Electromyographic Control of Prosthetic Voice after Total Laryngectomy

by

Cara Elizabeth Stepp

S.B., Engineering Science
Smith College, 2004

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment for the requirements for the degree of

Master of Science in Electrical Engineering and Computer Science
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2008

Signature of Author: ______________________________
Department of Electrical Engineering and Computer Science
August 7, 2008

Certified by: ______________________________
Joseph S. Perkell
Senior Research Scientist of RLE and Department of Brain and Cognitive Sciences
Thesis Supervisor

Certified by: ______________________________
James T. Heaton
Assistant Professor of Surgery, Harvard Medical School
Thesis Co-Supervisor

Accepted by: ______________________________
Terry P. Orlando
Professor of Electrical Engineering and Computer Science
Chair, Department Committee on Graduate Theses

Electromyographic Control of Prosthetic Voice after Total Laryngectomy

by

Cara Elizabeth Stepp

Submitted to the Department of Electrical Engineering and Computer Science
on August 7, 2008 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in
Electrical Engineering and Computer Science

Abstract

The electrolarynx (EL) is a common rehabilitative speech aid for individuals who have undergone laryngectomy, but typical devices lack pitch control and require the exclusive use of one hand. This study investigated the viability of surface electromyography (sEMG) to control the onset and offset of an EMG-controlled EL (EMG-EL) while attending to real-time sEMG biofeedback using sEMG collected from seven locations across the ventral neck and face surface in eight individuals at least 1 year past total laryngectomy.

Speech performance was assessed as the percentage of fully voiced words and successfully produced pauses. During use of the EMG-EL with biofeedback participants increased the sEMG during words and decreased the sEMG during pauses. Electrodes on the superior ventral neck, submental surface, and below the corner of the mouth showed consistently high performance across all participants. These results indicate promise for the applicability of the EMG-EL across a large segment of the laryngectomy population.

Thesis Supervisor: Joseph S. Perkell, Ph.D.
Title: Research Scientist of RLE and BCS

Thesis Co-Supervisor: James T. Heaton, Ph.D.
Title: Assistant Professor of Surgery, Harvard Medical School

Acknowledgements

This work could not be possible without the guidance, foresight, and intellect of Dr. James Heaton. He is the best type of advisor, allowing independence, but always taking the time to provide direction. He is also a very valued friend. Dr. Robert Hillman has also been essential to this work, providing insights and support. Also, he was kind enough to fund my involvement in this work, as well as all other aspects of the project.

Many thanks to Dr. Joseph Perkell at RLE for agreeing to be my EECS supervisor and providing helpful insights. Without him, this master's thesis would not exist.

To my SHBT classmates (shbt2004) and other SHBT friends, thanks for making the times out of lab (and sometimes in lab) so entertaining.

Every single person at the MGH Voice Center has helped this work to exist, and there are far too many to mention. The cooperation of all of the staff there was essential and I continue to be very grateful for it.

My parents and the rest of my family have always offered me their unconditional love and support, so it's hard to imagine being in Boston at all without them. Thank you!

Lastly, this work has been supported by a grant from the NIH (R01-DC006449), awarded to Robert Hillman.

Contents

Chapter1: Introduction

Chapter 2: Background

Chapter 3: Methods

Chapter 4: Results and Discussion

Chapter 5: Conclusions

Bibliography

List of Figures

Chapter 1

Introduction

In 2008, the American Cancer Society estimated that there were 12,250 new cases of laryngeal cancer in the U.S. (*Cancer facts & figures 2008*). A common surgical intervention for the management of invasive laryngeal cancer is total laryngectomy (removal of the larynx), which removes the natural sound source for speech production. The electrolarynx (EL) is a battery-powered unit that provides a mechanical voice source in many of these cases. Despite being widely used, the EL has several limitations. We have previously developed EL technology utilizing neck surface electromyography (sEMG) to control the activation, termination, and pitch of an EL, freeing both hands during speech and providing the ability to produce pitch-based intonational contrasts (Goldstein *et al.*, 2004). We have shown previously that individuals who had their laryngectomy surgery modified to preserve neck strap muscles bilaterally, using targeted muscle reinnervation (TMR) by rerouting of the recurrent laryngeal nerve (Goldstein et al., 2004; Goldstein *et al.*, 2007; Heaton *et al.*, 2004) on one side can gain effective control of EMG-EL initiation, termination, and pitch modulation using sEMG from both RLN- and naturally-innervated strap muscles (Goldstein et al., 2007; Kubert *et al.*, 2007). However, the availability of residual strap muscles or other vocally-active muscle groups after typical laryngectomy surgery is uncertain; thus, the applicability of the EMG-EL to the laryngectomy population at large is still undetermined. The primary goal of this study was to ascertain the onset and offset control capabilities of individuals who had undergone standard total laryngectomy, without special efforts to preserve musculature for EMG-EL control, using sEMG from neck and face locations to control the EMG-EL.

1.1 Thesis Organization

Chapter 2 presents the relevant background on the human voice and alaryngeal communication. This chapter also presents the basics for understanding electromyography and the current standards for collecting and processing electromyographic signals. Finally, this chapter summarizes the previous work developing and training on the EMG-EL.

Chapter 3 describes details regarding the participants of this work, the sEMG and acoustic recording procedures, the EMG-EL system, and the methodology utilized in perceptual judgments of speech performance and sEMG analysis.

Chapter 4 contains results from data analysis and discusses the possible implications of these results.

Chapter 5 provides a general overview of the results and suggests appropriate future work.

1.2 Thesis Contributions

The work described herein investigates the onset and offset control capabilities of individuals who have undergone standard total laryngectomy using sEMG from neck and face locations. We explored this topic by examining the following issues:

1) **Perceptual evaluation of word and pause production during serial speech using the EMG-EL for seven electrode positions on the neck and face.** Acoustic data were scored perceptually as to the percentage of attempted words that were fully voiced (no breaks) and the attempted pauses between words that were successfully produced.

2) **Perceptual evaluation of word and pause production during read sentences using the EMG-EL for "best" electrode positions on the neck and face.** Acoustic data of the sentences produced while the EMG-EL was controlled using the two best electrodes were scored perceptually as the percentage of attempted words that were fully voiced and the attempted pauses between sentences that were produced successfully.

3) **Quantification of sEMG at electrode positions prior to use of the EMG-EL during serial speech tasks.** The sEMG from all seven electrode positions collected during use of a traditional EL for serial speech tasks was analyzed as the percentage above baseline during words and pauses.

4) **Quantification of sEMG at electrode positions during use of the EMG-EL with visual biofeedback during serial speech tasks.** The sEMG from each electrode position as it controlled the EMG-EL for serial speech tasks was analyzed as the percentage above baseline during words and pauses.

Chapter 2

Background

This chapter presents the relevant background for understanding the motives of this project, as well as specifics about the nature of electromyography and current processing recommendations.

2.1 The Human Voice

The classic theory of speech production is the source-filter model (Fant, 1960; further described by Stevens, 2000). This model separates functionally and physically the source, or vocal stimulus, from the filter, which consists of the articulatory apparatus. The source, then, of most human speech is created through use of the larynx. The larynx is a system of suspended cartilages, lined with folds, acting as a valve between the airway and the pharynx. Laryngeal muscles can alter the position and shape of these folds. The larynx transforms airflow from the lungs into a series of air puffs which constitute the voice, which is the source for the articulatory filter of the upper airway. Airflow from the lungs passes through the vocal folds, causing them to open; the folds are then pulled back together due to Bernoulli forces and the elastic properties of the folds, cutting off the airflow and creating an air puff. These forces may be manipulated to change the characteristics of phonation.

2.2 Alaryngeal Communication

In 2008, the American Cancer Society estimated that there were 12,250 new cases of laryngeal cancer in the U.S. (*Cancer facts & figures 2008*). Common surgical intervention for the management of invasive laryngeal cancer is total laryngectomy (removal of the larynx), which removes the natural sound source for speech production. Total laryngectomy includes creating a stoma in the front of the neck that is attached to the patient's trachea allowing him or her to breathe. The oral cavity remains connected to the esophagus, permitting normal swallowing function in many cases. Thus, after laryngectomy, the airway is separated from the pharynx and an extra laryngeal voice source is needed to speak.

Options for voice rehabilitation after total laryngectomy include esophageal speech, tracheo-esophageal (TE) speech, and the use of an electrolarynx (EL). To produce esophageal speech, an individual must learn to inject air into an esophageal reservoir and to release it through the vibratory pharyngoesophageal segment, a skill which is difficult for many individuals to acquire (Gates *et al.*, 1982). TE speech is produced with the use of a TE prosthesis placed through the tracheo-esophageal wall. This allows pulmonary air to be shunted into the esophagus where it can be released through the pharyngoesophageal segment. Although TE speech is a clinically preferred method of voice rehabilitation, only a limited number of patients are rehabilitated in the long-term (Mendenhall *et al.*, 2002) due to reasons such as lack of tissue integrity or poor respiratory health (for reviews see (Gress & Singer, 2005; Monahan, 2005; Pou, 2004)). The EL is a battery-powered unit that provides a mechanical voice source through the tissues of the neck or directly into the mouth via a flexible tube. An EL is used for verbal communication in over half of laryngectomy cases (Gray & Konrad, 1976; Hillman *et al.*, 1998; Mendenhall et al., 2002; Morris *et al.*, 1992). Although it is widely used, the EL has several limitations. Most EL devices require the dedicated use of one hand and do not provide a means to control pitch while speaking, two issues noted in the top five deficits of EL speech communication for both users and for speech-language pathologists (Meltzner *et al.*, 2005). These deficits may contribute to the lowered quality of life scores seen in electrolarynx users relative to individuals using tracheoesophageal (TE) speech, particularly with respect to communication ability (Finizia & Bergman, 2001).

2.3 Electromyography

As a nerve impulse from an alpha motor neuron reaches the motor end plates of muscle fibers comprising a motor unit, the fibers innervated by that neuron discharge nearly synchronously. The electric potential field generated by the depolarization of the outer muscle-fiber membranes essentially reflects the alpha motor neuron activity; the electromyogram (EMG) is a representation of this "myoelectricity" as summed over a number of motor units and measured at some distance. Tissues separating the EMG signal sources (depolarized zones of the muscle fibers) act like spatial low-pass filters on the potential distribution, and constitute a volume conductor. Therefore, the EMG may be measured intramuscularly or at the surface of the skin, yielding different information based on the distance of the observation site from the muscle fibers. For surface detection particularly, the effect of the separating tissues becomes significant.

In order to remove interference sources and to compensate of the low-pass filtering effect of the tissue, surface signals are typically detected using a combination of electrodes, the simplest of which is a differential electrode (Farina *et al.*, 2004b). Bipolar surface EMG (sEMG) is dependent upon on the inter-electrode distance (Roeleveld *et al.*, 1997). In measuring the sEMG, filtering is introduced by finite electrode size, inter-electrode distance, electrode configuration, electrode location, and characteristics of the front-end EMG amplifier.

The European Union sponsored a project termed SENIAM (Surface Electromyography for the Noninvasive Assessment of Muscles), one outcome of which was a set of recommendations for sEMG recording. In general, SENIAM recommends a maximum electrode size of 10 mm in the muscle fiber direction, with an interelectrode distance of approximately 20 mm or ¼ the length of the muscle fiber, whichever is smaller (Hermens *et al.*, 1999). Other recent recommendations include using smaller electrodes (diameter less than 5mm) for sEMG, as the larger electrodes introduce temporal low-pass filtering (Merletti & Hermens, 2004).

Skin preparation techniques can enhance electrode-skin contact, resulting in a reduction of artifacts and less noise. SENIAM recommends shaving the skin surface if it is covered with hair, and cleaning the skin in question with alcohol (Hermens et al., 1999). Also preferred is the practice of slight skin abrasion or "peeling" with adhesive tape; this practice is known to reduce electrode-skin impedance, noise, DC voltages, and motion artifacts (Merletti & Hermens, 2004).

Recommendations for differential electrode placements are that the two electrodes be applied between the innervation zone and a tendon. In the past, sensors have been placed over the belly or over the innervation zone (motor end plate zone), since this was the best location to record "large" monopolar sEMG signals. It is now well known that this location is not suitable for differential recordings; it is not stable or reproducible because relatively small displacements of the sensors with respect to the innervation zone cause large effects on the amplitude of the sEMG signal (Merletti & Hermens, 2004). Thus, in order for sEMG signals to be accurate and repeatable, there must be a clear definition of electrode position relative to the innervation zones (Hermens et al., 1999). When the locations of innervation zones are unknown, use of double differential electrodes can diminish the effects of an ill-placed sensor (Farina *et al.*, 2004a).

Ideal sEMG recording procedures would first identify the innervation zones and find the optimal electrode position on a subject-by-subject basis, using multi-channel electrode arrays. Falla et al. (2002), for example, examined the sternocleidomastoid (SCM) muscles in this way in 11 healthy normal individuals. Based on their findings, they have offered the following recommendations to optimize sEMG recordings from SCM muscles: the electrode should be placed 1/3 of the distance from the sternal notch to the mastoid process, in the direction of the line from the sternal notch to the mastoid process (Falla et al., 2002). Recommendations of this type are not available for sEMG recordings of the extrinsic laryngeal muscles. With regard to ground locations, SENIAM recommends the wrist, the spinous process of C7, or the ankle as appropriate locations (Hermens et al., 1999).

Because of the variability surrounding neck surface electrode contact and participant neck mass, sEMG signals should be normalized to a reference contraction before they are compared between conditions and/or participants (Netto & Burnett, 2006). Most especially, the layers of subcutaneous fat present can have attenuating and widening effects on the signal seen at the surface (Farina & Rainoldi, 1999). Common references include maximal voluntary contraction (MVC) and some percentage of the MVC (usually 50% or 60%). Studies have shown that for more simple, one-joint systems, sub-maximal contractions are more reliable (Allison *et al.*, 1993; Yang & Winter, 1983). However, Netto and Burnett (2006) found that for anterior neck musculature, the MVC reference was more reliable both within-day and between-days. The authors speculate that this is likely due to the complex structure and synergistic action of neck musculature.

The typical sEMG signal has 95% of its power in the frequency range less than 400 Hz. The remaining 5% is mostly due to electrode and equipment noise. For this reason, it is common to lowpass filter the signal with a cutoff point around 500 Hz or 1000 Hz. Movement artifacts create signals in the 0 – 20 Hz range, and can be attenuated by a high pass filter with a cut-off around 10 – 20 Hz. However, there may be some relevant information in this range of the EMG spectrum, specifically the firing rates of the active motor units. Notch filters to remove 50 or 60 Hz interference should not be used due to the high power density of the EMG signal in this range, and the phase rotation introduced to the time waveform (Hermens et al., 1999).

Commonly used amplitude estimators are Average Rectified Value (ARV) and Root Mean Square Value (RMS; see Equation 1). In general, the "best" estimator of SEMG amplitude is RMS (smaller variance for Gaussian distributions, which sEMG approximates). The epoch used for amplitude estimation is recommended by SENIAM to be 0.25 – 0.5 sec for contraction levels above 50% MVC, or 1 – 2 sec for contraction levels below 50% MVC.

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \qquad \text{Equation 1}$$

2.4 EMG-EL Performance

Previousy, we have developed EL technology utilizing neck surface electromyography (sEMG) to control the activation, termination, and pitch of an EL, freeing both hands during speech and providing the ability to produce pitch-based intonational contrasts (Goldstein et al., 2004). The EMG-controlled EL (EMG-EL) uses sEMG to provide on/off control based on a threshold. Rather than using a single threshold for on/off control, the system employs an offset threshold that is an adjustable ratio of the onset threshold, creating a hysteresis band that allows the user to bias the EMG-EL toward maintaining voicing once initiated, reducing unwanted cutouts. Pitch is controlled by the level of suprathreshold sEMG energy. We have shown previously that individuals who had their laryngectomy surgery modified to preserve neck strap muscle bilaterally, using targeted muscle reinnervation (TMR) by rerouting of the recurrent laryngeal nerve (Goldstein et al., 2004; Goldstein et al., 2007; Heaton et al., 2004) on one side can gain effective control of EMG-EL initiation, termination, and pitch modulation using sEMG from both RLN- and naturally-innervated strap muscles (Goldstein et al., 2007; Kubert et al., 2007).

Chapter 3

Methods

This chapter describes details regarding the participants of this work, the sEMG and acoustic recording procedures, the EMG-EL system, and the methodology utilized in perceptual judgments of speech performance and sEMG analysis.

3.1 Participants

Participants were 8 individuals (2 females, 6 males) with a mean age of 61 years (R=48-80 years) who had undergone total laryngectomy at least 1 year previously and were proficient users of EL speech, even if it was not their primary mode of communication. The average time past laryngectomy was 5 years (R=1-17 years). Six participants used EL speech as their primary mode of communication. Two participants were proficient users of TE speech, and had used TE speech as their primary mode of communication for at least 1 year, but maintained EL use for backup communication. Five of the participants had a history of radiation therapy, 3 pre-surgery and 2 post-surgery. All participants reported that they were non-smokers during the time of the experiment, with no history of other speech, hearing, or language disorders.

3.2 Recording Procedure

Differential sEMG electrodes (Delsys DE2.1) consisting of two parallel bars (10 mm by 1 mm) spaced 10 mm apart were positioned at seven locations across the ventral neck and face surface, with preference to the side of the neck with the least anatomical change from surgery, based on participant information at the time of the recording. The differential recording bars of electrodes were aligned perpendicular to the long axis of the body. Example electrode locations are shown in Figure 1 on one participant and schematically. Electrode locations included positions 1 cm lateral to the neck midline (right and left) and just superior to the stoma (#1 and #2), centered on the sternocleidomastoid at one-third of the distance from the clavicle to the mastoid (#3), 1 cm lateral to the ventral neck midline at the superior-most location prior to the start of the submental surface (#4), 1 cm lateral to the submental midline (#5), just below the corner of the mouth (#6), and centered on the lateral jaw superficial to the masseter muscle (#7). All electrodes were referenced to a single ground electrode placed on the participant's wrist. Simultaneous acoustic signals from a headset microphone (AKG Acoustics C 420 PP) and the seven channels of sEMG signals were filtered and recorded digitally (20,000 Hz sampling rate) with Axon Instruments hardware (Cyberamp 380, Digidata 1200) and software (Axoscope). An example of the audio and sEMG data collected during experimentation are shown in Figure 2.

The participants used a commercially-available hand-held EL to produce serial speech (saying the days of the week, counting 1 – 10) with a pause between each word and 10 sentences randomly selected from the Yorkston and Beukelman test (Yorkston & Beukelman, 1981), with

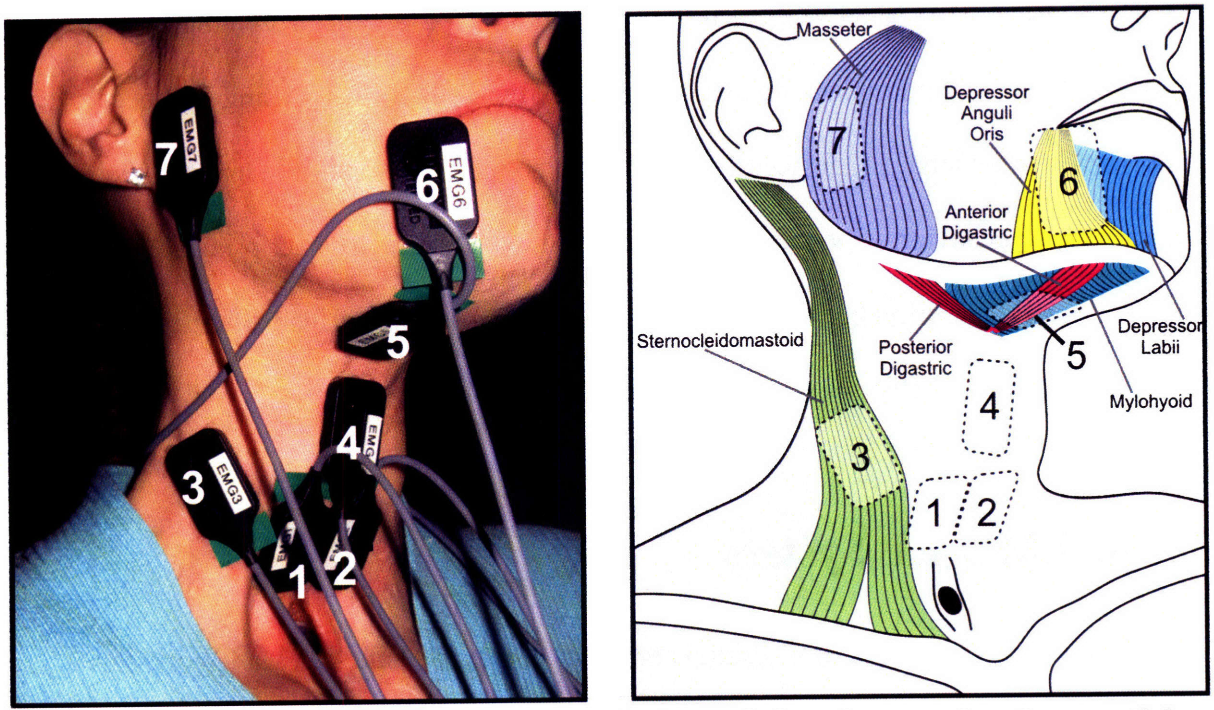

Figure 1. The left panel shows example sEMG electrode locations as placed on participant S3 during experimentation. The right panel shows a schematic of the expected residual muscle after total laryngectomy superficial to sEMG electrodes, depicted on the right side only. Absent from this depiction is the platysma, which is located in the subcutaneous tissue of the neck.

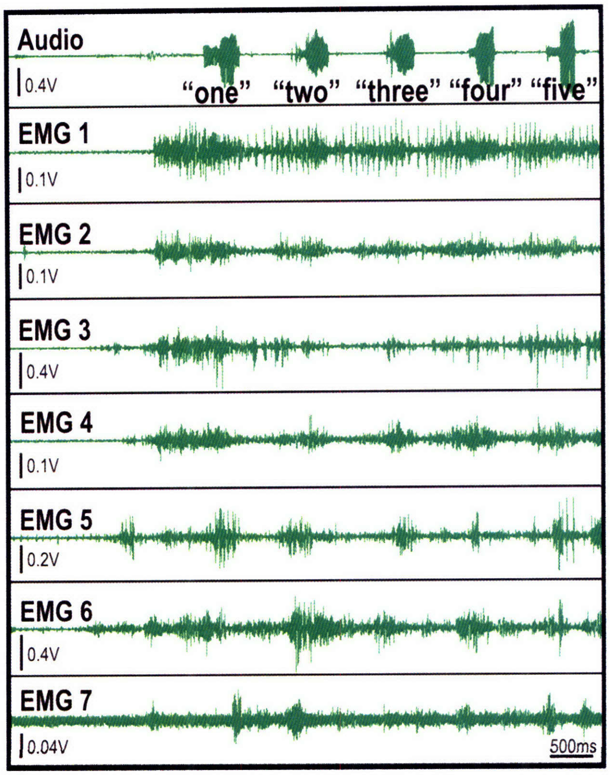

Figure 2. An example of the audio and raw sEMG data collected as a participant counted aloud using a typical EL. Traces indicated by EMG1 – EMG7 refer to sEMG collected from the seven electrode positions indicated in Figure 1.

instructions to leave a pause between each sentence. In the six individuals utilizing EL speech as their primary mode of communication and one of the individuals typically using TE speech, the EL used was their personal EL device. In these cases, no modifications were made to their typical settings or behavior. One individual (S3) who used TE speech as her primary mode of communication did not bring her backup EL to the recording session. This participant used a TruTone™ EL (Griffin Labs) from our clinical facility.

3.3 The EMG-EL System

The EMG-EL consisted of a desktop computer running MATLAB (MathWorks™), a digital signal processing board (DSP56311EVM), and an EL (NuVois). The selected sEMG signal received from the electrode being used to control the device was processed by the DSP to create a fast sEMG envelope (5 Hz low pass filter corner frequency) and a slow sEMG envelope (1 Hz low pass filter corner frequency). The fast envelope was used to control EL activation and termination whereas the slow envelope was used to modulate the EL pitch. Activation and termination thresholds were set independently for each electrode used to control the EMG-EL, with the termination threshold set at 60 – 70% of the activation threshold in order to assist in uninterrupted voicing. The EMG-EL was mountable to participants using a thick, flexible copper wire bent around the base of the neck, but was hand-held in these experiments to avoid interfering with the multiple neck sEMG recording locations. Participants were provided with some basic information in preparation for using the EMG-EL. Participants were informed that stronger muscle contractions would lead to the device turning on, that relaxation would lead to the device turning off, and that increases in muscle activity would lead to increases in pitch, but they were only asked to focus specifically on device onset and offset precision.

Testing of the control capabilities began with sequential control of the EMG-EL with sEMG from each electrode recording location as the participant attempted to produce serial speech with a pause between each word. The participant was presented real-time visual feedback of the RMS sEMG and EMG-EL threshold settings for the electrode position being tested using a video monitor placed approximately 1 m away. An example screenshot of this feedback is shown in Figure 3. After all seven electrode positions had been tested, the participant and the investigators determined the two channels they felt offered the participant the best control. For these two positions, the participant used the EMG-EL to produce 10 sentences selected randomly from the Yorkston and Beukelman test with the instruction to leave a pause between each sentence.

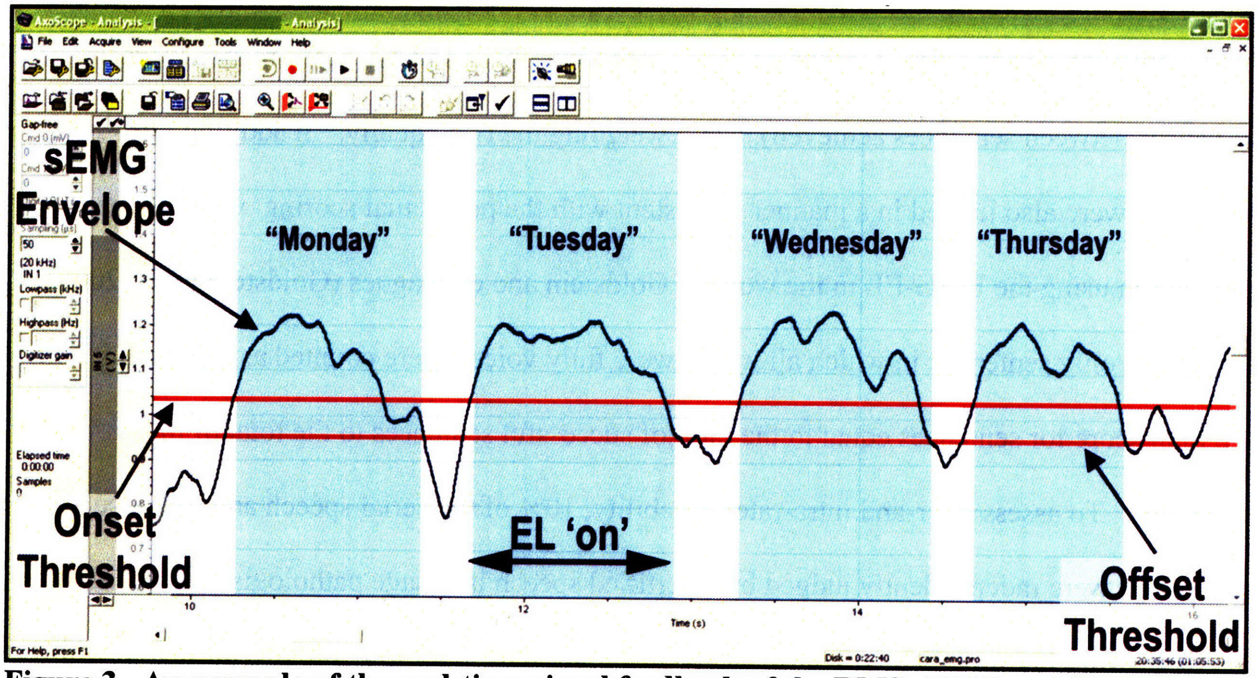

Figure 3. An example of the real-time visual feedback of the RMS sEMG and EMG-EL threshold settings is shown. The black line shows the sEMG envelope used to control the onset and termination of the EMG-EL. The two red lines specify the settings for onset threshold and offset threshold. The blue shading indicates time periods during which the device was activated.

3.4 Data Analysis

Audio recordings were scored perceptually by the first author in terms of the percentage of words that were fully voiced (uninterrupted EMG-EL activation) as well as the percentage of pauses between words or sentences that were achieved. Each participant/electrode combination was scored for a percentage of achieved voicing (percentage of fully voiced words of those attempted) and of the percentage of achieved pauses during serial speech. As a simple indicator of general control ability, “serial speech performance” was defined as the average of the voicing performance (%) and pause performance (%), weighing the two equally. In a few instances, participants were completely unable to produce serial speech using the EMG-EL from particular electrode locations (S1: positions 1, 2, and 3; S3: position 1; S6: position 1), and were scored as having 0% serial speech performance.

The 10 sentences were likewise scored in terms of performance, defined as the average of the voicing performance (percentage of fully voiced words) and pause performance (percentage of pauses between sentences achieved), again weighing the two equally. In addition, the sentences were also judged in a manner consistent with the perceptual scoring of sentence production using the EMG-EL in the work of Goldstein and colleagues (Goldstein et al., 2007). Inasmuch, only sentences in which all words were fully voiced were counted as successful, with the final score for sentences equal to the ratio of successful sentences to the total of those attempted. To assess inter-and intra-rater reliability, 10% of the serial speech and sentence recordings were independently judged by a certified speech language pathologist and again by the first author (approximately 6 months later) yielding inter-rater reliability as measured with Pearson’s R of 0.98 and intra-rater reliability of 0.99.

The sEMG gathered from each recording location during serial speech using a traditional EL was analyzed as the root-mean-square (RMS) during the task (words and pauses) as a percent above the baseline sEMG during the participant's rest. The RMS sEMG was collected for words and for pauses as selected manually by the first author using visual inspection of the audio signal while listening to the audio simultaneously. Visual inspection was used to locate periods in which the EL was activated, whereas listening to the audio enabled identification of articulatory cues such that intended words and pauses could be identified. The sEMG from each electrode during words and pauses was also collected for serial speech produced using the EMG-EL while being controlled by the electrode in question. In this case, device activation was not always concurrent with the intent to speak. Thus, the sEMG as a percent above baseline was only estimated for electrodes at which the participant had achieved at least 80% "serial speech performance" to avoid the addition of error due to listener uncertainty of voicing intent. The threshold of 80% was chosen arbitrarily *a priori*.

Chapter 4

Results and Discussion

This section provides an overview of the results obtained from this work, as well as a discussion of the possible implications of those results.

4.1 EMG-EL Control Performance

The serial speech performance (average of the percentage of appropriately voiced words and the percentage of appropriately unvoiced pauses) is plotted for each participant at each of the seven electrode positions in Figure 4. Serial speech performance at electrode positions #1 and #2 (inferior anterior neck), #3 (sternocleidomastoid), and #7 (masseter) varied greatly amongst participants, with values ranging from 0% to nearly 100%. Alternatively, positions #4 (superior anterior neck), #5 (jaw opening musculature), and #6 (lip depressing musculature) showed consistently high serial speech performance across all participants. These results match the subjective choices of top electrode positions made during the experiment by the participants and investigators. In all participants, at least one of the two electrode positions showing the highest serial speech performance values was one of the two electrode positions chosen during the time of the experiment; in six of the eight participants the two electrodes corresponded exactly.

A history of radiation therapy may be indicative of a loss of muscle integrity. Of the 8 participants, 5 individuals had previously undergone radiation therapy. In order to assess the possible interaction between radiation therapy and the number of viable electrode recording locations, a chi-square test was performed on the serial speech performance data, assessing the counts of electrode locations showing performance values greater or equal to 80% versus performance values less than 80% in the individuals with a history of radiation therapy versus individuals with no history of radiation therapy. Again, the cut-off of 80% was chosen arbitrarily a priori. The results of the chi-square test showed higher than expected counts of "successful" ($\geq$ 80%) electrode locations in the individuals with no history of radiation therapy ($df = 1$, $p = 0.009$).

The performance during sentences was assessed at the two electrode positions chosen for further testing at the time of the experiment, yielding consistently high results across participants for both electrode locations. The average speech performance was 97% (SD=2.3%). When scored as in Goldstein et al. (Goldstein et al., 2007), the sentence scores were far more varied, with a range from 20% and 100% (mean = 64%, SD = 21%).

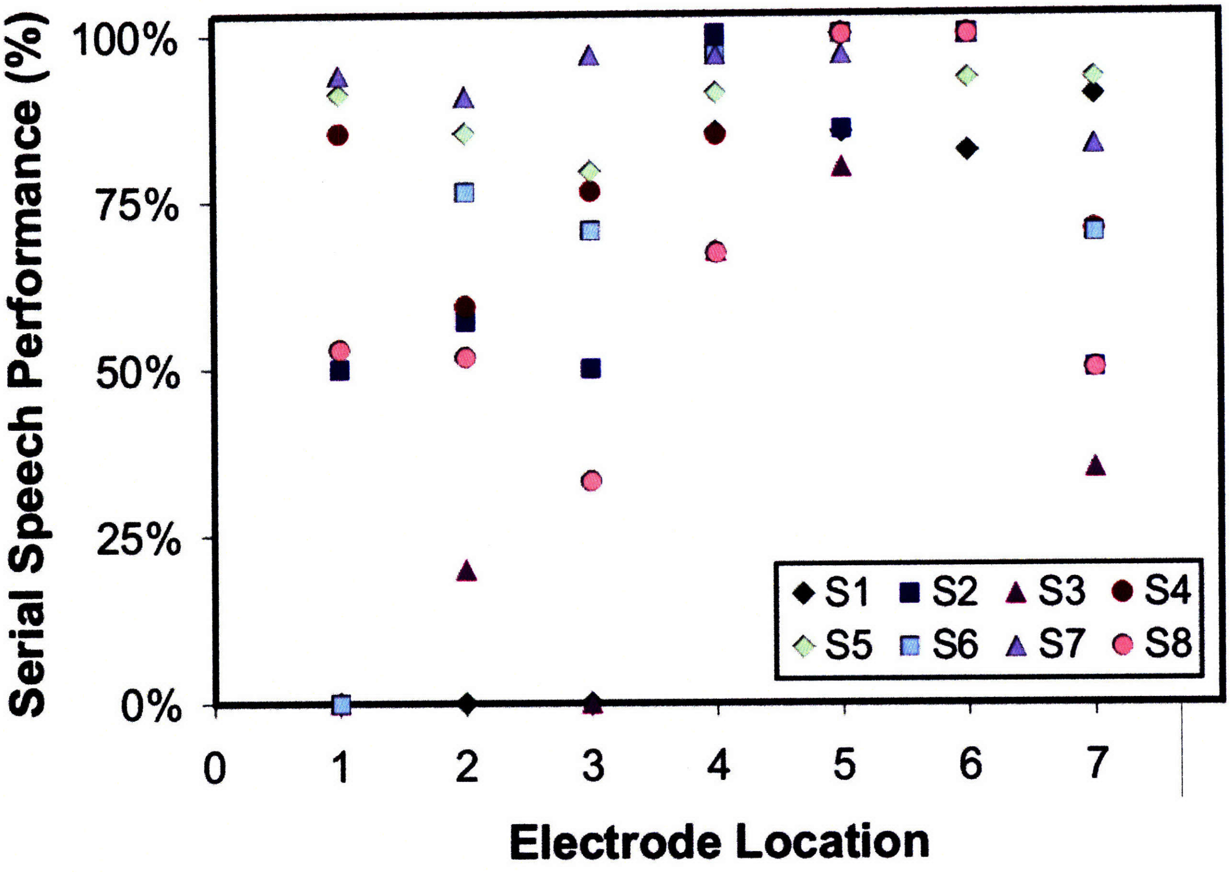

Figure 4. The serial speech performance (average of the percentage of appropriately voiced words and the percentage of appropriately unvoiced pauses) is shown for each participant at each of the seven electrode positions.

4.2 sEMG During Task Performance

During serial speech using a traditional EL as well as during the use of the EMG-EL with visual feedback, the percent above baseline RMS varied significantly over both participant and electrode position. In cases in which the participant was able to achieve at least 80% serial speech performance, sEMG changes seemed to occur for both words (increase) and pauses (reduction). The difference between words and pauses in the percent above baseline RMS sEMG from each recording location during serial speech using a traditional EL (labeled as

“Initial”) and during use of the EMG-EL with visual feedback (labeled as “During Feedback”) is shown in Figure 5. Specifically, in an effort to produce voiced serial speech with pauses between words, participants generally attempted to increase the sEMG during words, decrease the sEMG during pauses, or used a combination of the two approaches. A t-test comparing the percent above baseline RMS during word production initially to that during feedback found a statistically significant increase (mean=108%, $p = 0.03$, one-way paired t-test, $df = 56$). A t-test comparing the percent above baseline RMS during pause production initially to that during feedback found a non-significant decrease of 37% (one-way paired t-test, $p = 0.20$, $df=56$). Moreover, a comparison of the difference between the percent above baseline RMS during words and pauses showed a significant increase during feedback relative to the initial condition (mean = 145%, one-way paired t-test, $p < 0.0001$, $df = 56$).

4.3 Discussion

All participants showed high serial speech performance when using sEMG from electrode recording locations #4, #5, and #6. Vocal-related activity from these recording locations likely stemmed from residual suprahyoid and tongue root musculature (used for articulation and laryngeal control in healthy individuals) and possibly platysma (#4 and #5) as well as the depressor anguli oris (#6). For all participants, the superior ventral neck or submental surface (#4 or #5) was at least one of their two best control locations, leading to average serial speech performance of 95% (SD = 8%). The face recording location below the corner of the mouth (superficial to the depressor anguli oris; #6) was also an effective control location for all participants, facilitating average serial speech performance of 97% (SD = 6%). However, this location presents an unfavorably conspicuous electrode site and may be more prone to false

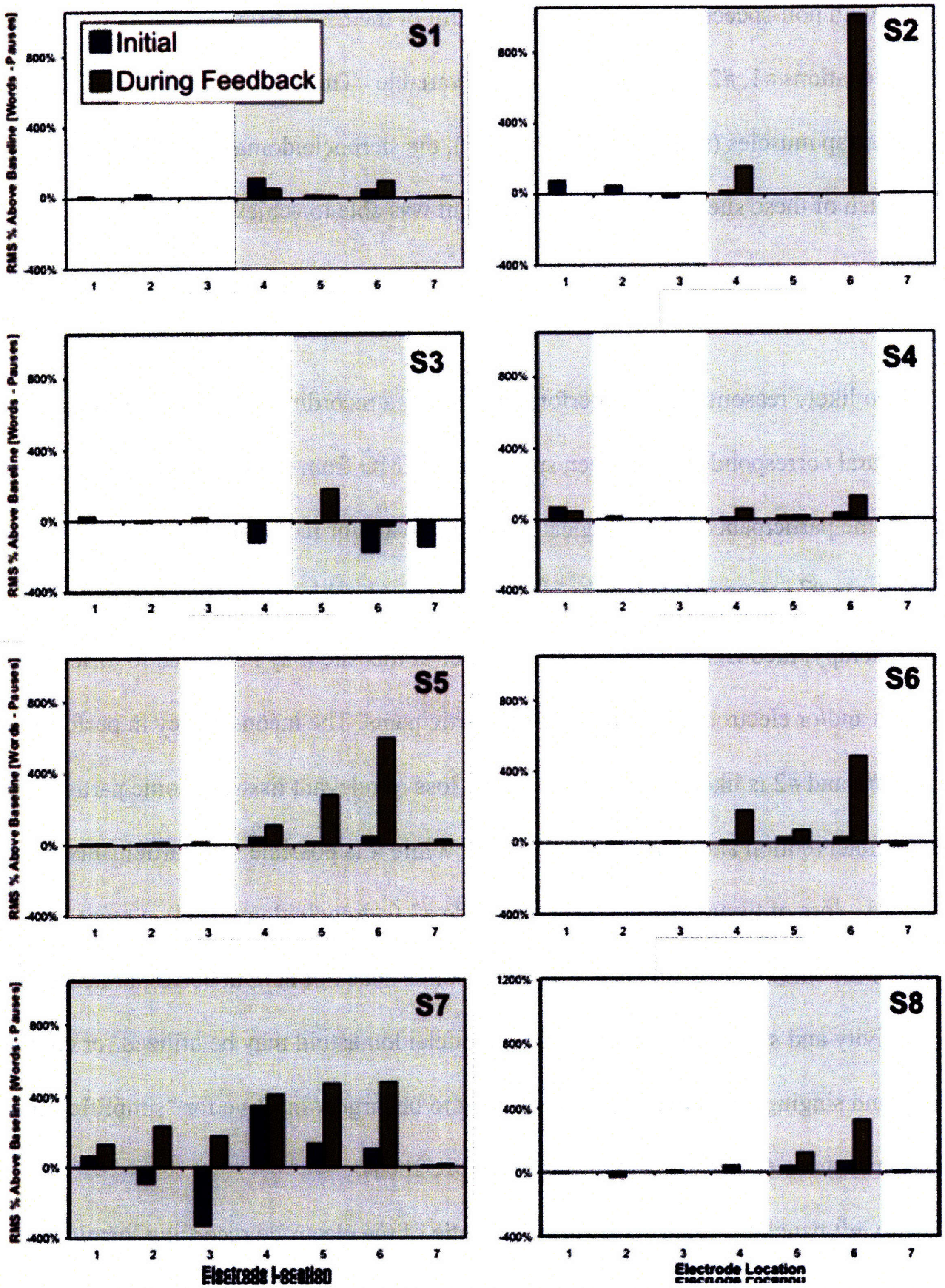

Figure 5. The difference between words and pauses in the percent above baseline RMS sEMG from each recording location during serial speech using a traditional EL (labelled as "Initial") and during use of the EMG-EL with visual feedback (labelled as "During Feedback") for each participant. Electrode positions for which each participant achieved at least 80% serial speech performance using the EMG-EL are indicated with background shading in grey.

triggering with non-speech lip movements. Control of the EMG-EL with sEMG from electrode recording locations #1, #2, #3, and #7 was more variable. These electrodes are placed to record from neck strap muscles (sternohyoid; #1 and #2), the sternocleidomastoid (#3), and the masseter (#7). At each of these sites, at least one participant was able to achieve serial speech performance at or above 80%, while others were unable to utilize the site effectively (see Figure 4).

Two likely reasons for poor performance using a recording site would be the result of a lack of natural correspondence between speech and sEMG from the site and/or loss of relevant tissue in some participants due to surgical intervention and/or radiation therapy. The tissue integrity at site #7 seems unlikely to be affected by most total laryngectomy surgeries or radiation therapy; inconsistency in the performance at this site may be related to differences in articulation and/or electrode placement across participants. The inconsistency in performance in electrodes #1 and #2 is likely to be a result of the loss of relevant tissue in some participants due to surgical intervention and/or radiation therapy. While it is possible that participants experienced a loss of tissue integrity near electrode #3 (sternocleidomastoid), a more likely explanation for variable performance at this location is a lack of natural correspondence between muscle activity and speech production. The sternocleidomastoid may be utilized for complex breathing and singing performance, but is thought to be largely inactive for "simplified speaking tasks" such as those attempted here (Pettersen *et al.*, 2005).

The left panel of Figure 1 shows a schematic of the electrode recording locations relative to the musculature thought to be left after total laryngectomy. Absent from this depiction is the platysma, which is a superficially located thin sheet of muscle in the subcutaneous tissue of the neck. It extends over the anterolateral aspect of the neck from the inferior border of the mandible to the superior aspect of the pectoralis major. During total laryngectomy, the neck is incised near the clavicle and an apron-shaped skin flap is raised toward the head, preserving most of the platysma length attached to the lower jaw, and maintaining its motor supply through the cervical branch of the facial nerve (Wong, 1996). The likely retraction of deeper neck muscles (e.g. strap muscles) after their division during laryngectomy and the probable survival and superficial location of the platysma makes it a potential source for the sEMG collected at electrode recording sites #1, #2, #3, #4, and possibly #5 in this study. Although the activation of the platysma during speech has been studied less than other laryngeal and orofacial musculature, it has been shown to be active during speech production (McClean & Sapir, 1980). It is thought to be an antagonist to the orbicularis oris inferior muscle, and has been shown to activate during lowering of the lower lip during speech and non-speech tasks (McClean & Sapir, 1980). It is possibly active during a large selection of phonemes created by lip movements. Therefore, platysma-based sEMG could provide consistent EMG-EL control during running speech, but would perhaps perform weakly for control of the EMG-EL for non-articulated speech such as prolonged vowels produced without lip rounding.

Comparison of overall serial speech performance between individuals with a history of radiation therapy and those without showed that those individuals who had had radiation tended to have fewer electrode recording positions yielding at least 80% serial speech performance (chi-square test, df = 1, p = .009). Perhaps the radiation therapy reduced the muscle integrity, leading to reduced overall control capabilities; however, it is also possible that the need for radiation therapy covaries with the need for more extensive surgical intervention. Therefore, we cannot determine that there is a causative relationship between a history of radiation therapy and the ability of an individual to control the EMG-EL.

Using their two best electrode recording locations for EMG-EL control, all participants were able to produce running speech (sentences) with few disrupted words due to breaks in voicing. This replicates the findings of Goldstein and colleagues, who found that individuals controlling the EMG-EL with neck strap muscles showed improvement in their ability to produce sentences without formal training, and that running speech may be more easily produced with the EMG-EL due to the ability of participants to anticipate pauses and adjust muscle activity accordingly (Goldstein et al., 2007). When sentence production was scored similarly to Goldstein et al. (Goldstein et al., 2007), the average score for all participants at both electrode sites tested was 64% (SD = 21%). The 3 individuals with laryngectomy studied by Goldstein et al. had an average score of 30% (SD = 26%) prior to training and 83% (SD = 21%) after training (Goldstein et al., 2007) using neck strap muscle surgically modified to be innervated by the RLN. Their training consisted of 4 – 10, 10 – 60 min training sessions. Without any formal training or surgical modifications, the individuals with laryngectomy studied here were able to produce sentences at least as well as the individuals in Goldstein et al. before training.

The fact that the difference between the percent above baseline RMS sEMG during words and pauses showed a highly significant increase during feedback relative to the initial condition suggests that while participants may have employed differing strategies (increasing sEMG during words or decreasing sEMG during pauses), the chief result was an increase in the dynamic range between the sEMG during words versus pauses. Further, these statistically significant changes suggest that attendance to relevant muscle groups using visual sEMG feedback can improve EMG-EL control, without the use of formal training protocols.

Chapter 5

Conclusions

This chapter summarizes the results of this thesis and explores the future work recommended based on those results.

5.1 Summary

All participants were able to produce running and serial speech with the EMG-EL controlled by sEMG from multiple recording locations, with the superior ventral neck or submental surface locations providing at least one of the two best control locations. Vocal-related activity from these recording locations likely stemmed from residual suprahyoid and tongue root musculature, and possibly the platysma. The face recording location below the corner of the mouth (superficial to the depressor anguli oris) was also an effective control location for all participants, but presents an unfavorably conspicuous electrode site and may be more prone to false triggering with lip movements. Each participant had multiple sEMG recording locations providing intuitive and effective prosthetic voice control perceived as natural as a typical handheld EL without formal training, indicating promise for use of an EMG-EL system across a large segment of the laryngectomy population.

5.2 Future Work

Without any formal training or surgical modifications, the individuals with laryngectomy studied here were able to produce sentences at least as well as the individuals in Goldstein et al. before training. Future work will include a study on the effects of training on our present participants' ability to control the EMG-EL. Mimicking the training protocol and outcome measures employed in studies of individuals controlling the EMG-EL with RLN-innervated neck strap muscle will reveal whether EMG-EL control capabilities are enhanced by modification of the laryngectomy surgery.

Bibliography

Allison, G. T., Marshall, R. N., & Singer, K. P. (1993). Emg signal amplitude normalization technique in stretch-shortening cycle movements. *J Electromyogr Kinesiol, 3*(4), 236-244.

Cancer facts & figures 2008. (2008).). Atlanta: American Cancer Society.

Falla, D., Dall'Alba, P., Rainoldi, A., Merletti, R., & Jull, G. (2002). Location of innervation zones of sternocleidomastoid and scalene muscles--a basis for clinical and research electromyography applications. *Clin Neurophysiol, 113*(1), 57-63.

Fant, G. (1960). *Acoustic theory of speech production.* The Hague: Mouton.

Farina, D., Merletti, R., & Disselhorst-Klug, C. (2004a). Multi-channel techniques for information extraction from the surface emg. In R. Merletti & P. Parker (Eds.), *Electromyography: Physiology, engineering, and noninvasive applications* (pp. 169-203). Hoboken, NJ: John Wiley & Sons, Inc.

Farina, D., Merletti, R., & Stegeman, D. F. (2004b). Biophysics of the generation of emg signals. In R. Merletti & P. A. Parker (Eds.), *Electromyography: Physiology, engineering, and noninvasive applications* (pp. 81-105). Hoboken, NJ: John Wiley & Sons, Inc.

Farina, D., & Rainoldi, A. (1999). Compensation of the effect of sub-cutaneous tissue layers on surface emg: A simulation study. *Med Eng Phys, 21*(6-7), 487-497.

Finizia, C., & Bergman, B. (2001). Health-related quality of life in patients with laryngeal cancer: A post-treatment comparison of different modes of communication. *Laryngoscope, 111*(5), 918-923.

Gates, G. A., Ryan, W., Cooper, J. C., Jr., Lawlis, G. F., Cantu, E., Hayashi, T., et al. (1982). Current status of laryngectomee rehabilitation: I. Results of therapy. *Am J Otolaryngol, 3*(1), 1-7.

Goldstein, E. A., Heaton, J. T., Kobler, J. B., Stanley, G. B., & Hillman, R. E. (2004). Design and implementation of a hands-free electrolarynx device controlled by neck strap muscle electromyographic activity. *IEEE Transactions on Biomedical Engineering, 51*(2), 325-332.

Goldstein, E. A., Heaton, J. T., Stepp, C. E., & Hillman, R. E. (2007). Training effects on speech production using a hands-free electromyographically controlled electrolarynx. *Journal of Speech, Language, and Hearing Research, 50*(2), 335-351.

Gray, S., & Konrad, H. R. (1976). Laryngectomy: Postsurgical rehabilitation of communication. *Arch Phys Med Rehabil, 57*(3), 140-142.

Gress, C., & Singer, M. (2005). Tracheoesophageal voice restoration. In P. Doyle & R. L. Keith (Eds.), *Contemporary consideration in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing* (pp. 431-452). Austin: Pro-Ed.

Heaton, J. T., Goldstein, E. A., Kobler, J. B., Zeitels, S. M., Randolph, G. W., Walsh, M. J., et al. (2004). Surface electromyographic activity in total laryngectomy patients following laryngeal nerve transfer to neck strap muscles. *Annals of Otology, Rhinology, & Laryngology, 113*(9), 754-764.

Hermens, H. J., Freriks, B., Merletti, R., Stegeman, D. F., Blok, J., Gunter, R., et al. (1999). *European recommendations for surface electromyography: Results of the seniam project*. Enschede: Roessingh Research and Development b.v.

Hillman, R. E., Walsh, M. J., Wolf, G. T., Fisher, S. G., & Hong, W. K. (1998). Functional outcomes following treatment for advanced laryngeal cancer. Part i--voice preservation in advanced laryngeal cancer. Part ii--laryngectomy rehabilitation: The state of the art in the va system. Research speech-language pathologists. Department of veterans affairs laryngeal cancer study group. *Annals of Otology, Rhinology, & Laryngology Supplement, 172*, 1-27.

Kubert, H., Heaton, J. T., Stepp, C. E., Zeitels, S. M., Walsh, M., Prakash, S. R., et al. (2007). *Electromyographic control of a hands-free electrolarynx using neck strap muscles.* Paper presented at the American Speech-Language-Hearing Association Convention, Boston, MA.

McClean, M. D., & Sapir, S. (1980). Surface electrode recording of platysma single motor units during speech. *Journal of Phonetics, 8*, 169-173.

Meltzner, G. S., Hillman, R. E., Heaton, J. T., Houston, K. M., Kobler, J. B., & Qi, Y. (2005). Electrolaryngeal speech: The state of the art and future directions for development. In P. C. Doyle & R. L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing* (pp. 571-590). Austin, TX: PRO-ED.

Mendenhall, W. M., Morris, C. G., Stringer, S. P., Amdur, R. J., Hinerman, R. W., Villaret, D. B., et al. (2002). Voice rehabilitation after total laryngectomy and postoperative radiation therapy. *J Clin Oncol, 20*(10), 2500-2505.

Merletti, R., & Hermens, H. J. (2004). Detection and conditioning of the surface emg signal. In R. Merletti & P. Parker (Eds.), *Electromyography: Physiology, engineering and non-invasive applications* (pp. 107-131). Hoboken: Wiley-IEEE Press.

Monahan, G. (2005). Clinical troubleshooting with tracheoesophageal puncture voice prostheses. In P. Doyle & R. L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing* (pp. 481-502). Austin: Pro-Ed.

Morris, H. L., Smith, A. E., Van Demark, D. R., & Maves, M. D. (1992). Communication status following laryngectomy: The iowa experience 1984-1987. *Ann Otol Rhinol Laryngol, 101*(6), 503-510.

Netto, K. J., & Burnett, A. F. (2006). Reliability of normalisation methods for emg analysis of neck muscles. *Work, 26*(2), 123-130.

Pettersen, V., Bjorkoy, K., Torp, H., & Westgaard, R. H. (2005). Neck and shoulder muscle activity and thorax movement in singing and speaking tasks with variation in vocal loudness and pitch. *J Voice, 19*(4), 623-634.

Pou, A. M. (2004). Tracheoesophageal voice restoration with total laryngectomy. *Otolaryngol Clin North Am, 37*(3), 531-545.

Roeleveld, K., Stegeman, D. F., Vingerhoets, H. M., & Van Oosterom, A. (1997). Motor unit potential contribution to surface electromyography. *Acta Physiol Scand, 160*(2), 175-183.

Stevens, K. N. (2000). *Acoustic phonetics*. Cambridge, MA: MIT Press.

Wong, F. (1996). Total laryngectomy. In B. Bailey, K. Calhoun, A. Coffey & J. G. Neely (Eds.), *Atlas of head & neck surgery--otolaryngology* (pp. 934). Philadelphia: Lippincott-Raven.

Yang, J. F., & Winter, D. A. (1983). Electromyography reliability in maximal and submaximal isometric contractions. *Arch Phys Med Rehabil, 64*(9), 417-420.

Yorkston, K. M., & Beukelman, D. R. (1981). Communication efficiency of dysarthric speakers as measured by sentence intelligibility and speaking rate. *J Speech Hear Disord, 46*(3), 296-301.