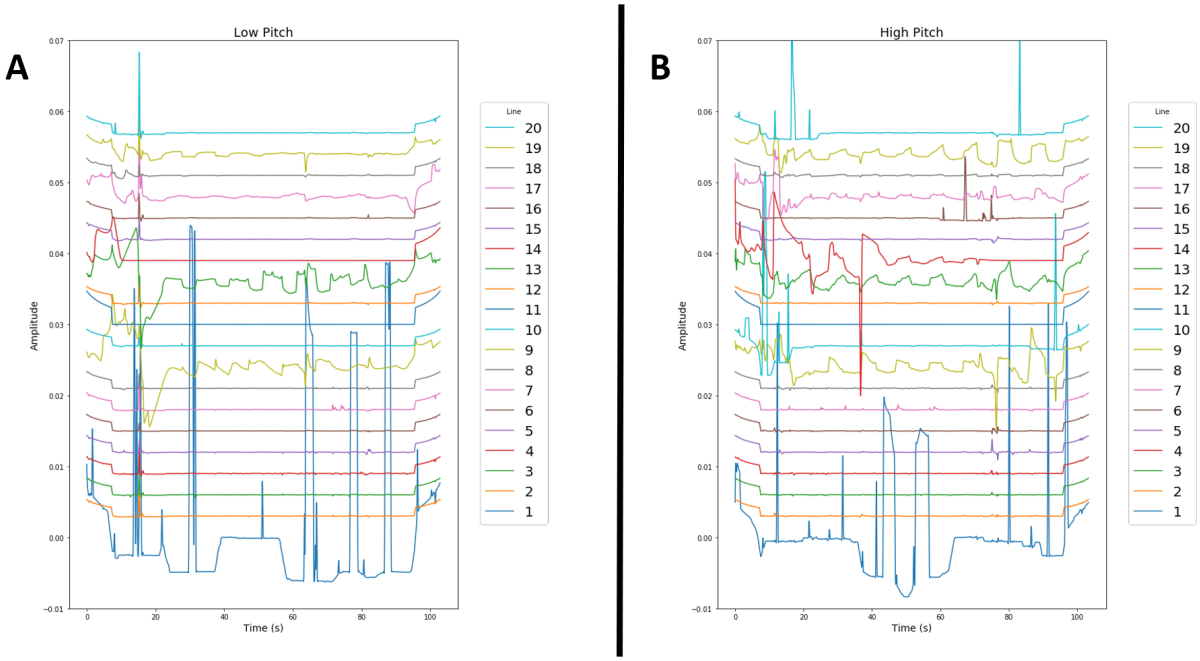

Figure 4.12: Both plots depict the entirety of the low- and high-pitch data recorded from the FES and miniaturized system after processing.

An energy map during a single time-point of high-pitch phonation can be seen in fig. 4.13 side A. Greater amplitudes were seen in the upper rows of the array throughout all phonation segments. However unlike the previous dataset, the energy seemed skewed towards the sides instead of the center. Applying the LLR method from Ch. 2.2 produced an ROC curve shown in fig. 4.13, side B. The calculated AUC was 0.73.

Removal of the array was relatively less painful than the traditional array and much quicker to clean up (<5 minutes). Readying the system for the next experiment was significantly faster as well. In the traditional array set-up, an additional 10 minutes were spent disconnecting the electrodes and wiping down the cables for the next experiment. If the traditional HDsEMG

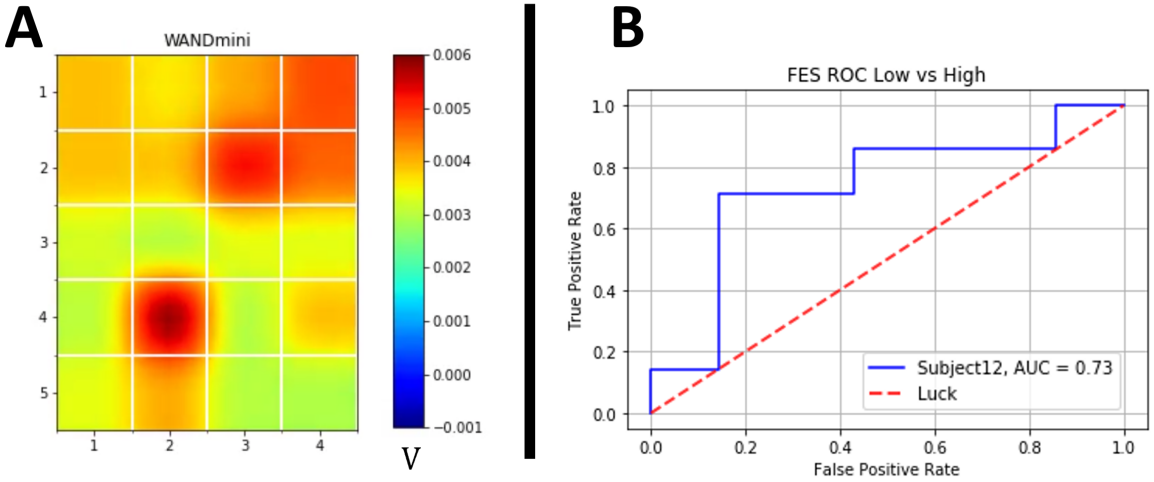

Figure 4.13: A: This is the energy map of the high-pitch amplitudes across all channels during a single time-point. Note how it is skewed to the sides. B: This is the ROC curve between low and high-pitch phonation. The AUC was 0.73

arrays were pre-made, the wait time between experimental subjects was about 30 minutes. For the FES and miniaturized system, the wait time reduced to about 15-20 minutes, depending on the user's dexterity. This is due to the fact that the FES electrodes are wired to one connector, so disconnecting from the system simply involved pulling a single connector, instead of individual wires.

4.2.4 Discussion

The 20-channel FES coupled with the miniaturized signal acquisition system was capable of monitoring the neck during low and high-pitch phonation. Section 2.2, section Signal Processing, discusses the presence of oscillatory patterns that appeared to have larger amplitudes in areas near the laryngeal muscles during phonation (Fig. 2.8). Additionally, the channels detecting these patterns varied between the two types of phonation. Taking advantage of these properties, an LLR was used to compare the two distributions and solve for an ROC curve. This same method was applied to the data recording from the FES. In the processed time-series data in Figure 4.12,

the phonation segments are evident. The AUC of the ROC curve between low and high-pitch is 0.73 for a single subject, much lower than the AUC of Section 2.2 of 0.97 for the full array. Its also 0.04 lower than the AUC obtained from only two channels.

This brings about questions regarding the utility of an FES array. Given these two AUCs of 0.73 (FES) and 0.77 (2-ch traditional electrodes), it may seem easy to simply say a 2-ch system is more convenient than an FES HDsEMG array given the complexity of its fabrication and lack of availability in the clinic. However its important to take a step back and understand the big picture. Yes, the 2-ch traditional set-up produced a higher AUC, but this was the averaged AUC across all subjects' ROC curve. If one looks at the legend of Fig. 2.11's 2-ch ROC, the range of AUCs varies from 0.51 to 0.99. Two subjects had an AUC just above chance. The 0.73 AUC of the FES ROC is of a single subject. Though this is a significant limitation to this pilot study, there may be a possibility that when more subjects are recruited, the averaged AUC of the FES ROC will be competitive with the 2-ch configuration. Moreover with an array, there is a negligible risk of "missing" the laryngeal muscles as compared to only monitoring the muscles with a few channels. Therefore, the possibility of getting an AUC near 0.5 is low in an FES array and should still be considered for monitoring laryngeal EMG non-invasively.

Experimental preparation was an additional benefit to using an FES array. In emergency settings, an FES array can be envisioned as a "patch" that can quickly be placed on the subject and connected to a system. In an ambulatory setting, one can imagine the time it takes to prepare an array composed of traditional electrodes to be unappealing, or even an impossible task, for the elderly or populations dealing with voice and swallow disorders. Moreover, the subject reported the FES array to be significantly more comfortable to wear than the traditional array, describing the FES to feel like "those moisturizing peel-off face masks". They also stated not to feel discomfort, or pull at the skin, from the array cable, as compared to the slight irritation and tug they felt with the old array. The ease-of-use, quick placement and removal, and greater comfort gained through an FES array will likely maintain patient compliance for a laryngeal

ambulatory monitoring system and increase its applicability for a variety of settings.

A limitation to the FES array is its delicateness and more fragile polymore substrate when compared to Tegaderm. Much like when working with the "peel-off face masks", those placing the array must be dexterous and quick, therefore the training period may increase slightly. However this can be addressed with future improvements to the substrate on which the FES is made. Another limitation is the inability to manage or remove faulty electrodes. In a old HDsEMG array, a bad electrode could easily be switched out. This is not possible with the FES array. A bad electrode, if its deemed to be in a critical location (over the laryngeal muscles) would require a replacement of the whole array. This issue is believed to have caused the noisy and drifting in channels 1 and 20 in Fig. 4.12.

The signal acquisition system used in this experiment was simple to set-up, manage, and record from. However, this research device is still in development. It also transmits data via Bluetooth, which is always at risk of signal loss. Despite this set-back, the system is promising, a novel step towards miniaturizing biopotential recording systems, and was still able to detect the oscillatory phonation patterns needed for pitch differentiation.

4.2.5 Conclusion

FES HDsEMG is a far more comfortable experience for both patient and clinicians. It is capable of providing the necessary spatiotemporal information without the constraints of a traditional array. With further developments, FES used in conjunction with a miniaturized signal acquisition system, has the potential to revolutionize the applications of noninvasive laryngeal EMG. Whether in the clinic, therapy, or at home, the ability to accurately measure laryngeal muscles during voice and swallow will help improve treatment decision-making and our understanding of voice and swallow disorders.

4.3 Acknowledgments

The author thanks Dr. Pedro Cabrales (UCSD, Department of Bioengineering) for allowing us to use his BIOPAC system, and Dr. Alex Williams for his guidance in using the system.

Chapter 4, Section 4.1, in part is currently being prepared for submission for publication of the material. Ornelas, Gladys; Ma, Xiaoyi; Rios-Ramirez, Ruby; Weissbrod, Philip A.; Coleman, Todd P., The dissertation author was the primary investigator and first author of this paper.

Chapter 4, Section 4.2, in part is currently being prepared for submission for publication of the material. Ornelas, Gladys; Kurniawan, Jonas F.; Nguyen, Andrew; Pham, Timothy; Sit, Nathan L.J.; Rios-Ramirez, Ruby; Weissbrod, Philip A.; Coleman, Todd P., The dissertation author was the primary investigator and first author of this paper.

Chapter 5

Conclusion

Voice and swallow disorders can arise from a myriad of reasons such as overuse in educators and singers, nerve injury, neuromuscular disorders, stroke, cancer, or aging. Unfortunately, the complexity of such origins and anatomy of the throat means that diagnosis is time-consuming, costly, and uncomfortable, relying on subjective and invasive techniques. When possible, treatment for voice or swallow disorders tackles the underlying condition. However in instances such as aging or neuromuscular disorders, treatment may be limited to blocking the dysfunctional muscle through Botox, or managing symptoms through voice and swallow therapy. In the elderly, these solutions may not be possible depending on other limitations such as expense, time, traveling difficulties, or prioritization of other health issues. Moreover in swallow disorders, Ekberg et al. found that 39% of those surveyed thought that dysphagia was untreatable [40]. As a result, some elderly may resign themselves to accept their speech and swallow difficulties, leading to under-treatment and potentially increasing the risks of developing further complications. In children, studying these disorders via traditional methods is difficult due to their fear of invasive equipment, or intolerance to pain and discomfort. Lastly, not all those with voice and swallowing problems are affected enough to seek medical attention [135]. In neuromuscular disorders such as Parkinson's disease, early diagnosis leads to an improved quality of life compared to those who

were diagnosed later [28, 45]. Therefore ignoring symptoms of voice and swallow dysfunction, which could be indicative of a worst underlying condition, can worsen a patient's prognosis. Even primary care physicians may not understand the significance of such symptoms, or may be dealing with multiple medical issues in their elderly patients that they may not discuss voice and swallowing problems, and fail to provide the patient with proper early-care intervention [135].

There is a need for improved screening tools and better education of voice and swallow disorders in both the general public and primary care physicians. These tools should be capable of monitoring the symptoms over time and during treatment to understand the outcomes for voice and swallowing problems. This thesis presents the use of HDsEMG for monitoring voice and swallow disorders. HDsEMG overcomes some of the limitations encountered by traditional sEMG through a multi-electrode array (MEA). In providing spatiotemporal data, clinicians are reassured that one of the many electrodes is capturing the muscle of interest and not cross-talk contamination from neighboring muscles. Improvements in MEA analysis have increased HDsEMG's applications and widened the breath of information they can detect [37].

Chapter 2 of this thesis presented high- and low-pitch phonation data recorded from a 20-channel MEA in 10 and 11 subjects. In the first half of the chapter, HDsEMG recorded lateral symmetric EMG activity, reflective of the laryngeal muscles' orientation. Additionally, there were visual differences in the time-series data after processing. Energy maps of the amplitudes also presented local active regions hypothesised to be the laryngeal muscles. The second half of this chapter confirmed the detection of pitch differences through HDsEMG using the log likelihood ratio (LLR) hypothesis testing. In solving for the distributions of low and high-pitch data, we were able to train and test the LLR across the 20 phonation trials for each channel and subject. The resulting receiver operating curve (ROC) had an area under the curve (AUC) of 0.97. To compare against a traditional SEMG set-up, only the two electrodes on one side placed over the cricothyroid muscle (CT), were used for the LLR. The AUC of the ROC dropped to 0.77. There was also an increased variance across subjects and two of the subjects had a AUC just above

chance. These results illustrate the issues with traditional sEMG in which a few electrodes can miss the target muscles, therefore experience a greater risk of cross-talk contamination, and be more affected by low signal quality. Limitations in this study include a small subject population, and no data from unhealthy subjects. It is important to verify these results with the gold-standard of invasive EMG. Nonetheless, HDsEMG has the potential to not only revolutionize how it is used in the clinical and therapy, but when combined with acoustic analysis, could improve applications of speech recognition through deep learning models (e.g. LSTM) [6, 70].

In deep learning models, different types of variables can be simultaneously used as model inputs. Therefore future studies could combine time-varying HDsEMG features with voice recording waveforms for binary or multi-class characterization of vocal dysfunction. In addition, features learned in pathology characterization may be utilized, through transfer learning, to improve existing voice recognition paradigms (e.g. Apple Siri, Google Alexa). Transfer learning involves the use of information learned in one classification task for a separate but related task, typically enabling breakthrough performance [6, 54, 70], and in this case addressing the vulnerabilities in existing voice recognition technology. Furthermore, state space modeling can characterize relationships between features of voice recordings and the underlying latent state, HDsEMG [79, 113]. With these relationships, voice classification tasks can be performed even when only voice recording data is available, thus widely extending the reach of our applied methodology.

Chapter 3 investigated the use of HDsEMG for swallow analysis. Section 3.1 studied "dry" or saliva swallows performed by the same cohort of chapter 2. Results indicated the presence of symmetric activity, much like in voice, as well as higher energy in the upper rows of the array. The second half of chapter 3, section 3.2, investigated swallow classification of five different types of swallows. This was achieved with an average probability error of 0.16. This error was higher for complex swallows such as fruit cup and cracker swallows. The error was much lower on simple swallows such as water and apple sauce. These results give credence towards

a potential application in the swallow therapeutic realm. Currently there is a lack of objective assessment tools that can determine the success of swallow therapy across various swallow disorders. Attempts have been made through sEMG as a biofeedback tool, but the results are susceptible to bias and lack standardization. Though limited by a small population size, the swallow classification results demonstrate speech and language pathologists could use HDsEMG during therapy to better understand a subject's swallow disorder and treatment progression. If made portable to be used at home, such a system could monitor patient compliance to their therapy exercises, and provide continuous information about their symptoms and response to various types of food. This knowledge will help in treatment decision making, and will shed light on the outcomes of treatment for voice and swallow disorders.

Chapter 4 aimed to validate the HDsEMG laryngeal protocol and begin development of an ambulatory system. Unfortunately the COVID-19 pandemic prevented subject recruitment and restricted laboratory work. Section 4.1 demonstrated my attempt to work around the pandemic by designing a data analysis workflow through forearm EMG. This study placed an HDsEMG array, along with a few needle electrodes, to record the forearm during finger extension. The needles were placed in the muscles associated with pinky, middle finger, and thumb extension. Results demonstrated strong concordance of EMG onset detection between the needle electrode and its nearest surface electrodes. The concordance increased during activation of the target muscle, thus supporting our hypothesis that spatiotemporal data detects localized regional activity. Though we lacked the necessary invasive information from the throat, when applying this same workflow to the phonation dataset from chapter 2, we were able to see some distinction between low and high-pitch phonation counts. The AUC of the ROC was 0.84 and the amplitude of the onset counts were smaller ($<$20). This may be due to the small size of the laryngeal muscles, significant overlap of surrounding muscles, and depth of muscles relative to the skin. Therefore additional processing may be required for the laryngeal data. Its important to note the forearm data was from a single subject subject, a major limitation due to COVID-19. Nonetheless, the results still hold

promise: HDsEMG can differentiate low and high-pitch phonation with onset counts in addition to the results from chapter 2. This is novel for applications in unhealthy subjects as we would expect little pitch differentiation or symmetry in the cases of nerve injury and other disorders. It is our hope that future studies will validate these results with needle EMG on the throat.

The second half of chapter 4, section 4.2 focused on the use of Flexible Electronic Sensors (FES) and a miniaturized signal acquisition system, for ambulatory monitoring. Not accounting for the time needed to fabricate the array, the FES were much quicker to prepare, apply, and remove, than the array composed of standard 3M electrodes. During experimentation, the FES was more comfortable and painless than the traditional array, especially upon removal. A limitation to the array was the frailness of the FES. Some training will be required to teach future researchers how to handle the delicateness of the FES, but future refinements on the substrate can mediate this. Other issues include signal loss during Bluetooth transmission of the device, large drift in some of the channels, and an increased susceptibility to movement and electrical noise. The signal loss could be due to Bluetooth and can be avoided through other forms of data transmission such as WiFi, or by recording the data in the device and not transmitting it at all. This of course may be an issue for real-time data analysis, but at the moment that is not a priority. Additionally, the system is a research device in development and is not housed within any container. This may be why the signal quality was lower. Though only restricted to one subject (also due to COVID-19), the AUC of the low and high-pitch differentiation LLR ROC was 0.73. By improving the acquisition system, we predict the ROC will improve and lead to the development of an ambulatory monitoring system. This system can be applied for voice and swallow disorders to disambiguate treatment success and symptom progression.

In conclusion, HDsEMG is a crucial ingredient for recording noninvasive laryngeal activity. It is able to differentiate low and high pitch phonation, as well as classify different types of swallows. Its applications are not restricted to a clinical expert and can be used in a variety of settings. Information from HDsEMG can be used to inform the general public about voice

and swallow disorders through a cost-effective, comfortable, and easy-to-use assessment system. Coupled with novel processing tools, machine learning algorithms, and miniaturized recording devices, HDsEMG will improve our understanding of voice and swallow disorders, as well as speech recognition tools.

Bibliography

[1] U. R. Acharya, H. Fujita, O. S. Lih, Y. Hagiwara, J. H. Tan, and M. Adam. Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network. *Information Sciences*, 405:81–90, 2017.

[2] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. San Tan. A deep convolutional neural network model to classify heartbeats. *Computers in Biology and Medicine*, 89:389–396, 2017.

[3] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, and H. Adeli. Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in Biology and Medicine*, 100:270–278, 2018.

[4] A. S. Agrusa, A. A. Gharibans, A. Allegra, D. C. Kunkel, and T. Coleman. A deep convolutional neural network approach to classify normal and abnormal gastric slow wave initiation from the high resolution electrogastrogram. *IEEE Transactions on Biomedical Engineering*, 2019.

[5] A. S. Agrusa, D. C. Kunkel, and T. P. Coleman. Robust regression and optimal transport methods to predict gastrointestinal disease etiology from high resolution egg and symptom severity. *In Review*, 2021.

[6] L. A. Alexandre. 3d object recognition using convolutional neural networks with transfer learning between input channels. In *Intelligent Autonomous Systems 13*, pages 889–898. Springer, 2016.

[7] U. C. Allard, F. Nougarou, C. L. Fall, P. Giguère, C. Gosselin, F. Laviolette, and B. Gosselin. A convolutional neural network for robotic arm guidance using SEMG based frequency-features. In *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*, pages 2464–2470. IEEE, 2016.

[8] A. A. Alrahim, R. A. Alanazi, and M. H. Al-Bar. Hoarseness among school teachers: A cross-sectional study from dammam. *Journal of family & community medicine*, 25(3):205, 2018.

[9] M. Alsulaiman. Voice pathology assessment systems for dysphonic patients: detection, classification, and speech recognition. *IETE Journal of Research*, 60(2):156–167, 2014.

[10] S. K. Archer, R. Garrod, N. Hart, and S. Miller. Dysphagia in duchenne muscular dystrophy assessed objectively by surface electromyography. *Dysphagia*, 28(2):188–198, 2013.

[11] J. Ashford, D. McCabe, K. Wheeler-Hegland, T. Frymark, R. Mullen, N. Musson, T. Schooling, and C. S. Hammond. Evidence-based systematic review: Oropharyngeal dysphagia behavioral treatments. part iii–impact of dysphagia treatments on populations with neurological disorders. *Journal of Rehabilitation Research & Development*, 46(2), 2009.

[12] N. Audag, C. Goubau, M. Toussaint, and G. Reychler. Screening and evaluation tools of dysphagia in children with neuromuscular diseases: a systematic review. *Developmental Medicine & Child Neurology*, 59(6):591–596, 2017.

[13] N. Audag, C. Goubau, M. Toussaint, and G. Reychler. Screening and evaluation tools of dysphagia in adults with neuromuscular diseases: a systematic review. *Therapeutic advances in chronic disease*, 10:2040622318821622, 2019.

[14] K. K. Baker, L. O. Ramig, E. S. Luschei, and M. E. Smith. Thyroarytenoid muscle activity associated with hypophonia in parkinson disease and aging. *Neurology*, 51(6):1592–1598, 1998.

[15] M. P. Balata, H. J. Silva, L. A. Pernambuco, H. P. Oliveira, and S. R. Moraes. Normalization patterns of the surface electromyographic signal in the phonation. *Journal of Voice*, 29(1): 129.e1–29.e8, 2015.

[16] M. P. Balata et al. Electrical activity of extrinsic laryngeal muscles in subjects with and without dysphonia. *Journal of Voice*, 29(1):129.e9–29.e17, 2015.

[17] P. M. M. Balata, H. J. Silva, G. K. O. Nascimento, K. J. R. Moraes, L. A. Pernambuco, M. C. R. Freitas, L. M. Lima, R. S. Braga, S. R. Souza, and S. R. A. Moraes. Incomplete swallowing and retracted tongue maneuvers for electromyographic signal normalization of the extrinsic muscles of the larynx. *Journal of Voice*, 26(6):813–e1, 2012.

[18] L. A. Bateman. Parallels between singing and phonetic terminology. *Working Papers of the Linguistics Circle*, 15:79–84, 2002.

[19] P. M. Bath, H. S. Lee, and L. F. Everton. Swallowing therapy for dysphagia in acute and subacute stroke. *Cochrane Database of Systematic Reviews*, (10), 2018.

[20] J. K. Benfield, L. F. Everton, P. M. Bath, and T. J. England. Does therapy with biofeedback improve swallowing in adults with dysphagia? a systematic review and meta-analysis. *Archives of physical medicine and rehabilitation*, 100(3):551–561, 2019.

[21] N. Bhattacharyya. The prevalence of dysphagia among adults in the united states. *Otolaryngology–Head and Neck Surgery*, 151(5):765–769, 2014.

[22] N. Bhattacharyya. The prevalence of voice problems among adults in the united states. *The Laryngoscope*, 124(10):2359–2362, 2014.

[23] H. Bogaardt, W. Grolman, and W. Fokkens. The use of biofeedback in the treatment of chronic dysphagia in stroke patients. *Folia Phoniatrica et Logopaedica*, 61(4):200–205, 2009.

[24] R. Boostani and M. H. Moradi. Evaluation of the forearm emg signal features for the control of a prosthetic hand. *Physiological measurement*, 24(2):309, 2003.

[25] D. J. Bracken, G. Ornelas, T. P. Coleman, and P. A. Weissbrod. High-density surface electromyography: A visualization method of laryngeal muscle activity. *The Laryngoscope*, 129(10):2347–2353, 2019.

[26] M. Bryant. Biofeedback in the treatment of a selected dysphagic patient. *Dysphagia*, 6(3): 140–144, 1991.

[27] C. Carreiras. BioSPPy: Biosignal processing in Python, 2015–. URL `https://github.com/PIA-Group/BioSPPy`.

[28] J. J. Chen. Parkinson's disease: health-related quality of life, economic cost, and implications of early treatment. *The American journal of managed care*, 16:S87–93, 2010.

[29] G. Constantinescu, W. Hodgetts, D. Scott, K. Kuffel, B. King, C. Brodt, and J. Rieger. Electromyography and mechanomyography signals during swallowing in healthy adults and head and neck cancer survivors. *Dysphagia*, 32(1):90–103, 2017.

[30] M. A. Crary, G. D. Carnaby, M. E. Groher, and E. Helseth. Functional benefits of dysphagia therapy using adjunctive semg biofeedback. *Dysphagia*, 19(3):160–164, 2004.

[31] M. A. Crary, G. D. Carnaby, L. A. LaGorio, and P. J. Carvajal. Functional and physiological outcomes from an exercise-based dysphagia therapy: a pilot investigation of the mcneill dysphagia therapy program. *Archives of physical medicine and rehabilitation*, 93(7): 1173–1178, 2012.

[32] A. N. Crespo, P. A. Kimaid, A. J. M. Júnior, and A. E. Wolf. Laryngeal electromyography: are the results reproducible? *Journal of Voice*, 29(4):498–500, 2015.

[33] C. J. De Luca and P. Contessa. Hierarchical control of motor units in voluntary contractions. *Journal of neurophysiology*, 107(1):178–195, 2012.

[34] R. Ding, C. R. Larson, J. A. Logemann, and A. 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.

[35] S. H. Doeltgen, E. Ong, I. Scholten, C. Cock, and T. Omari. Biomechanical quantification of mendelsohn maneuver and effortful swallowing on pharyngoesophageal function.

Otolaryngology–Head and Neck Surgery, 157(5):816–823, 2017.

[36] B. Dong and S. Biswas. Analyzing breathing signals and swallow sequence locality for solid food intake monitoring. *Journal of medical and biological engineering*, 36(6): 765–775, 2016.

[37] G. Drost, D. F. Stegeman, B. G. van Engelen, and M. J. Zwarts. Clinical applications of high-density surface emg: a systematic review. *Journal of Electromyography and Kinesiology*, 16(6):586–602, 2006.

[38] P. Du, S. Calder, T. R. Angeli, S. Sathar, N. Paskaranandavadivel, G. O'Grady, and L. K. Cheng. Progress in mathematical modeling of gastrointestinal slow wave abnormalities. *Frontiers in physiology*, 8:1136, 2018.

[39] R. Edwin. Cross training for the voice. *Journal of Singing*, 65(1):73, 2008.

[40] O. Ekberg, S. Hamdy, V. Woisard, A. Wuttge-Hannig, and P. Ortega. Social and psychological burden of dysphagia: its impact on diagnosis and treatment. *Dysphagia*, 17(2): 139–146, 2002.

[41] C. Ertekin. Electrophysiological techniques to evaluate swallowing in central and peripheral nervous system disorders. *Journal of Clinical Neurophysiology*, 32(4):314–323, 2015.

[42] D. Farina, W. Muhammad, E. Fortunato, O. Meste, R. Merletti, and H. Rix. Estimation of single motor unit conduction velocity from surface electromyogram signals detected with linear electrode arrays. *Medical and Biological Engineering and Computing*, 39(2): 225–236, 2001.

[43] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya. Deep learning for healthcare applications based on physiological signals: a review. *Computer Methods and Programs in Biomedicine*, 2018.

[44] D. O. Francis, A. E. Sherman, K. L. Hovis, K. Bonnet, D. Schlundt, C. G. Garrett, and L. Davies. Life experience of patients with unilateral vocal fold paralysis. *JAMA Otolaryngology - Head and Neck Surgery*, 144(5):433–439, 2015.

[45] H. Gage, A. Hendricks, S. Zhang, and L. Kazis. The relative health related quality of life of veterans with parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 74(2):163–169, 2003.

[46] R. G. Gallager. *Principles of Digital Communication*. Cambridge University Press, 2008.

[47] T. Gay, M. Strome, H. Hirose, and M. Sawashima. Electromyography of the intrinsic laryngeal muscles during phonation. *Annals of Otology, Rhinology & Laryngology*, 81(3): 401–409, 1972.

[48] M. Gazzoni, D. Farina, and R. Merletti. A new method for the extraction and classification

of single motor unit action potentials from surface emg signals. *Journal of neuroscience methods*, 136(2):165–177, 2004.

[49] A. A. Gharibans, S. Kim, D. C. Kunkel, and T. P. Coleman. High-resolution electrogastrogram: a novel, noninvasive method for determining gastric slow-wave direction and speed. *IEEE Transactions on Biomedical Engineering*, 64(4):807–815, 2017.

[50] A. A. Gharibans, S. Kim, D. C. Kunkel, and T. P. Coleman. High-resolution electrogastrogram: a novel, noninvasive method for determining gastric slow-wave direction and speed. *IEEE Transactions on Biomedical Engineering*, 64(4):807–815, 2017.

[51] A. A. Gharibans, B. L. Smarr, D. C. Kunkel, L. J. Kriegsfeld, H. M. Mousa, and T. P. Coleman. Artifact rejection methodology enables continuous, noninvasive measurement of gastric myoelectric activity in ambulatory subjects. *Scientific Reports*, 8(1):5019, 2018.

[52] A. A. Gharibans, T. P. Coleman, H. Mousa, and D. C. Kunkel. Spatial patterns from high-resolution electrogastrography correlate with severity of symptoms in patients with functional dyspepsia and gastroparesis. *Clinical Gastroenterology and Hepatology*, 2019.

[53] M. Gonzalez-Fernandez and S. K. Daniels. Dysphagia in stroke and neurologic. *Phys Med Rehabil Clin N Am*, 19:867–888, 2008.

[54] V. Guedes, F. Teixeira, A. Oliveira, J. Fernandes, L. Silva, A. Junior, and J. P. Teixeira. Transfer learning with audioset to voice pathologies identification in continuous speech. *Procedia Computer Science*, 164:662–669, 2019.

[55] G. M. Guindi, T. W. Higenbottam, and J. K. Payne. A new method for laryngeal electromyography. *Clinical Otolaryngology & Allied Sciences*, 6(4):271–278, 1981.

[56] G. L. Hands and C. E. Stepp. Effect of age on human–computer interface control via neck electromyography. *Interacting with computers*, 28(1):47–54, 2016.

[57] Y. Heman-Ackah, S. Mandel, R. Manon-Espaillat, M. Abaza, and R. Sataloff. Laryngeal electromyography. *Otolaryngologic clinics of North America*, 40(5):1003–1023, 2007.

[58] Y. D. Heman-Ackah and A. Barr. The value of laryngeal electromyography in the evaluation of laryngeal motion abnormalities. *Journal of Voice*, 20(3):452–460, 2006.

[59] Y. D. Heman-Ackah and A. Barr. The value of laryngeal electromyography in the evaluation of laryngeal motion abnormalities. *Journal of Voice*, 20(3):452–460, 2006.

[60] D. N. Henrich. Mirroring the voice from garcia to the present day: Some insights into singing voice registers. *Logopedics Phoniatrics Vocology*, 31(1):3–14, 2006.

[61] A. D. Hillel. The study of laryngeal muscle activity in normal human subjects and in patients with laryngeal dystonia using multiple fine-wire electromyography. *Laryngoscope*, 111(S97):1–47, 2009.

[62] A. D. Hillel, L. R. Robinson, and P. Waugh. Laryngeal electromyography for the diagnosis and management of swallowing disorders. *Otolaryngology–Head and Neck Surgery*, 116 (3):344–348, 1997.

[63] H. Hirose and T. Gay. The activity of the intrinsic laryngeal muscles in voicing control. *Phonetica*, 25(3):140–164, 1972.

[64] H. Hirose and T. Gay. Laryngeal control in vocal attack: An electromyographic study. *Haskins Laboratories Status Report on Speech Research*, SR-29/30:49–60, 1972.

[65] H. Hirose, T. Gay, and M. Strome. Electrode insertion techniques for laryngeal electromyography. *The Journal of the Acoustical Society of America*, 50(6A):1449–1550, 1971.

[66] I. Hiroto, M. Hirano, Y. Toyozumi, and T. Shin. Electromyographic investigation of the intrinsic laryngeal muscles related to speech sounds. *Annals of Otology, Rhinology & Laryngology*, 76(4):861–872, 1967.

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

[68] C. J. Houtman, D. F. Stegeman, J. Van Dijk, and M. Zwarts. Changes in muscle fiber conduction velocity indicate recruitment of distinct motor unit populations. *Journal of Applied Physiology*, 95(3):1045–1054, 2003.

[69] D. M. Hull. Thyroarytenoid and cricothyroid muscular activity in vocal register control. 2013.

[70] S. Hussein, K. Cao, Q. Song, and U. Bagci. Risk stratification of lung nodules using 3d cnn-based multi-task learning. In *International conference on information processing in medical imaging*, pages 249–260. Springer, 2017.

[71] I. N. Jacobs and R. S. Finkel. Laryngeal electromyography in the management of vocal cord mobility problems in children. *The Laryngoscope*, 112(7):1243–1248, 2002.

[72] H. Jongprasitkul and W. Kitisomprayoonkul. Effectiveness of conventional swallowing therapy in acute stroke patients with dysphagia. *Rehabilitation research and practice*, 2020, 2020.

[73] D. Y. Kang, Y.-S. Kim, G. Ornelas, M. Sinha, K. Naidu, and T. P. Coleman. Scalable microfabrication procedures for adhesive-integrated flexible and stretchable electronic sensors. *Sensors*, 15(9):23459–23476, 2015.

[74] R. Kaveh, J. Doong, A. Zhou, C. Schwendeman, K. Gopalan, F. L. Burghardt, A. C. Arias, M. M. Maharbiz, and R. Muller. Wireless user-generic ear eeg. *IEEE Transactions on Biomedical Circuits and Systems*, 14(4):727–737, 2020.

[75] S. M. Khoddami, S. Talebian, F. Izadi, and N. N. Ansari. Validity and reliability of surface electromyography in the assessment of primary muscle tension dysphonia. *Journal of Voice*, 31(3):386–e9, 2017.

[76] D.-H. Kim, N. Lu, R. Ma, Y.-S. Kim, R.-H. Kim, S. Wang, J. Wu, S. M. Won, H. Tao, A. Islam, et al. Epidermal electronics. *science*, 333(6044):838–843, 2011.

[77] Y.-S. Kim, J. Lu, B. Shih, A. Gharibans, Z. Zou, K. Matsuno, R. Aguilera, Y. Han, A. Meek, J. Xiao, et al. Scalable manufacturing of solderable and stretchable physiologic sensing systems. *Advanced Materials*, 29(39):1701312, 2017.

[78] P. A. Kimaid, E. M. Quagliato, A. N. Crespo, A. Wolf, M. A. Viana, and L. A. Resende. Laryngeal electromyography in movement disorders: preliminary data. *Arquivos de neuro-psiquiatria*, 62(3A):741–744, 2004.

[79] G. Kitagawa and W. Gersch. *Smoothness priors analysis of time series*, volume 116. Springer Science & Business Media, 1996.

[80] K. A. Kochis-Jennings, E. M. Finnegan, H. T. Hoffman, S. Jaiswal, and D. Hull. Cricothyroid muscle and thyroarytenoid muscle dominance in vocal register control: preliminary results. *Journal of Voice*, 28(5):652–e21, 2014.

[81] L. M. Kopf, C. Jackson-Menaldi, A. D. Rubin, J. Skeffington, E. J. Hunter, M. D. Skowronski, and R. Shrivastav. Pitch strength as an outcome measure for treatment of dysphonia. *Journal of Voice*, 31(6):691–696, 2017.

[82] J. Kreiman, B. R. Gerratt, G. B. Kempster, A. Erman, and G. S. Berke. Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *Journal of Speech, Language, and Hearing Research*, 36(1):21–40, 1993.

[83] J. F. Kurniawan, B. Tjhia, V. M. Wu, A. Shin, N. L. Sit, T. Pham, A. Nguyen, C. Li, R. Kumar, M. Aguilar-Rivera, et al. An adhesive-integrated stretchable silver-silver chloride electrode array for unobtrusive monitoring of gastric neuromuscular activity. *Advanced Materials Technologies*, 6(5):2001229, 2021.

[84] P. Lindestad, B. Fritzell, and A. Persson. Evaluation of laryngeal muscle function by quantitative analysis of the emg interference pattern. *Acta oto-laryngologica*, 109(5–6): 467–472, 1909.

[85] J. A. Logemann. Screening, diagnosis, and management of neurogenic dysphagia. In *Seminars in neurology*, volume 16, pages 319–327. © 1996 by Thieme Medical Publishers, Inc., 1996.

[86] L. W. Lopes, L. B. Simões, J. D. da Silva, D. da Silva Evangelista, A. C. d. N. e Ugulino, P. O. C. Silva, and V. J. D. Vieira. Accuracy of acoustic analysis measurements in the evaluation of patients with different laryngeal diagnoses. *Journal of Voice*, 31(3):382–e15,

2017.

[87] K. A. Ludwig, R. M. Miriani, N. B. Langhals, M. D. Joseph, D. J. Anderson, and D. R. Kipke. Using a common average reference to improve cortical neuron recordings from microelectrode arrays. *Journal of neurophysiology*, 101(3):1679–1689, 2009.

[88] S. Marmor, K. J. Horvath, K. O. Lim, and S. Misono. Voice problems and depression among adults in the united states. *The Laryngoscope*, 126(8):1859–1864, 2015.

[89] T. Masuda and T. Sadoyama. The propagation of single motor unit action potentials detected by a surface electrode array. *Electroencephalography and clinical Neurophysiology*, 63(6): 590–598, 1986.

[90] S. J. McCoy. *Your voice, an inside view: Multimedia voice science and pedagogy*. Inside View Press, 2004.

[91] G. H. McCullough, E. Kamarunas, G. C. Mann, J. W. Schmidley, J. A. Robbins, and M. A. Crary. Effects of mendelsohn maneuver on measures of swallowing duration post stroke. *Topics in stroke rehabilitation*, 19(3):234–243, 2012.

[92] D. L. Menkes and R. Pierce. Needle emg muscle identification: A systematic approach to needle emg examination. *Clinical neurophysiology practice*, 4:199–211, 2019.

[93] R. Merletti, D. Farina, and M. Gazzoni. The linear electrode array: a useful tool with many applications. *Journal of Electromyography and Kinesiology*, 13(1):37–47, 2003.

[94] P. W. Mirowski, Y. LeCun, D. Madhavan, and R. Kuzniecky. Comparing SVM and convolutional networks for epileptic seizure prediction from intracranial EEG. In *Machine Learning for Signal Processing, 2008. MLSP 2008. IEEE Workshop on*, pages 244–249. IEEE, 2008.

[95] E. Mullin. Why siri won't listen to millions of people with disabilities, 2016. URL `https://www.scientificamerican.com/article/why-siri-won-t-listinn-to-millions-of-people-with-disabilities/`.

[96] M. C. Munin, T. Murry, and C. A. Rosen. Laryngeal electromyography: diagnostic and prognostic applications. *Otolaryngologic Clinics of North America*, 33(4):759–770, 2000.

[97] S. Neumann. Swallowing therapy with neurologic patients: results of direct and indirect therapy methods in 66 patients suffering from neurological disorders. *Dysphagia*, 8(2): 150–153, 1993.

[98] D. Park, H. H. Lee, S. T. Lee, Y. Oh, J. C. Lee, K. W. Nam, and J. S. Ryu. Normal contractile algorithm of swallowing related muscles revealed by needle emg and its comparison to videofluoroscopic swallowing study and high resolution manometry studies: a preliminary study. *Journal of Electromyography and Kinesiology*, 36:81–89, 2017.

[99] D. Patel, S. Krishnaswami, E. Steger, E. Conover, M. Vaezi, M. Ciucci, and D. Francis. Economic and survival burden of dysphagia among inpatients in the united states. *Diseases of the Esophagus*, 31(1):dox131, 2018.

[100] R. R. Patel, S. N. Awan, J. Barkmeier-Kraemer, M. Courey, D. Deliyski, T. Eadie, D. Paul, J. G. Švec, and R. Hillman. Recommended protocols for instrumental assessment of voice: American speech-language-hearing association expert panel to develop a protocol for instrumental assessment of vocal function. *American Journal of Speech-Language Pathology*, 27(3):887–905, 2018.

[101] B. Pourbabaee, M. J. Roshtkhari, and K. Khorasani. Deep convolutional neural networks and learning ECG features for screening paroxysmal atrial fibrillation patients. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, (99):1–10, 2017.

[102] J. Rafiee, M. Rafiee, F. Yavari, and M. Schoen. Feature extraction of forearm emg signals for prosthetics. *Expert Systems with Applications*, 38(4):4058–4067, 2011.

[103] J. Ramírez, D. Rodriquez, F. Qiao, J. Warchall, J. Rye, E. Aklile, A. S.-C. Chiang, B. C. Marin, P. P. Mercier, C. Cheng, et al. Metallic nanoislands on graphene for monitoring swallowing activity in head and neck cancer patients. *ACS nano*, 12(6):5913–5922, 2018.

[104] Y. Ren and Y. Wu. Convolutional deep belief networks for feature extraction of EEG signal. In *Neural Networks (IJCNN), 2014 International Joint Conference on*, pages 2850–2853. IEEE, 2014.

[105] V. V. Ribeiro, A. G. de Oliveira, J. da Silva Vitor, A. C. Ramos, A. G. Brasolotto, and K. C. A. Silverio. Effectiveness of voice therapy associated with electromyographic biofeedback in women with behavioral dysphonia: randomized placebo-controlled double-blind clinical trial. *Journal of Voice*, 33(3):381–e11, 2019.

[106] N. Roy, R. M. Merrill, S. D. Gray, and E. M. Smith. Voice disorders in the general population: prevalence, risk factors, and occupational impact. *The Laryngoscope*, 115(11): 1988–1995, 2005.

[107] N. Roy, J. Stemple, R. M. Merrill, and L. Thomas. Dysphagia in the elderly: preliminary evidence of prevalence, risk factors, and socioemotional effects. *Annals of Otology, Rhinology & Laryngology*, 116(11):858–865, 2007.

[108] M. Saconato, B. M. Chiari, H. M. Lederman, and M. I. R. Gonçalves. Effectiveness of chin-tuck maneuver to facilitate swallowing in neurologic dysphagia. *International archives of otorhinolaryngology*, 20(1):13–17, 2016.

[109] K. Saitou, T. Masuda, D. Michikami, R. Kojima, and M. Okada. Innervation zones of the upper and lower limb muscles estimated by using multichannel surface emg. *Journal of human ergology*, 29(1-2):35–52, 2000.

[110] T. S. Saponas, D. S. Tan, D. Morris, and R. Balakrishnan. Demonstrating the feasibility of using forearm electromyography for muscle-computer interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 515–524, 2008.

[111] R. Sataloff, S. Mandel, E. Mann, and C. Ludlow. Practice parameter: laryngeal electromyography (an evidence-based review). *Otolaryngology-Head and Neck Surgery*, 130: 770–779, 2004.

[112] R. T. Sataloff et al. Laryngeal electromyography: clinical application. *Journal of Voice*, 24 (2):228–234, 2010.

[113] G. Schamberg, D. Ba, and T. P. Coleman. A modularized efficient framework for non-markov time series estimation. *IEEE Transactions on Signal Processing*, 66(12):3140–3154, 2018.

[114] G. Schlotthauer, M. E. Torres, and M. C. Jackson-Menaldi. A pattern recognition approach to spasmodic dysphonia and muscle tension dysphonia automatic classification. *Journal of voice*, 24(3):346–353, 2010.

[115] E. Schulte, D. Farina, G. Rau, R. Merletti, and C. Disselhorst-Klug. Single motor unit analysis from spatially filtered surface electromyogram signals. part 2: conduction velocity estimation. *Medical and Biological Engineering and Computing*, 41(3):338–345, 2003.

[116] C. Schultheiss, T. Schauer, H. Nahrstaedt, and R. O. Seidl. Automated detection and evaluation of swallowing using a combined emg/bioimpedance measurement system. *The Scientific World Journal*, 2014, 2014.

[117] A. R. Scott, B. A. Kudak, S. Skinner, and J. D. Sidman. Use of intraoperative laryngeal electromyography to evaluate stridor in children with arthrogryposis. *Annals of Otology, Rhinology & Laryngology*, 120(3):150–154, 2011.

[118] N. Sengar, M. K. Dutta, and R. Burget. Detection of neuro mascular disease using emg signals in wavelet domain. In *2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON)*, pages 624–627. IEEE, 2017.

[119] T. Shipp and R. E. McGlone. Laryngeal dynamics associated with voice frequency change. *Journal of Speech and Hearing Research*, 14(4):761–768, 1971.

[120] R. Shrivastav. The use of an auditory model in predicting perceptual ratings of breathy voice quality. *Journal of Voice*, 17(4):502–512, 2003.

[121] S. W. Smith. The scientist and engineer's guide to. *Digital Signal Processing,. California Technical Publishing, San, Diego, CA*, 1997.

[122] C. M. Steele, J. W. Bennett, S. Chapman-Jay, R. Cliffe Polacco, S. M. Molfenter, and

M. Oshalla. Electromyography as a biofeedback tool for rehabilitating swallowing muscle function. *Applications of EMG in clinical and sports medicine*, pages 311–328, 2012.

[123] C. E. Stepp. Surface electromyography for speech and swallowing systems: measurement, analysis, and interpretation. *Journal of Speech, Language, Hearing & Research*, 55: 1232–1246, 2012.

[124] C. E. Stepp, J. T. Heaton, R. G. Rolland, and R. E. Hillman. Neck and face surface electromyography for prosthetic voice control after total laryngectomy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 17(2):146–155, 2009.

[125] C. E. Stepp, J. T. Heaton, M. E. Jetté, J. A. Burns, and R. E. Hillman. Neck surface electromyography as a measure of vocal hyperfunction before and after injection laryngoplasty. *Annals of Otology, Rhinology & Laryngology*, 119(9):594–601, 2010.

[126] C. E. Stepp, R. E. Hillman, and J. 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.

[127] C. E. Stepp, J. T. Heaton, M. N. Braden, M. E. Jetté, T. K. Stadelman-Cohen, and R. E. Hillman. Comparison of neck tension palpation rating systems with surface electromyographic and acoustic measures in vocal hyperfunction. *Journal of Voice*, 25(1):67–75, 2011.

[128] C. E. Stepp, J. T. Heaton, T. K. Stadelman-Cohen, M. N. Braden, M. E. Jetté, and R. E. Hillman. Characteristics of phonatory function in singers and nonsingers with vocal fold nodules. *Journal of voice*, 25(6):714–724, 2011.

[129] L. Sulica. Laryngoscopy, stroboscopy and other tools for the evaluation of voice disorders. *Office Procedures in Laryngology, An Issue of Otolaryngologic Clinics-E-Book*, 46(1):21, 2012.

[130] J. P. Teixeira and P. O. Fernandes. Acoustic analysis of vocal dysphonia. *Procedia Computer Science*, 64:466–473, 2015.

[131] M. Testa, T. Geri, L. Gizzi, and D. Falla. High-density emg reveals novel evidence of altered masseter muscle activity during symmetrical and asymmetrical bilateral jaw clenching tasks in people with chronic nonspecific neck pain. *The Clinical journal of pain*, 33(2):148–159, 2017.

[132] I. R. Titze, J. Jiang, and D. G. Drucker. Preliminaries to the body-cover theory of pitch control. *Journal of Voice*, 1(4):314–319, 1988.

[133] T. Tokunaga, S. Baba, M. Tanaka, K. Kashiwagi, K. Kimura, and T. Kawazoe. Two-dimensional configuration of the myoneural junctions of human masticatory muscle detected with matrix electrode. *Journal of oral rehabilitation*, 25(5):329–334, 1998.

[134] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig. Novel speech signal processing algorithms for high-accuracy classification of parkinson's disease. *IEEE transactions on biomedical engineering*, 59(5):1264–1271, 2012.

[135] R. Turley and S. Cohen. Impact of voice and swallowing problems in the elderly. *Otolaryngology - Head and Neck Surgery*, 140(1):33–36, 2009.

[136] M. Vaiman. Standardization of surface electromyography utilized to evaluate patients with dysphagia. *Head & face medicine*, 3(1):1–7, 2007.

[137] M. Vaiman and E. Eviatar. Surface electromyography as a screening method for evaluation of dysphagia and odynophagia. *Head & face medicine*, 5(1):1–11, 2009.

[138] M. Vaiman, E. Eviatar, and S. Segal. Surface electromyographic studies of swallowing in normal subjects: a review of 440 adults. report 1. quantitative data: timing measures. *Otolaryngology—Head and Neck Surgery*, 131(4):548–555, 2004.

[139] E. Van Houtte, S. Claeys, E. D'haeseleer, F. Wuyts, and K. Van Lierde. An examination of surface emg for the assessment of muscle tension dysphonia. *Journal of Voice*, 27(2): 177–186, 2013.

[140] V. Varadarajan, J. H. Blumin, and J. M. Bock. State of the art of laryngeal electromyography. *Current Otorhinolaryngology Reports*, 1(3):171–177, 2013.

[141] B. C. Watson, S. D. Schaefer, F. J. Freeman, J. Dembowski, G. Kondraske, and R. Roark. Laryngeal electromyographic activity in adductor and abductor spasmodic dysphonia. *Journal of Speech, Language, and Hearing Research*, 34(3):473–482, 1991.

[142] C. J. Wentland, P. C. Song, and C. J. Hartnick. Pediatric voice and swallowing disorders related to vocal fold immobility: the use of laryngeal emg. *Current Treatment Options in Pediatrics*, 2(4):339–351, 2016.

[143] G. N. White, F. O'Rourke, B. S. Ong, D. J. Cordato, and D. K. Chan. Dysphagia: causes, assessment, treatment, and management. *Geriatrics*, 63(5), 2008.

[144] D. A. Winter. *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.

[145] P. Xia, J. Hu, and Y. Peng. EMG-Based estimation of limb movement using deep learning with recurrent convolutional neural networks. *Artificial Organs*, 42(5):E67–E77, 2018.

[146] M. Yamada, K. Kumagai, and A. Uchiyama. The distribution and propagation pattern of motor unit action potentials studied by multi-channel surface emg. *Electroencephalography and clinical neurophysiology*, 67(5):395–401, 1987.

[147] J. Yousefi and A. Hamilton-Wright. Characterizing emg data using machine-learning tools. *Computers in biology and medicine*, 51:1–13, 2014.

[148] A. P. Zarzur, A. C. Duprat, G. Shinzato, and C. A. Eckley. Laryngeal electromyography in adults with parkinson's disease and voice complaints. *The Laryngoscope*, 117(5):831–834, 2007.

[149] A. P. Zarzur, A. de Campos Duprat, B. O. Cataldo, D. Ciampi, and E. Fonoff. Laryngeal electromyography as a diagnostic tool for parkinson's disease. *The Laryngoscope*, 124(3): 725–729, 2014.

[150] A. Zhou, S. R. Santacruz, B. C. Johnson, G. Alexandrov, A. Moin, F. L. Burghardt, J. M. Rabaey, J. M. Carmena, and R. Muller. A wireless and artefact-free 128-channel neuromodulation device for closed-loop stimulation and recording in non-human primates. *Nature biomedical engineering*, 3(1):15–26, 2019.

[151] H. Zhu, Y. Yang, J. Rao, L. Liu, Y. Wang, W. Shao, and J. Zhang. Effect of surface electromyographic biofeedback on the pharyngeal phase activities in patients with dysphagia after stroke. *Chinese Journal of Cerebrovascular Diseases*, (11):572–576, 2015.

[152] M. Zhu, W. Yang, O. W. Samuel, Y. Xiang, J. Huang, H. Zou, and G. Li. A preliminary evaluation of myoelectrical energy distribution of the front neck muscles in pharyngeal phase during normal swallowing. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1700–1703. IEEE, 2016.

[153] M. Zhu, F. Liang, O. W. Samuel, S. Chen, W. Yang, L. Lu, H. Zou, P. Li, and G. 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.

[154] M. Zhu, B. Yu, W. Yang, Y. Jiang, L. Lu, Z. Huang, S. Chen, and G. Li. Evaluation of normal swallowing functions by using dynamic high-density surface electromyography maps. *Biomedical engineering online*, 16(1):1–18, 2017.

[155] M. Zhu, L. Lu, Z. Yang, X. Wang, Z. Liu, W. Wei, F. Chen, P. Li, S. Chen, and G. 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.

[156] M. Zhu, O. W. Samuel, Z. Yang, W. Lin, Z. Huang, P. Fang, J. Tan, P. Li, M. C. Tong, K. K. Leung, et al. Using muscle synergy to evaluate the neck muscular activities during normal swallowing. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2454–2457. IEEE, 2018.

[157] M. Zhu, H. Zhang, X. Wang, X. Wang, Z. Yang, C. Wang, O. W. Samuel, S. Chen, and G. Li. Towards optimizing electrode configurations for silent speech recognition based on high-density surface electromyography. *Journal of Neural Engineering*, 18(1):016005, 2021.

[158] J. Zhuang, M. Zhu, X. Wang, D. Wang, Z. Yang, X. Wang, L. Qi, S. Chen, and G. Li. Comparison of contributions between facial and neck muscles for speech recognition using high-density surface electromyography. In *2019 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, pages 1–5. IEEE, 2019.

[159] M. J. Zwarts and D. F. Stegeman. Multichannel surface emg: basic aspects and clinical utility. *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, 28(1):1–17, 2003.