33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

# Development of a Wireless Electromyographically Controlled Electrolarynx Voice Prosthesis

James T. Heaton, Mark Robertson, and Cliff Griffin

*Abstract*— **The most common artificial voice source for post-laryngectomy speech rehabilitation is the hand-held buzzer or electrolarynx (EL). EL speech is often described as mechanical-sounding (robotic), and typically lacks pitch variation, making it monotone and unnatural. Prior studies have shown improved perceptual ratings of speech naturalness when pitch variation is added to EL speech, and a proof-of-concept EL prosthesis has been developed to provide pitch variation and voice on/off control in relation to neck muscle electromyographic (EMG) signals. The goal of the present study was to design a new wireless version of the EMG-controlled EL (EMG-EL) that could provide a flexible mixture of manual (push button) and automatic (EMG-based) control options for voice onset/offset and pitch, and that could be manufactured at a reasonable cost for widespread patient use. This paper describes both technical and human factors considered while designing the new EMG-EL voice prosthesis.**

## I. INTRODUCTION

Individuals who have lost their voice box (larynx) due to disease or trauma require an artificial voice source in order to speak. Although there are several options for producing a prosthetic voice, the most common artificial voice source is the hand-held buzzer or electrolarynx (EL; see Figure 1) [1]. This device generates sound by electromechanically vibrating a plastic diaphragm when activated by a push-button switch, using technology that has remained largely unchanged since becoming commercially available in the mid 1940s [2]. It is relatively easy to use and enables most laryngectomees to verbally communicate, but with reduced intelligibility and markedly degraded naturalness [for review see 1]. In addition, EL speech is often described as robotic and monotone, attributable largely to the lack of normal fundamental frequency (F0) variation/control, abnormal voice source timing (due to voice onset and offset being controlled by a push button) and abnormal spectral characteristics (e.g., reduced energy below 500 Hz). Previous research has demonstrated a reduction in speech naturalness [3] and intelligibility [4] when F0 variation is stripped from a normal laryngeal voice source, as well as improved naturalness [3] and intelligibility [5] of EL-assisted speech when F0 modulation is added to the speech signal. Current EL devices typically lack dynamic F0 control, or only provide it through manual variation of push-button pressure (TruTone™, Griffin Laboratories), which can be difficult for many users to master. Moreover, current EL devices require manual activation of voice onset/offset, which often introduces voice timing errors. In combination, these deficiencies in EL voice control substantially limit the users' communication capabilities – particularly when speaking with someone who is unfamiliar with EL speech.

Manuscript received March 26th, 2011. This work was supported by 1R41DC011212-01 to Mark Robertson.

J. T. Heaton is with the Massachusetts General Hospital Department of Surgery and Harvard Medical School, One Bowdoin Square Floor 11, Boston, MA 02114 USA. (phone: 617-726-0211; fax: 617-643-0681; email: james.heaton@mgh.harvard.edu).

Mark Robertson and Cliff Griffin are with Griffin Laboratories, Inc., 27636 Ynez Rd., #L7-199, Temecula, Ca, 92591, USA. (email: mr@griffinlab.com, cg@griffinlab.com).

Figure 1. Examples of electrolarynx (EL) devices. From left to right: the Western Electric, Neovox, Servox, and TruTone ELs. Note the difference in size between the two older models on the left (1950s) versus the two models on the right sold today. Quarter provides size reference.

Our previous research has focused on improving EL communication by providing a neural interface for EL control [6-9], demonstrating that neck surface electromyographic (EMG) signals can serve as an intuitive and effective control source for EL voice onset/offset timing and F0 modulation (i.e. intoning interrogative versus declarative statements) [9-12]. Our initial platform for providing EMG control of an EL (EMG-EL) has provided an encouraging proof-of-concept, but has required several refinements in preparation before being manufactured for patient use. Specifically, the EMG-EL hardware has been either non-portable (e.g. attached to a desktop computer system) or housed in an enclosure that was relatively large, heavy, and included several controls that were not appropriate for patient adjustment. Moreover, up to this point our EMG sensor and ground electrode has been wired to the signal processing hardware, which was likewise wired to the EL transducer, making the system both conspicuous and constraining. Finally, although our prototype system has provided effective EMG-based voice control, it has not simultaneously offered a conventional (push-button) control interface, which would be an important option to maintain.

The goal of this research was to build upon our previous work by developing an effective, practical, and user-friendly EMG-EL that can provide improved control over voice source timing and F0 for more intelligible and natural-sounding EL speech. This paper describes the technical and human factors that were considered in developing the design and control options of the new wireless EMG-EL voice prosthesis.

## II. EMG-EL System Design

### A. *Original EMG-EL System Designs and Performance*

Our initial design and implementation of an EMG-EL [9] used a commercial differential skin surface EMG sensor (Delsys Inc., Boston MA) attached to custom signal processing hardware (Draper Laboratories, Boston, MA), which controlled the function of an EL transducer (NuVois, Mountain Precision Mfg., ID). The EL transducer was held to the neck surface using a flexible brace for hands-free operation, and the processor could be worn on a belt or fanny pack. The EMG sensor could be placed on any number of body surface locations for EMG-based EL control.

EMG activity was band-pass filtered (10-500 Hz), amplified, rectified, and low-pass filtered along two parallel pathways for the creation of envelopes that tracked EMG activity with “slow” (1 Hz corner frequency) and “fast” (1-9 Hz corner frequency) time constraints. These envelopes were used to control the fundamental frequency of the EL transducer and the EL onset/offset, respectively. Vocal amplitude was adjustable via a knob on the EMG processing unit, but was not dynamically controlled by EMG activity because it is already modulated by change in mouth gape during EL speech [13], and EL amplitude is often insufficient, so we did not want to exacerbate amplitude deficiencies by attempting to place it under dynamic control.

We have studied the ability of individuals who had undergone complete surgical removal of the larynx (Laryngectomees) to speak using the prototype EMG-EL prosthesis using multiple prospective EMG control sources. These have included both naturally and RLN-innervated neck strap muscles superficial to the neck ventral surface [10, 11], as well as submental muscles (under the chin) and on the lower face [12]. Of these potential EMG-EL control sources, the submental muscles provide several distinct advantages, including: 1) their availability after most laryngectomy surgeries without needing special surgical steps, 2) sensor position at this location is less conspicuous than on the face and is less prone to false triggering from non-speech behavior (e.g. smiling), and 3) our

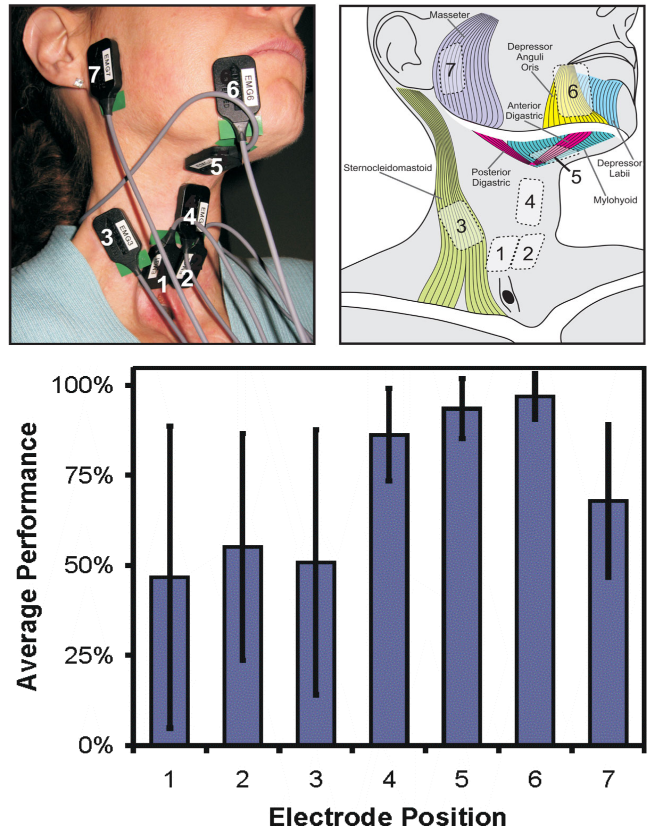


Figure 2. *(Top)* Seven sensor locations are shown that were tested for EMG-EL control, along with expected residual muscles after total laryngectomy superficial to the EMG electrodes (except for the platysma and pharynx). *(Bottom)* The average performance (N=8) for each EMG location on serial speech tasks, indicating that locations 4-6 supported the best EMG-EL control. Modified from IEEE Trans Neural Syst Rehabil Eng. 2009 Apr:17(2):146-155.

demonstration of good control performance with this location in a test of 8 Laryngectomee participants, without the need for training (see sensor location 5 in figure 2). Therefore, we will use the submental surface as the preferred control location when testing future iterations of EMG-EL prosthetic hardware.

The original EMG-EL analog control circuit produced an internal threshold-dependent hysteresis band whereby the threshold for voice offset was some degree lower than for onset (based on the initial setting of the onset threshold) to avoid rapid oscillations in the device output and unintentional voice cut-outs. This dual onset/offset threshold was implemented as a controllable parameter in the subsequent DSP version of the EMG-EL [11], and proved important for maintaining intended EL voicing, because vocal-related EMG activity tends to peak at the start of an utterance, drops to an intermediate level for the remainder of intended voicing, and falls again upon intended termination. Thus, dual onset/offset thresholds will be an important feature to maintain in future EMG-EL versions.

### B. *New EMG-EL System Design*

The new EMG-EL system obtains EMG signals from a wireless sensor (described below) which communicates with an EL that is either hand-held or neck-mounted. The EMG processing steps for EL control are described in Figure 3. The EL portion of the system is based on an upcoming version of the TruTone™ soon to be offered by Griffin Laboratories (Temecula, Ca). This new EL design replaces the analog controls of the previous version with digital controls via a microcontroller, and is modified in the following ways to provide EMG-based control: 1) the addition of an RF transceiver, 2) special firmware to accept the EMG data and utilize it for F0 and voice onset-offset control, and 3) slight enlargement of the enclosure to hold the additional RF circuitry and adjustment points for EMG-EL parameters.

One important aspect of this system is the maintenance of manual (push-button) control options for the EL so that it can be used in a conventional manner or via a flexible combination of push-button and EMG-based controls. For example, at one extreme the user can turn off all EMG-based control and use the device as a typical TruTone™, or the EL can be mounted on the neck and controlled completely hands-free via EMG activity. In between these extremes are several options whereby the manual button can be used to supplement or gate the EMG-based control. Table 1 summarizes the control combinations that the EMG-EL system offers, which will accommodate a wide range of user needs and capabilities.

Table 1. EMG-EL Control Options for Combining Manual and EMG-Based Interfaces

| **Control Mode** | **Voice Onset** | **Voice Offset** | **F0 Modulation** |
|---|---|---|---|
| Completely Manual | Button Press | Button Release | Button Pressure Modulation |
| Manual on-off with EMG-Based F0 | Button Press | Button Release | EMG Envelope |
| Completely EMG-Controlled | Supra-Threshold EMG | Sub-Threshold EMG | EMG Envelope |
| Dual Activation | Button Press **or** Supra-Threshold EMG | Button Release **or** Sub-Threshold EMG | Button Pressure **or** EMG Envelope |
| Manual Gaiting of EMG-Controlled Onset | Button Press **and** Supra-Threshold EMG | Button Release **or** Sub-Threshold EMG | Button Pressure **or** EMG Envelope |

### C. *Wireless EMG Sensor Design*

The EMG sensor is 28 x 23 x 10 mm and weighs 7.6 g. It has parallel electrode bars of stainless-steel, 1 mm wide by 10 mm long, spaced 10 mm apart for differential acquisition of skin-surface EMG signals (see Figure 4) in reference to half of the battery voltage. The rechargeable battery

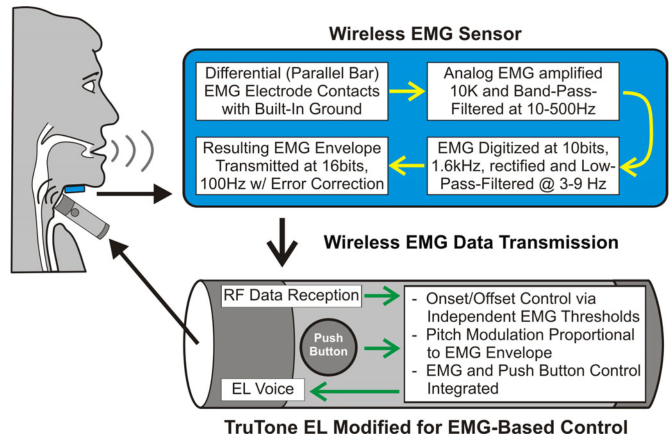


Figure 3. EMG-EL system diagram with the sensor positioned on the submental surface (under the chin), wireless communication between the sensor and EL, and the EL transmitting sound through the neck wall into the vocal tract.

powering the sensor is 3.7 V, LiPoly, with 20 mAh. The sensor filters and amplifies the surface EMG signal, generates an EMG envelope using a simple diode-model detector with a 20 ms decay rate, and digitizes the envelope at 1.6 kHz (10 bits). The envelope is then transmitted to the EL using an RF transmitter (nRF2402, Nordic Semiconductor ASA, Tiller, Norway) in the 2.4-2.5 GHz band at 16 bits (to reduce rounding errors and increase dynamic range). This RF transmitter is set for a low transmission distance (approximately 15 cm) and low sampling rate (100 Hz), drawing < 1 mA and thereby enabling approximately 18 hours of battery life for the sensor between charges. The short transmission distance also helps prevent RF cross-talk among multiple EMG-EL users or other RF receivers.

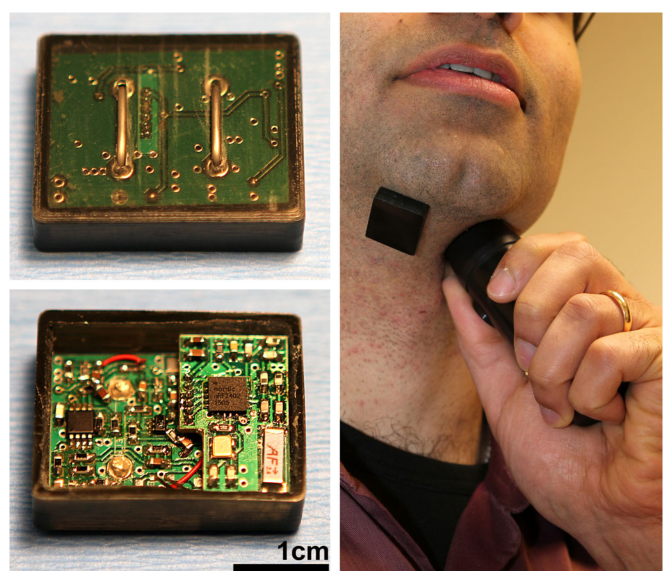


Figure 4. Wireless EMG sensor. Top Left: Bar electrodes protruding through the semi-transparent sensor enclosure. Bottom Left: Circuit boards in opened enclosure with the rechargeable battery removed. Right: Sensor placed on the submental surface with an EL positioned for speaking.

## III. RESULTS OF DEVICE TESTING

In preliminary recordings of surface EMG signals in an anatomically intact adult male, placement of the sensor on the neck surface above the thyroid cartilage of the larynx or under the chin produced envelopes that closely corresponded to voice onset/offset (see Figure 5), and were on average >360% of baseline for individual words, and >380% of baseline for running speech for both sensor locations. Vocal-related signals obtained with the sensor placed on the face superficial to peri-oral musculature (such as the depressor anguli oris and orbicularis oris) were substantially stronger (twice or more) than the neck or submental locations, but face placement was also prone to frequent non-speech-related signals associated with facial expression.

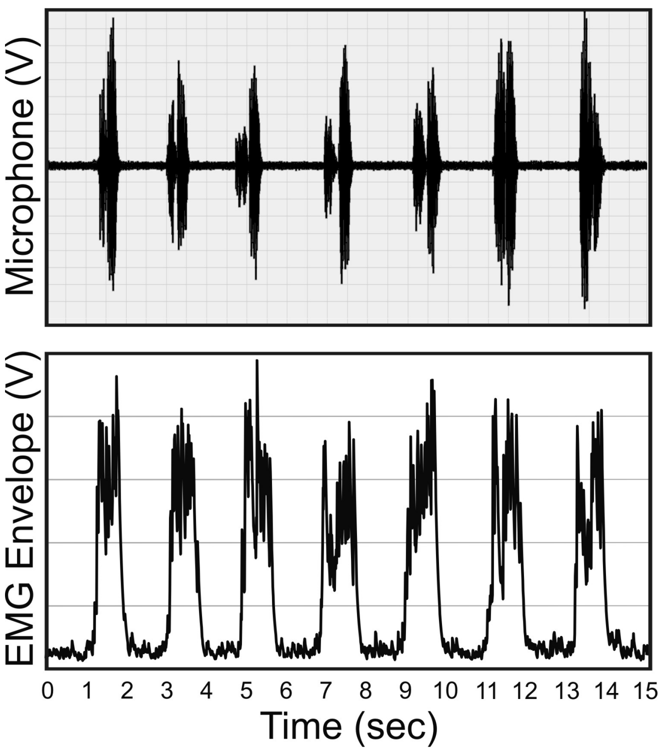


Figure 5. Simultaneous microphone (top plot) and submental surface EMG (bottom plot) recordings of an adult male saying the days of the week (Sunday through Saturday) with his normal voice. The EMG envelope generated from the wireless sensor (bottom plot) corresponded closely to each utterance, with relatively low baseline between words, making it an appropriate signal for EL onset/offset control.

## IV. CONCLUSIONS AND FUTURE WORK

The EMG-EL system is able to detect natural vocal-related EMG activity from the face, neck or submental surface and transmit it as an envelope to a modified commercial EL for control of prosthetic voice onset, offset, and pitch modulation. These signals correlating with voice/speech were typically > 360% of resting baseline in preliminary tests, which should be strong enough for reliable vocal control, because prior research has found that signals 200% of baseline (or greater) are adequate for effective EMG-EL control [9].

Upcoming experiments with laryngectomee participants will compare EL voice timing control, serial speech capability, speech naturalness, and intonational control using several different EMG-EL control modes ranging from completely manual to completely automatic. Naive listeners will then judge the naturalness and intelligibility of recorded EMG-EL speech samples produced under these different control modes. We will also obtain feedback from EMG-EL users with formal questionnaires concerning speech quality and device form/function to help guide the final steps toward commercial availability of the EMG-EL system.

## V. REFERENCES


[1] R. E. Hillman, M. J. Walsh, and J. T. Heaton, "Laryngectomy Speech Rehabilitation: A Review of Outcomes," in *Contemporary Considerations in the Treatment and Rehabilitation of Head and Neck Cancer*, P. C. Doyle and R. L. Keith, Eds. Austin, TX: ProEd, 2005, pp. 75-90.

[2] R. L. Keith, J. C. Shanks, and P. C. Doyle, "Historical Highlights: Laryngectomy Rehabilitation," in *Contemporary Considerations in the Treatment and Rehabilitation of Head and Neck Cancer*, P. C. Doyle and R. L. Keith, Eds. Austin, TX: ProEd, 2005, pp. 17-58.

[3] G. S. Meltzner and R. E. Hillman, "Impact of aberrant acoustic properties on the perception of sound quality in electrolarynx speech," *J Speech Lang Hear Res,* vol. 48, pp. 766-79, Aug 2005.

[4] P. J. Watson and R. S. Schlauch, "The effect of fundamental frequency on the intelligibility of speech with flattened intonation contours," *Am J Speech Lang Pathol,* vol. 17, pp. 348-55, Nov 2008.

[5] P. J. Watson and R. S. Schlauch, "Fundamental frequency variation with an electrolarynx improves speech understanding: a case study," *Am J Speech Lang Pathol,* vol. 18, pp. 162-7, May 2009.

[6] G. Meltzner, R. E. Hillman, J. T. Heaton, K. Houston, J. B. Kobler, and Y. Qi, "Electrolarynx speech: The state-of-the-art and future directions for development," in *Laryngectomy Rehabilitation*, P. C. Doyle and R. L. Keith, Eds. Austin, TX: ProEd, 2005, pp. 571-590.

[7] J. T. Heaton, E. A. Goldstein, J. B. Kobler, S. M. Zeitels, G. W. Randolph, M. J. Walsh, J. E. Gooey, and R. E. Hillman, "Surface electromyographic activity in total laryngectomy patients following laryngeal nerve transfer to neck strap muscles," *Ann Otol Rhinol Laryngol,* vol. 113, pp. 754-64, Sep 2004.

[8] J. T. Heaton, J. B. Kobler, E. A. Goldstein, T. A. McMahon, D. T. Barry, and R. E. Hillman, "Recurrent laryngeal nerve transposition in guinea pigs," *Ann Otol Rhinol Laryngol,* vol. 109, pp. 972-80., 2000.

[9] E. A. Goldstein, J. T. Heaton, J. B. Kobler, G. B. Stanley, and R. E. Hillman, "Design and implementation of a hands-free electrolarynx device controlled by neck strap muscle electromyographic activity," *IEEE Trans Biomed Eng,* vol. 51, pp. 325-32, Feb 2004.

[10] E. A. Goldstein, J. T. Heaton, C. E. Stepp, and R. E. Hillman, "Training effects on speech production using a hands-free electromyographically controlled electrolarynx," *J Speech Lang Hear Res,* vol. 50, pp. 335-51, Apr 2007.

[11] H. L. Kubert, C. E. Stepp, S. M. Zeitels, J. E. Gooey, M. J. Walsh, S. R. Prakash, R. E. Hillman, and J. T. Heaton, "Electromyographic control of a hands-free electrolarynx using neck strap muscles," *J Commun Disord,* Jan 19 2009.

[12] 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 Trans Neural Syst Rehabil Eng,* vol. 17, pp. 146-55, Apr 2009.

[13] Y. Saikachi, K. N. Stevens, and R. E. Hillman, "Development and perceptual evaluation of amplitude-based F0 control in electrolarynx speech," *J Speech Lang Hear Res,* vol. 52, pp. 1360-9, Oct 2009.