Training Effects on Speech Production Using a Hands-Free Electromyographically Controlled Electrolarynx

Ehab A. Goldstein
Harvard University, Cambridge, MA

James T. Heaton
Massachusetts General Hospital, Boston, and Harvard Medical School, Cambridge, MA

Cara E. Stepp
Harvard–MIT Division of Health Sciences and Technology

Robert E. Hillman
Massachusetts General Hospital, Boston, and Harvard Medical School

Purpose: The *electrolarynx (EL)* is a widely used device for alaryngeal speech, but it requires manual operation and produces voice that typically has a constant fundamental frequency. An electromyographically controlled EL (EMG-EL) was designed and implemented to provide hands-free control with dynamic pitch modulation.
Method: Three participants who underwent total laryngectomy surgery and 4 participants with normal voice were trained to produce EMG-EL speech through a multiple-baseline, successive-stage protocol. Baseline performance was established through 3 testing probes, followed by multiple hour-long training sessions.
Results: At the end of the training, all participants learned to initiate, sustain, and terminate EMG-EL activation in correspondence with articulation, and most were able to modulate the pitch to produce intonational contrasts. After completing the testing/training protocol, 1 of the 3 participants who underwent total laryngectomy was encouraged to independently use the EMG-EL at his residence. This participant sustained his performance for an additional 6 weeks and also used the EMG-EL successfully to communicate over the phone.
Conclusions: Our findings suggest that some participants with laryngectomies and vocally normal individuals can learn to produce hands-free speech using the EMG-EL device within a few hours and that significant additional gains in device control (particularly pitch modulation) are attainable through subsequent training sessions.

KEY WORDS: alaryngeal speech treatment, artificial larynx, electromyography, laryngectomy, speech

Electrolarynx (EL) speech continues to play a major role in laryngectomy rehabilitation; multiple studies report that more than half of patients use an EL for verbal communication after total laryngectomy surgery (Gray & Konrad, 1976; Hillman, Walsh, Wolf, Fisher, & Hong, 1998; Mendenhall et al., 2002; Morris, Smith, Van Demark, & Maves, 1992). Despite the popularity of EL speech, it still suffers from inadequacies that could be addressed by advanced biomedical technology. A key disadvantage in conventional EL speech is the requirement to use one hand to produce it. Every time they want to speak, EL users must retrieve their device and manually apply it to the neck or oral cavity. This action is not only cumbersome and sometimes awkward but also occupies the use of one hand, thus precluding manual tasks that require the use of both hands while speaking. A survey described by Meltzner et al., 2005, found that the inconvenience of use was ranked most problematic by EL users, followed by the monotonic nature of the speech. In addition to

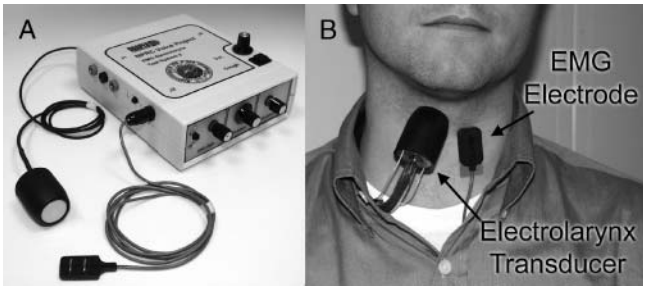
1092-4388/07/5002-0335

occupying one hand, when compared with other alaryngeal voice sources, the EL has been shown to be the least effective of laryngectomy rehabilitation methods in providing contrastive and intonational stress to emphasize different parts of an utterance, mostly due to the constant voice fundamental frequency (pitch; Gandour & Weinberg, 1983, 1984). These previous findings indicate the potential benefit in providing hands-free EL control with the ability to dynamically modulate pitch.

The three main components of a handheld EL are the transducer that mechanically produces the sound, an electric circuit that drives the transducer, and a battery pack that powers the device. Typically, on/off control has been achieved using a thumb button switch that produces a constant pitch. Previous attempts at improving EL devices include modification of the switch for an intraoral EL device so that it could be operated "hands-free" using elbow movements (McRae & Pillsbury, 1979); however, this device still largely occupies the use of one arm and produces constant pitch. A few EL devices provide manually adjustable pitch, such as the UltraVoice (www.ultravoice.com), the Western Electric No. 5, and others (for reviews, see Keith, Shanks, & Doyle, 2005; Meltzner et al., 2005). There have also been a number of attempts to control the pitch of the EL using airflow from the stoma (Takahashi et al., 2001b; Uemi & Takahashi, 1994) and electromyographic (EMG) signals from the neck strap muscles (Min et al., 1994a, 1994b), but success has been limited. Other EL pitch control techniques have included a force-sensitive transducer mounted either on the on/off switch button, such as the TruTone (Griffen Labs) and other experimental EL devices (Choi, Park, Lee, & Kim, 2001; Goldstein, 1998; Takahashi et al., 2001a), or on the sound-source neck interface (Goldstein, 1998). While these EL devices offer variable pitch, they still use a manual switch for on/off control.

We have previously shown that it is possible to provide both hands-free on/off control as well as dynamic pitch modulation of an EL by using voice-related EMG activity as a control signal. Our EMG-controlled EL system (EMG-EL; Goldstein, Heaton, Kobler, Stanley, & Hillman, 2004) utilizes neck muscle electrical signals detected with a skin surface electrode to provide on/off control using a threshold with an internal threshold-dependent hysteresis band. In addition, pitch control is achieved in proportion to the level of suprathreshold EMG energy. The EL transducer is mounted on a brace that is worn around the base of the neck (under clothing) by holding the transducer head (buzzer) to the neck surface for hands-free control (see Figure 1). The motivation for utilizing EMG signals for the voice prosthesis stems from the success of EMG signals in the control of limb prostheses (Jacobsen, Knutti, Johnson, & Sears, 1982; Koike & Kawato, 1995; Latwesen & Patterson, 1994; Saridis & Gootee, 1982; Yamada, Niwa, & Uchiyama, 1983) and from the successful application of EMG signals in human–computer interface technology (Junker, 1995; Scargle, 1998).

Figure 1. (A) The EMG-EL system, showing DelSys surface EMG electrode, the EMG-EL processor, and EL transducer. (B) The transducer and EMG electrode are worn on the neck for hands-free EL speech. EMG-EL = electromyographically controlled electrolarynx. From "Design and Implementation of a Hands-Free Electrolarynx Device Controlled by Neck Strap Muscle Electromyographic Activity," by E. A. Goldstein, J. T. Heaton, J. B. Kobler, G. B. Stanley, and R. E. Hillman, 2004, *IEEE Transactions on Biomedical Engineering*, *51*(2), p. 328. Copyright 2004 by IEEE. Adapted with permission.



This article assesses participant performance by using an EMG-EL system for speech production through a single-participant, multiple-baseline, successive-stage training protocol (Kearns, 2000). Two groups of participants were tested: 4 vocally normal participants (N1–N4) and 3 participants who had undergone total laryngectomy (L1–L3). The participants were tested and trained for their proficiency in using the EMG-EL on voice and speech tasks of supposedly increasing difficulty, including vowel initiation and termination, words, sentences, a paragraph, and intonational contrast production. Aside from guidance around the operation of the EMG-EL during task performance, structured feedback and guidance were withheld from participants for the first three recording sessions to establish baseline performance. The remaining sessions were then used to train participants on each successive skills stage (described in detail shortly). Training was conducted on an individualized basis according to task performance assessed at the outset of each session. Performance levels in relation to training were therefore analyzed for individual participants as well as by participant group.

Method

Participants

Seven participants participated in the study. Four control participants (N1–N4) were graduate students with normal neck anatomy, including 2 men (age 26 years) and 2 women (ages 22 and 23 years) with no history of voice or neuromuscular disorders. Those participants were chosen based on their ready availability. The 3 remaining participants (L1–L3) were men (ages 61, 73, and 58 years) approximately 30, 31, and 32 months after a modified total laryngectomy surgery, respectively.

Normal healthy participants were studied for two purposes: (a) to validate the experimental procedure and the operation of the equipment used for training and (b) to use the normal participant outcomes as a proxy for the potential performance of participants who undergo a modified total laryngectomy in which some neck strap muscles (sternohyoid, sternothyroid, and/or omohyoid) are preserved.

After completing the training protocol for the 4 participants with normal neck anatomy (N1–N4), Participants L1–L3 were recruited for training. These individuals had received an experimental modification to their total laryngectomy surgery that entailed rerouting of the recurrent laryngeal nerve on one side of the neck into a set of host strap muscles through the distal trunk of the ansa cervicalis nerve. The full details of the surgical modifications and their outcomes are described elsewhere (Heaton et al., 2004), where participants L1, L2, and L3 are referred to as Participants TL3, TL4, and TL8, respectively. Participants L1 and L3 initially used a neck EL until they were fitted with a TE valve a few months after the surgery; both participants were highly intelligible with both methods of alaryngeal speech. Participant L2 used a Cooper-Rand mouth-type EL and was highly intelligible with it as well as neck-type devices. All procedures were performed according to protocols approved by the Internal Review Board for human studies at the Massachusetts Eye and Ear Infirmary.

Instrumentation

The EMG-EL system and its components are described in a separate article (Goldstein et al., 2004). Briefly, this system utilizes a surface EMG electrode to detect muscular activity and processes the EMG activity to produce a smooth and proportional EMG envelope. The envelope is then used to both trigger the EL device on and off and vary its fundamental frequency in response to the level of EMG activity. As a result, the EMG-EL enables the user to start and stop the EMG-EL and to vary the fundamental frequency (pitch) of the EL voice source, solely through muscle contraction rather than through manual operation.

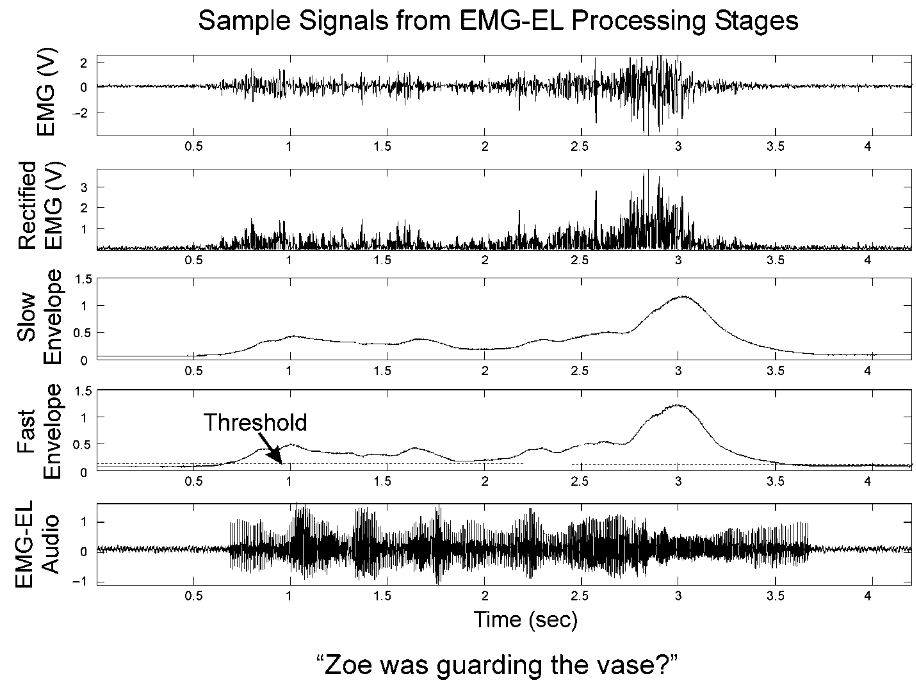
Each participant was seated in a sound-attenuating chamber with a video monitor (Sony Trinitron GVM 2020) placed 1 m away. Speech signals were recorded with a condenser microphone (Sony ECM-50PSW) suspended in front of the patient at approximately 30° from the sagittal plane and approximately 15 cm from the mouth. The video monitor was used to present the stimulus material to the participant as well as the reaction time scores during the training sessions as a form of performance feedback. Data acquisition and analysis for vocal reaction times were performed using a PC. MATLAB software controlled the PC sound card (Aureal Semiconductor, Vortex AU8830 PCI) as a two-channel input device. For timing accuracy, a photocell was mounted on the video monitor to detect the timing of visual stimulus presentation. The photocell output was digitized and recorded along with speech signals. Using the acoustic signal from the microphone and the electrical signal from the photocell, the time delays between the stimulus and voice initiation and termination were measured with a resolution of 1 ms. In addition, the video output of the computer was mixed with video of the participant and was recorded on tape for judging the synchronization between voice timing and articulation offline.

EMG signals were detected with a skin surface electrode (DE2.1, DelSys Inc.) placed over the neck strap muscles. The EMG electrode was positioned along the midline of the neck, directly below the thyroid prominence of the vocally normal participants, partially overlying the strap muscles on both the right and the left (sternohyoid, sternothyroid, and omohyoid). Although the neck strap muscles were assumed to have symmetrical activation patterns, the midline electrode position was intended to mitigate potential asymmetry in strap muscle activation and provide maximal spatial displacement from the sternocleidomastoid muscles located laterally on each side. To accommodate their altered neck anatomy, in the participants with laryngectomies, the EMG electrode was positioned lateral to the neck midline directly above the RLN-innervated strap muscles (sternohyoid, sternothyroid, and/or omohyoid) in their particular position after laryngectomy. An Ag/AgCl gel ground electrode (Kendall LTP) was placed on the superior surface of the shoulder. The EMG signal was processed by the EMG-EL device to control initiation, termination, and pitch of prosthetic voice (see Figure 2). The threshold was set at the baseline +10% of the observed voice-induced envelope amplitude range. For example, if the baseline was 30 mV and the maximum EMG envelope was 500 mV, then the threshold was set at 30 mV + (500 – 30) / 10 mV = 77 mV, with an accuracy of 2 mV. Maximum EMG activity from the neck strap muscles was elicited by asking participants to produce (normal participants) or visualize (laryngectomy participants) a low-pitch sustained vowel sound. The EMG and audio signals were recorded on an 8-channel digital audio tape for backup (TEAC Model RD-111T).

Stimulus Materials

Participants were tested and trained on seven successive stages for the production of fluent EMG-EL speech,

Figure 2. Sample signals from the EMG-EL, showing the band-pass filtered EMG from the neck strap muscles of participants with normal voice uttering the question "Zoe was guarding the vase?" using the EMG-EL. The corresponding rectified EMG signal is shown, along with the slow envelope controlling EMG-EL voice pitch and the fast envelope triggering the device on and off. The bottom trace shows the EMG-EL audio output.



starting with (a) vowel initiation, (b) vowel duration, and (c) vowel termination, and then with (d) words, (e) sentences, (f) a paragraph, and (g) intonation contrasts. The goals and criteria for each stage are described in detail next.

A reaction time procedure was used to test vowel initiation, duration, and termination in a fashion similar to previous studies on stuttering (Watson, 1994). Within each trial, a sequence of visual commands with different background colors was displayed on the video monitor. Each trial started with a rest period of 10 s, followed by a get-ready period of 1 or 2 s. The variable ready period was chosen to reduce the effect of the participant's anticipation (Izdebski, 1980; Watson, 1994). The ready period was followed by the response cue, in which the display showed the command, "Say /a/" on a bright-green background. The response interval was 2, 2.5, 3, 3.5, or 4 s. Ending the response period was the stop signal, which was a display of the word *Stop* on a red background for 2 s. The stimulus sequence was repeated in random order for the five response intervals and two ready periods, yielding 10 trials per run.

To elicit words and sentences, materials from the Yorkston and Beukelman Assessment for Intelligibility of Dysarthric Speech (Yorkston & Beukelman, 1981) were used. The test included a pool of 600 words, grouped into 50 subsets of 12 similar-sounding words. Each of the 10 stimulus words used in the present study were randomly chosen from the pool of 600 by first randomly selecting one of the 50 sets then randomly picking 1 of the 12 words within that subset. Random selection was performed using a built-in function in MATLAB.

The Yorkston and Beukelman test also includes 10 groups of 100 sentences each, in which the number of words per sentence (6–15) is constant within sentence group. One sentence was randomly chosen from each of the 10 groups of constant word-length sentences, resulting in a set of 10 random sentences with varying word lengths to be used for testing the participants during each session.

The paragraph used in this study was the first section of *The Rainbow Passage* (Fairbanks, 1960), which contains 10 phrases separated by commas or full stops. Intonation was tested using a pool of 20 short declarative sentences that could be spoken as statements or as questions, depending on the intonational pattern imposed. These short sentences were composed entirely of voiced phonemes in order to avoid voiced/unvoiced ambiguity while using the EMG-EL. The words, sentences, and intonation exercises were performed by displaying the stimulus material on the video monitor in front of the participant, whereas the paragraph was read from a printed page given to the participant.

Training Protocol

The experimental protocol involved ten 10- to 60-min sessions, performed at the same time each day for 10 consecutive days. A probe designed to assess performance was administered during the first 10 min of each session. Probes were composed of one exercise from each of the seven skills and were administered with basic instructions (see Appendix A) but without any structured feedback. The exercise for each of the seven skill types was scored on a scale of 10 (scoring was dichotomous for each utterance, but a total of 10 points was possible across a set of 10 utterances). The first four sessions were solely aimed at establishing baseline performance and were thus composed of the probe exercises without any training or feedback. For the purposes of the study, baseline performance was defined as the extend to which neck surface EMG signals spontaneously provided accurate control of the EL sound source during speech tasks prior to any training or specific orientation to device operation. Training was started after the fourth probe administration and involved approximately 50 min during which participants repeated the exercises while receiving structured feedback about performance and instruction on how to improve performance. Each participant initially trained on vowel initiation, then moved on to the successive stages once an 80% success rate was achieved at each level, either during the probe exercise or during training. If a participant managed to perform all skills successfully at some point before the end of the study, training was continued on the weakest skills through the remaining sessions. On the other hand, if a participant showed no signs of improvement in performance over three consecutive training sessions in a certain skill, the training was automatically advanced to the next skill in order to ensure that all skills received training by the end of the experiment. Such automatic advancement was only necessary for all three vowel sound skills of Participant L2 and for the vowel termination skill of Participants N3, N4, and L1–L3. In the event of automatic advancement, training was moved on to the next skill in addition to maintaining the training on the current skill throughout subsequent sessions. A final performance measurement probe was done at the end of the 10th training session (i.e., probes were run at both the start and end of the last session).

Participants N1–N4 and L1 all underwent the 10-day protocol. However, 2 of the participants with laryngectomies (L2 and L3) lived prohibitively far from the laboratory site, so their testing/training sessions were reduced from 10 sessions over 10 days to 4 sessions conducted over a 2-day period. Three probes were administered in the first session of their abbreviated protocol (separated by 15-min breaks), followed by 4 probes punctuating 3 training sessions each 1 hr in duration (separated by breaks of at least 1 hr). Therefore, 2 of the 3 participants with laryngectomies were tested with 7 probes instead of the typical 10 probes and received 3 hr of training instead of the typical 7 hr due to constraints in their availability. To accommodate this shortened training period, automatic advancement to the next skill because of insufficient performance improvement was allowed after 1 training session rather than the typical 3 sessions.

In summary, each skill was tested with 10 tokens, and each token was judged based on the rules described in Table 1, which are discussed further in the next section.

Training Criteria and Techniques

Vowel initiation. The aim in this initial stage was to teach the participant how to voluntarily produce adequate neck muscle EMG activity to start the device. The training focused on how to contract the neck muscles to turn the device on quickly and consistently. In general, participants were encouraged to visualize producing low-pitch and/or loud speech when controlling the EMG-EL prosthesis. Moreover, participants were reminded to relax their laryngeal and neck muscles when not trying to activate the voice prosthesis in order to facilitate rapid device termination and minimize unintentional triggering. The metric for vowel initiation performance was the ability to initiate EMG-EL voicing within a certain time window after a visual stimulus was issued. Results of a previous experiment with 7 adults with normal voice using

Table 1. Criteria used for judging EMG-EL speech tokens.

Skill	Criterion for Success
Vowel initiation	0 ms < VIT < 390 ms
Vowel duration	Device on from onset to stop signal, in addition to successful vowel initiation
Vowel termination	0 ms < VTT < 330 ms, with successful vowel initiation and duration
Words	Device triggered at beginning of articulation and sound maintained throughout the whole word
Sentences	Same as words, but pauses between words were allowed
Paragraph	Same as sentences
Intonation	Sentences must be read with correct triggering, and pitch must rise at the end of a question relative to a statement

Note. EMG-EL = electromyographically controlled electrolarynx; VIT = voice initiation time; VTT = voice termination time.

an identical testing protocol in our lab yielded average voice initiation time (VIT) values of 270 ±60 ms (Goldstein et al., 2004). This finding agreed with a published report (Watson, 1994) that used similar methods and found VIT values for producing a vowel of approximately 283 ms and 266 ms at ready periods of 1 and 2 s, respectively. On the basis of these results, we chose a VIT response window for successful voice initiation to be the normal-voice VIT mean plus two standard deviations (270 ms + 120 ms = 390 ms). If the participant managed to achieve a VIT that was less than 390 ms, then the token was scored as a success. Otherwise, tokens with VIT values that were negative or longer than 390 ms were scored as failures. Here, negative VIT values correspond to false starts of the device. While scoring was dichotomous for each utterance, the score out of 10 points was possible by looking across a set of 10 utterances.

Vowel duration. Once the participant was able to voluntarily contract the neck muscles to turn the device on, the next goal was to achieve sustained phonation. The stimulus length was set for 2, 2.5, 3, 3.5, and 4 s of voicing, which was within the range of times reported for average reading breath-group duration (Hodge & Rochet, 1989; Hoit & Hixon, 1987; Hoit, Hixon, Altman, & Morgan, 1989; Mitchell, Hoit, & Watson, 1996; Solomon & Hixon, 1993). The metric for success included starting with a successful onset followed by sustained voicing throughout the stimulus until the stop command was issued. Training focused on maintaining the EMG activity level high enough during the stimulus period in order to keep the EMG-EL device on.

Vowel termination. After gaining the skills to both trigger and maintain the output of the device, the training focused on appropriate device output termination, intended to achieve meaningful interword pauses during running speech. The metric involved starting the device with a successful onset, maintaining the voicing throughout the stimulus, and then stopping within a set time window of 330 ms after the visual stop signal. This criterion was chosen based on the mean plus two standard deviations (240 ms + 90 ms = 330 ms) of voice termination time (VTT) measured in a previous study of 7 participants with normal voice tested with an identical protocol (Goldstein et al., 2004). The training at this stage focused on how to relax the neck muscles in a timely fashion when intending to stop vocalization.

Words. This stage aimed at combining the skills of EMG-EL vowel voicing with word articulation. Using word lists from the Yorkston and Beukelman (1981) intelligibility test, the participant was prompted to read 10 randomized single words using the EMG-EL device. Word production was evaluated by the first author as the words were produced. The investigator judged the correspondence between articulation and EMG-EL voice production by viewing the participant's mouth and listening to the EMG-EL audio output. The success of each token was measured using three metrics: (a) timely EMG-EL onset perceived as simultaneous with the articulation of the first phoneme in the word, (b) uninterrupted EMG-EL voicing throughout the articulation of the word, and (c) EMG-EL voice termination perceived as occurring after articulation of the last phoneme in the word. The investigator's scoring of all speech material was evaluated at a later time by having a certified speech-language pathologist independently judge approximately 7% of the speech samples using an offline video recording of the session (see the *Score Verification* section). Note that the electrolarynx sound transducer does not provide the user with the ability to produce unvoiced phonemes. As a result, this stage (and all subsequent stages) did not evaluate the ability of the participant to produce unvoiced phonemes. Rather, the aim was to get the participant to utter the words clearly while voicing throughout both the voiced and the unvoiced phonemes within each word.

Sentences. In a fashion similar to the previous stage, randomized sentences from the Yorkston and Beukelman (1981) intelligibility test were used as stimulus material to be read by the participant. The metrics for success were the same as those used for words, for which the participant was required to produce EMG-EL voicing while articulating multiple words, with the exception that pauses were allowed between words within the sentence. As a result, EMG voice interruptions that occurred during intended word vocalization were scored as a failure. The stimulus sentences were of different lengths (6 words to 15 words). The participants were asked to read through the sentences without trying to repeat any parts if they failed to trigger the EMG-EL. The focus of the training in this stage was to maintain the suprathreshold EMG activity throughout the words within the sentences.

Paragraph. The paragraph read by participants was the first paragraph of *The Rainbow Passage* (Fairbanks, 1960), which is composed of five sentences divisible into 10 phrases. Performance was tested in a similar fashion to the sentences stage, with attention paid to the participant's ability to string multiple phrases together in an accurate and timely fashion. The participant was instructed to read the paragraph from a printed page provided at the time of the exercise. The exercise was scored based on the number of phrases uttered successfully. The phrases were deemed successful if there were no interruptions in voicing during any of the words. Voicing pauses were allowed between words and between phrases.

Intonation. At the sentence level, the ability of the participant to inject prosodic information into EMG-EL speech was tested using sentences uttered first as a statement and then as a question. The protocol used was based on the work of Gandour and Weinberg (1983, 1984) that assessed prosodic control for various forms of alaryngeal speech. During the intonation exercise, 10 out of the pool of 20 sentences were randomly selected and presented on the computer monitor to the participant, first as a statement ending with a period and then as a question ending with a question mark.

In addition to timely EMG-EL voice output (as for the sentence and paragraph exercises), the metrics for success included the requirement of an audibly perceptible variation in the fundamental frequency toward the end of the sentence to indicate whether it was a statement or a question. For each sentence, the question version was compared with the statement version by the first author, and the token was considered a success only if the question had a rising fundamental frequency toward the end of the sentence when compared with the statement.

Training at this stage emphasized the EMG-EL built-in proportional relation between the muscle contraction and the fundamental frequency (see the *Instrumentation* section). The feedback component of this exercise comprised of the pitch variation in the sound output of the EMG-EL corresponding to the level of suprathreshold EMG activity.

Score Verification

Because of the subjective nature of the judging for the words, sentences, paragraph, and intonation, the reliability of the real-time judgments of the first author was tested by randomly selecting 1 out of the 11 probes for 5 of the participants (N1–N4 and L1) were judged independently by a speech-language pathologist with 9 years of clinical experience at the Massachusetts Eye and Ear Infirmary's Voice and Speech Lab. The clinician was familiar with EL speech and the scoring criteria for each task, but she was not aware of the training phase or experience level of participants for each judged token (i.e., not biased toward scoring participants more favorably through the course of their training). The criteria used to verify the scores can be found in Appendix B. Two hundred tokens were checked in total (4 exercises × 10 tokens per exercise × 5 participants), resulting in 93% agreement between the first author and the speech-language pathologist. The 14 mismatches in scoring occurred as follows: 4 in words, 7 in sentences, 2 in the paragraph, and 1 in intonation.

Results

General Outcome

The participants using the device in this study generally responded positively about its hands-free nature and the quality of voicing once triggered. However, participants also mentioned frustration with EMG electrode contact instability (generating unintentional, high-pitch device activation), and the participants with laryngectomies, in particular, expressed displeasure with lingering voice associated with termination difficulty. The experimental outcomes are based on the probe scores of each participant in each of the exercises. Because of the individualized nature of the training, the data are first presented and analyzed for each participant. Figure 3 shows the probe scores for the Participants N1–N4, and Figure 4 shows the data for Participants L1–L3. In this analysis, multiple t tests were conducted without control for alpha inflation. Although this statistical approach is generally a cause for caution, it can be justified in this case by the early stage of inquiry and small sample sizes. Furthermore, the multiple t tests conducted were two-tailed, which partially mitigated alpha inflation associated with multiple statistical comparisons.

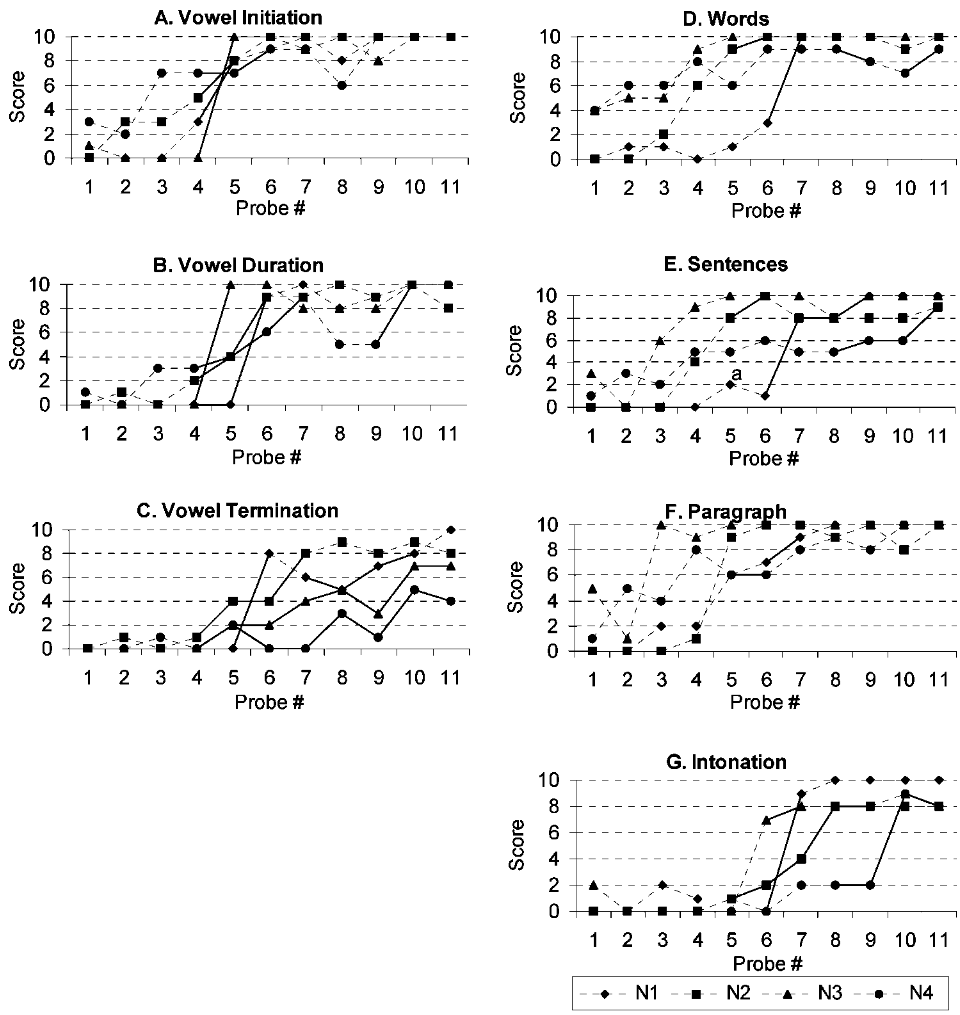
The criterion for success within each skill was the achievement of an 80% (8 out of 10) score. As can be seen in Figure 3, all participants in the normal-voice group managed to reach the 80% score in all the tasks, except for Participants N3 and N4 in vowel termination. For these 2 participants, the highest vowel termination score was 70% for N3 and 50% for N4. Formal training for each participant is indicated by a solid line connecting the scores for probes given before and after the indicated training. The positive effect of training on performance is clearly seen in the positive slope of the training lines, in most cases. It is interesting to note that a significant amount of learning took place without the application of formal training, especially in the tasks of the words, sentences, and paragraph. For example, Participants N2, N3, N4, and L3 all managed to achieve an 80% score in the words and the paragraph exercises before any training was administered for those skills.

Two out of the 3 participants with laryngectomies (L1 and L3) learned how to start and maintain the EMG-EL device output with reaction times that were similar to those in the normal-voice group, but all 3 failed to stop EMG-EL voicing as fast as those in the normal-voice group (see Figure 4). In addition, all 3 participants with laryngectomies achieved the success criterion in the words, sentences, and paragraph by the end of the experiment, and 2 these 3 participants (L2 and L3) scored 50% or more even before training was administered. The same 2 participants (L2 and L3) also managed to score 80% or more on intonation after some training, whereas 1 participant with a laryngectomy (L1) was only able to score a maximum of 20% on intonation control.

Training Effect

To compare pretraining with posttraining performance statistically, within-participant probe scores were tested for the first three sessions (pretraining) versus the last three sessions (posttraining) for each skill. The last three probes indicated Sessions 9, 10, and 11, except for Participants L2 and L3, in which it implied Sessions 5, 6, and 7. Across-participant comparisons were also conducted

Figure 3. Scores of Participants N1–N4 over the 11 probes for (A) vowel initiation, (B) vowel duration, (C) vowel termination, (D) words, (E) sentences, (F) paragraph, and (G) intonation. Points connected with a solid line indicate that formal training was administered between the two probes for that skill.

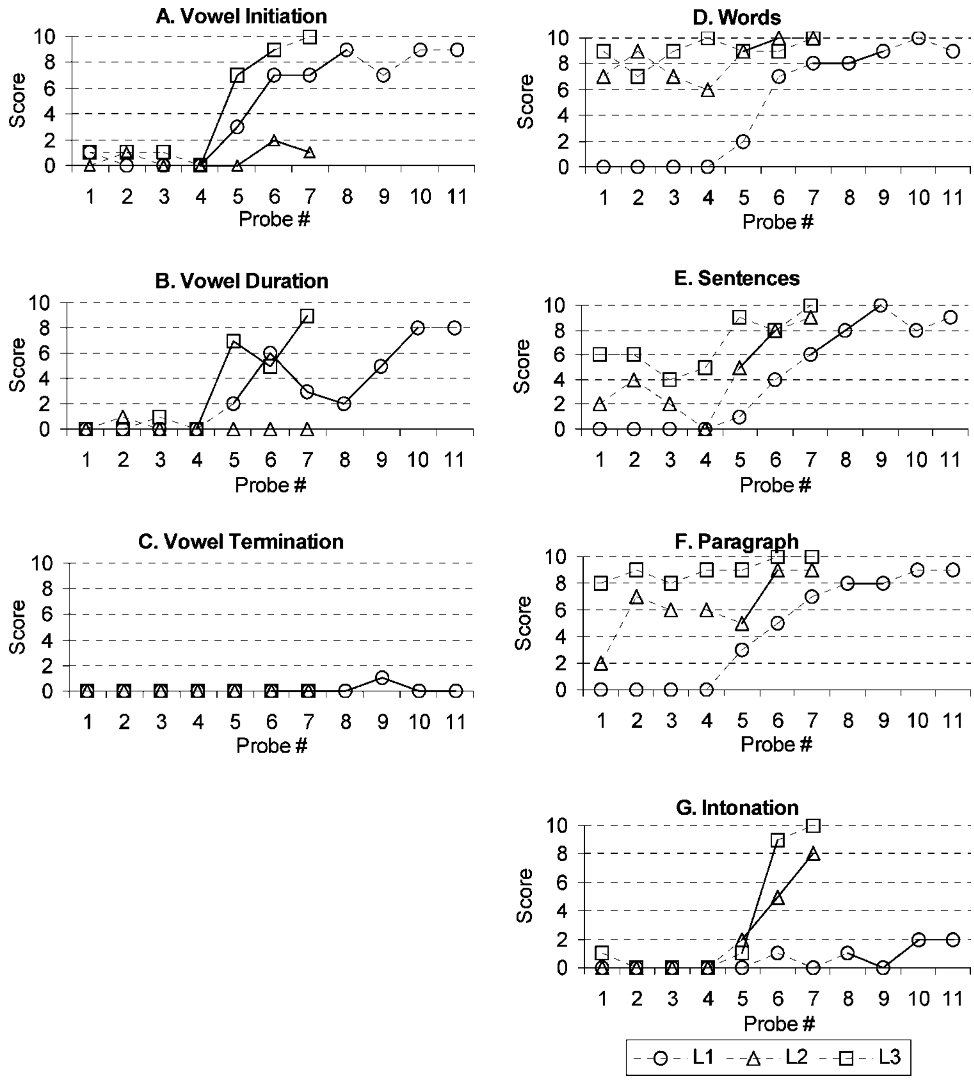


by grouping Participants N1–N4 and Participants L1–L3 to test for general training effect. A significant difference ($p < .05$, two-tailed t test) in the pre- versus the posttraining average scores was used as an indicator for the presence of a training effect. In most cases, the absence of a training effect was the result of either a low posttraining average score or a high pretraining score, potentially due to spontaneous learning. Figure 5 summarizes the results by showing the individual participant and the group average scores. High pretraining and low posttraining scores that led to insignificant training effects are labeled with their value next to the symbol.

For Participants N1–N4, the training effect was statistically significant for each of the seven skills at the group level. The training effect was also statistically significant for most within-participant individual scores. The results of participant unpaired t tests, assuming unequal variances, are shown in Table 2. Exceptions to statistical significance are Participant N4's vowel initiation and Participant N3's sentences and paragraph, which failed to differ from pre- to posttraining scores. Low posttraining scores precluded a significant training effect for vowel termination of Participants N3 and N4 (which never reached 80%) and intonation of Participant N4.

Participants L1–L3 achieved scores that showed a statistically significant group training effect for vowel initiation, duration, words, sentences, and paragraph. Again, t-test results are catalogued in Table 2. Vowel termination and intonation exercises proved challenging for all participants in this group, as reflected by the lack of a statistically significant training effect for all 3 participants. However, Participants L1 and L3 managed to achieve a significant training effect for vowel initiation and duration. All 3 participants achieved the 80% score in words,

Figure 4. Scores of the 3 participants N1–N3 with laryngectomies for (A) vowel initiation, (B) vowel duration, (C) vowel termination, (D) words, (E) sentences, (F) paragraph, and (G) intonation. Points connected with a solid line indicate that formal training was administered between 2 probes for that skill. Participant L1 underwent 10-day training with 11 probes and 7 hr of training, whereas Participants L2 and L3 took part in a 2-day intensive version of the training that had 7 probes and 3 hr of training.



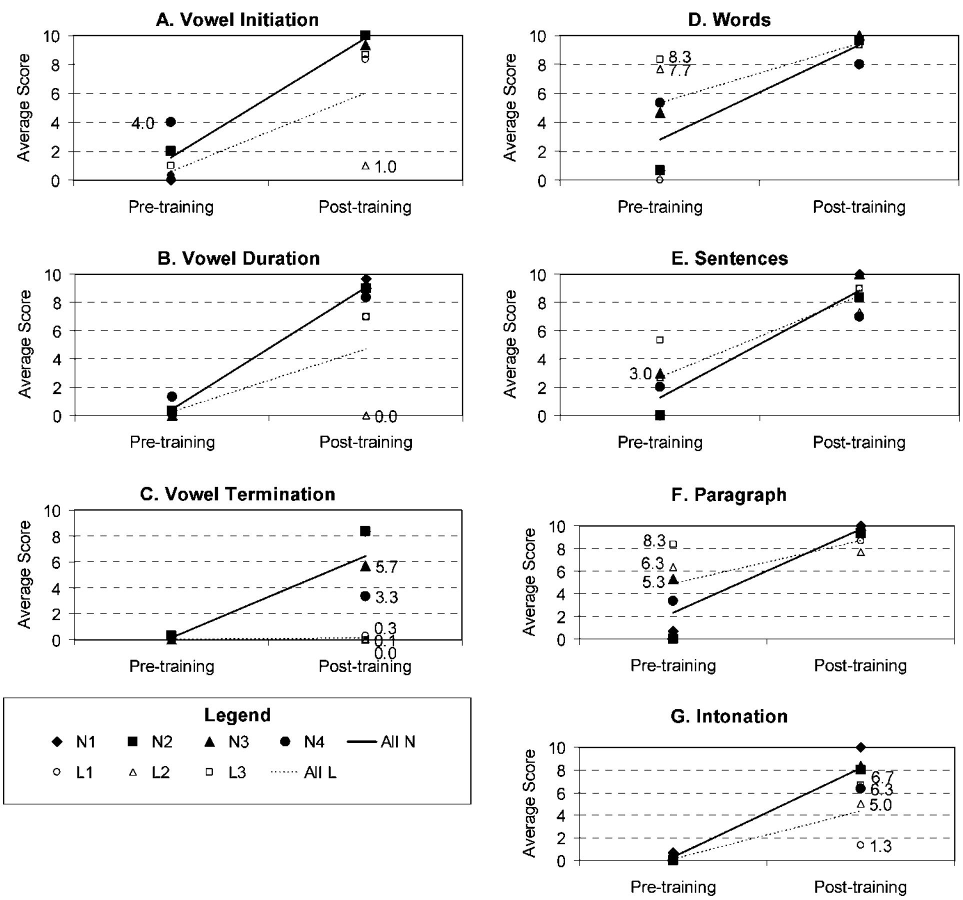
sentences, and paragraphs, but Participants L2 and L3 did not show a significant training effect due to the high pretraining scores in words and paragraphs.

Based on the numerous possible causes for a relatively high failure rate for termination (see the Discussion section), the termination criterion of <330 ms was arguably too strict. Thus, in order to more usefully compare the termination performance of participants with laryngectomies versus performance of participants with normal voice, termination performance was plotted across the continuum of possible termination criteria (see Figure 6). As shown in the plot, voice termination performance of the participants with laryngectomies did not reach 100% success until the criterion was 1,970 ms and reached 80% success 1,120 ms after the cue to terminate was given. Conversely, the performance of participants with normal voice reached 100% success when the criterion was 1,010 ms and reached 80% success at a termination criterion of 405 ms. Therefore, VTT times for participants with laryngectomies were approximately double those of participants with normal neck anatomy.

Practical Use of the EMG-EL Prosthesis

After completing 10 sessions of testing and training, Participant L1 was given an EMG-EL prosthesis to take

Figure 5. Effect of training on average scores. Symbols indicate individual pre- and posttraining scores of normal participants (N1–N4, solid symbols) and participants with laryngectomies (L1–L3, hollow symbols). The solid line indicates the trend across participants with normal voice, and the dotted line indicates the trend across subjects with laryngectomies. Most participants showed a significant training effect for most skills (paired *t* tests, $df = 3$, $p < .01$). Data labels indicate insignificant change across training, presumably because of either high pretraining scores or low posttraining scores.



home. The goal was to test the practicality and feasibility of a participant donning the device and using it without any assistance. For a period of approximately 1.5 months, Participant L1 used the EMG-EL between one and five times a week for periods ranging from 30 min to a couple of hours and recorded his comments on the challenges and successes in a log book. Participant L1 was also urged to call the first author on the phone once or twice a week using the EMG-EL. The purposes of these calls were twofold: (a) to assess the participant's ability to attach and

Table 2. Results of participant unpaired *t* tests, assuming unequal variances, first three probes versus the last three probes.

Participant	Initiation	Duration	Termination	Words	Sentences	Paragraph	Intonation
N1	<.01	<.01	.01	<.01	<.01	.01	.01
N2	.02	<.01	<.01	<.01	<.01	.01	<.01
N3	<.01	.01	.05	<.01	.06	.21	<.01
N4	.06	.03	.14	.04	.02	.02	.10
L1	<.01	.02	.42	<.01	<.01	<.01	.18
L2	.39	.42	—	.07	.04	.65	.10
L3	.01	.03	—	.27	.01	.05	.16

Note. Dashes indicate that the participant scored all zeroes over the course of the first three as well as the last three probes.

Figure 6. Termination performance of participants with laryngectomies and participants with normal voice plotted as a function of variable termination criteria. Termination performance is calculated as the percentage of vowel tokens passing at each possible voice termination time criterion level out of the total number of vowel tokens that had met onset and duration criteria across all probes.

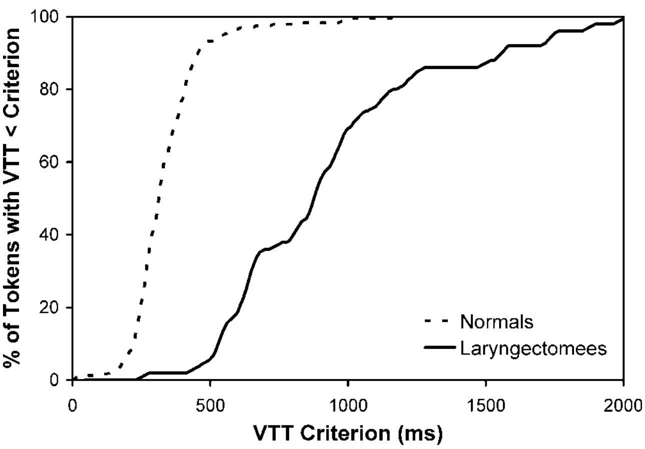
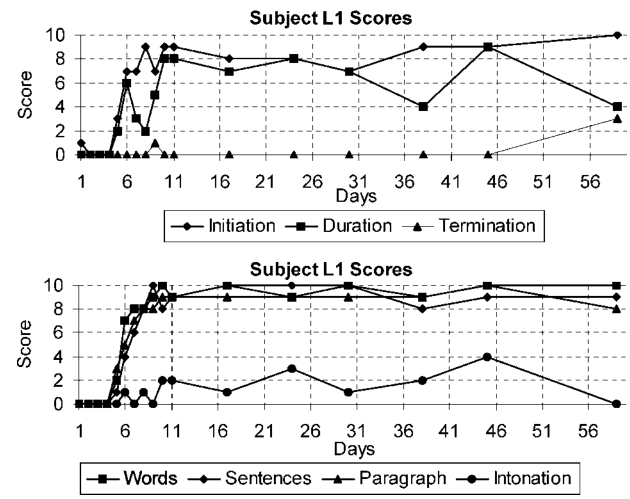


Figure 7. Scores of Participant L1 during the 10 days of initial training and through the posttraining monitoring period that continued for a month and a half after the end of training (Day 11 onward). Results show the maintenance of the posttraining scores for each of the skills learned during the training period.



set up the device correctly and (b) to evaluate subjectively the participant's EMG-EL communication ability through a short conversation. Through the phone calls, the participant was also provided with additional guidance in attaching and operating the device whenever needed. On the basis of the phone conversations, Participant L1 demonstrated the independent ability to attach the electrodes, set up the EMG-EL parameters, and use the device for an intelligible phone conversation. In addition to the phone conversations, the performance of Participant L1 was formally assessed through probe exercises administered at the lab on a weekly basis to track posttraining skill levels.

At all follow-up intervals, Participant L1's scores in vowel initiation, duration, words, sentences, and the paragraph were sustained at their high posttraining levels (see Figure 7) without any additional formal training sessions. Similarly, the skills of vowel termination and pitch modulation for intonational stress that were not achieved during the training period did not seem to change over the posttraining period. These results indicate that EMG-EL performance reached a stable plateau by the end of the 10 sessions of formal training for Participant L1.

Training Intensity

Table 3 shows the number of training sessions in which a certain skill received formal training, as a percentage of the total number of possible training sessions, which was seven sessions for Participants N1–N4 and L1 but only three for Participants L2 and L3 (see the Method section). The results in Table 3 shed light on the fact that the skills receiving most of the attention during training were vowel termination, duration, and initiation, with vowel termination receiving training during more than 80% of the training sessions. The skills receiving the smallest amount of training, on the other hand, were those related to paragraph, words, and sentences. The intonation skill fell somewhere in the middle, on average, in terms of relative training intensity. Note that during a training session, formal training exercises were administered for multiple skills, depending on the progress of the participant and his or her ability to reach the 80% criterion. As a result of the overlap in training on multiple skills, the results in Table 3 do not add up to 100% for each participant across the seven skills.

Discussion

The fact that all participants in this study achieved a high degree of proficiency in the production of words, sentences, and the paragraph demonstrates their ability to learn quickly to use their neck muscles to control the EMG-EL for basic speech production. The success of all but 1 of the participants (L1) in the intonation exercise also indicated that learning to modulate the pitch of the EMG-EL voice was likewise typically achieved within the training period. Statistical testing for learning effects across participants revealed a significant improvement in the average scores of each of the skills for both groups, with the exception of vowel termination and intonation for L1–L3. Vowel termination was the most difficult skill to acquire for both groups, although that parameter received

Table 3. Number of sessions in which training was administered as a percentage of total number of potential training sessions.

Subject	Initiation	Duration	Termination	Words	Sentences	Paragraph	Intonation
	Percent of potential training sessions used						
N1	14%	29%	57%	4%	43%	14%	14%
N2	14%	29%	43%	4%	14%	14%	57%
N3	14%	14%	100%	0%	0%	0%	14%
N4	29%	57%	100%	9%	43%	0%	43%
Average N	**18%**	**32%**	**75%**	**4%**	**25%**	**7%**	**32%**
L1	57%	71%	71%	9%	29%	14%	43%
L2	100%	100%	100%	3%	33%	33%	67%
L3	67%	100%	100%	0%	0%	0%	33%
Average L	**75%**	**90%**	**90%**	**1%**	**21%**	**16%**	**48%**
Average ALL	**42%**	**57%**	**82%**	**7%**	**23%**	**11%**	**39%**
Rank	**5**	**6**	**7**	**2**	**3**	**1**	**4**

Note. Seven training sessions were used for Participants N1–N4 and L1 and for Participants L2 and L3. The rank order of the skills is deduced from the average across all 7 participants, showing the areas of focus for training throughout the experiment. Words, sentences, and paragraph received the least amounts of training, relative to all the other skills.

the greatest average amount of dedicated training time. Possible factors influencing skill acquisition for EMG-EL voice activation and pitch control, and the role of training in performance improvements, are discussed in the next three sections.

Control of Voice Onset/Offset

Rapid voice termination was the most difficult skill for both participant groups to acquire. One impediment to rapid termination of EMG-EL activation was a circuit design feature that produced an internal threshold-dependent hysteresis band. This threshold-dependent hysteresis caused the threshold for termination to be some degree lower than the threshold for onset (based on the initial setting of the onset threshold) to avoid rapid oscillation in device output and to facilitate continuous output throughout intended voicing. However, the improved maintenance of intended voicing afforded by the hysteresis band potentially came at the expense of inappropriately prolonged output, resulting in increased termination error. Our initial experience with participant use of the EMG-EL indicated that people did not view the two types (Type I and Type II) of device output error equally. Persons with laryngectomies and anatomically intact individuals preferred increased unintentional buzzing (Type I error) over device cutouts (Type II error; Goldstein et al., 2004), and the threshold hysteresis band biased error toward this preference. Moreover, while this control strategy is biased toward producing Type I error, its reduction in Type II error is such that the total error function has the potential to be consistently less than in the single-threshold case (Goldstein, 2003). Further studies are needed to systematically assess participant performance under different thresholding strategies.

A further fact that delayed voice termination was the low-pass filtering of the EMG signal. This additional delay was the result of the corner frequency of the three-pole low-pass filter (1–9 Hz range) installed in the EMG-EL to achieve the smoothness of the EMG envelope needed to reduce jitter in EMG-EL operation. Simulations of the three-pole low-pass filter across the variable range with a step function input showed potential delays of up to 160 ms before decaying down to 1/e (37%) of the step input. Based on exponential decay with a time constant of 160 ms, the delay imposed by the filter is approximately 368 ms. This value represents the time it takes the energy to decay from the peak level used to determine threshold to a level at threshold, which is 10% above baseline. However, this potential delay of 368 ms is a maximum value that would occur only under conditions in which EMG decay started from a peak level, which was rarely noted toward the end of most declarative utterances. Question intonations, on the other hand, would be affected to a greater degree by the filter delay due to a relatively high-terminal EMG level. Prolonged EMG-EL activation during question production was not particularly problematic for our participants, showing that termination delays did not preclude device usefulness even in this context.

The VTT data for vowel production presented in Figure 6 reveal that the low-pass filter delay was often negligible when recognizing that VTT is a combination of filter decay and reaction time to the voice termination prompt. If voice termination reaction times for controlling neck muscles resemble reaction times for hand muscles (voice offset via EL button release), then reaction time delay would contribute approximately 315 ms to VTT based on our prior results using the same reaction time task (Goldstein, 2003). Given that normal participants terminated EMG-EL voice within 315 ms for 52% of their vowel tokens, it appears that the filter delay was often slight for this group. Participants with laryngectomies, however, might have experienced greater filter delay effects if their EMG decay starting point had been closer to their peak level. These participants had, on average, lower peak EMG levels and, therefore, less dynamic range for EMG-EL control compared with anatomically normal participants. The relatively reduced peak levels in participants who underwent laryngectomy were likely due to a combination of factors, including participant age and strap muscle trauma sustained during surgery and radiotherapy. If the reduced EMG range caused the participants with laryngectomies to maintain a habitual EMG level closer to their peak values during vowel production, then the approximately twofold difference between termination latencies of normal and laryngectomy participants shown in Figure 6 may have been at least partially due to a difference in the impact of the filter delay.

The EMG-EL required neck muscle relaxation for voice termination, despite the fact that normal voice termination is typically an active process rather than simply a passive process. Although this approach may seem reasonable (turn the EMG-EL off by stopping the effort to turn it on), it does not take into account the physiological role of the nerves and muscles contributing to the EMG control signal. For example, we normally stop voice production not only by reducing subglottic air pressure but also by RLN-mediated contraction of muscles that abduct or forcefully adduct the vocal folds to stop their vibration. Therefore, attempted laryngeal strategies for actively terminating voice may have inadvertently prolonged EMG activity (and thus EMG-EL voicing) in our participants with laryngectomies due to the RLN innervation of their recorded neck strap muscles. In contrast, naturally innervated strap muscles are not known to play a role in voice termination, so the difference in strap muscle nerve supply between the two participant groups could indeed have contributed to group differences in vowel termination performance.

Fortunately, the difficulty faced by most participants in achieving normal vowel sound termination did not seem to pose a problem during the subsequent word, sentence, paragraph, and intonation tasks, indicating that the VTT target might have been more stringent than what is required to produce conversational speech. This outcome may be attributed to the ability of the participants to more easily anticipate pauses during natural speech (words, sentences, and paragraph) and, therefore, to adjust neck muscle activity to produce timely voice termination. The ease with which participants learned to produce natural speech also implies that the integration of articulation movements to transform EMG-EL voice into speech was intuitive and easy to learn.

Control of Vocal Intonation

Participants N1–N4 managed to inject enough pitch modulation into their speech to allow listeners to distinguish a statement from a question. The average posttraining score for the normal-voice participants in the intonation exercise was 82%, which was significantly larger than the average pretraining score. Two of the 3 participants who underwent laryngectomy surgery also reached the 80% success criterion for intonation. Participant L1, however, only managed to get a few tokens right after three training sessions, thus yielding an average posttraining score of 13%. The limited ability of Participant L1 to produce EMG-EL intonation might be explained by the small size of the preserved neck strap muscle (omohyoid), in his particular case, and the correspondingly challenging task of generating a wide range of EMG activity from this limited muscle mass. Participants L2 and L3, on the other hand, had much more muscle mass because of the preservation of all medial neck strap muscles (omohyoid, sternothyroid, and sternohyoid) ipsilateral to the EMG recoding site.

The success of Participants N1–N4 in producing EMG-EL intonation might have been due to their intact set of naturally innervated strap muscles, which normally exhibit a range of EMG amplitudes correlated with voice pitch (Erickson, 1993; Goldstein et al., 2004; Hirano, Koike, & von Leden, 1967; Hirano, Ohala, & Vennard, 1969; Roubeau, Chevrie-Muller, & Lacau Saint Guily, 1997). As a general trend, the activity of the neck strap muscles recorded in this study is usually reported to be inversely related to vocal pitch, in which greater levels of EMG activity are recorded during lower pitch utterances and vocalizations. The interesting point here is that even though the EMG-EL is designed with a linear positive correlation between EMG activity level and EMG-EL pitch, these participants were able to learn to contract their neck muscles in a fashion that enabled them to modify the EMG-EL pitch to reflect the correct declarative versus interrogative intentions (i.e., elevate strap muscle activity to generate high-pitch prosthetic voice rather than low-pitch voice, as is normally the case).

Performance Improvements and the Role of Training

The clear improvement in probe score performance over successive sessions was the result of a combination of factors, only one of which was the formal training administered to the participants. An example of performance improvement unrelated to training can be seen in the results shown in Panels D, E, and F of Figures 3 and 4. Those displays show that most of the improvement in the probe scores for words, sentences, and paragraph appeared to occur without any administration of formal training. Such improvement might be due to auditory feedback of the EMG-EL device during the probe sessions. After initiation of training, indirect learning could be attributed to participants' generalization of skills from formal training on the vowel production skills to performance with words, sentences, and paragraphs.

Spontaneous skill acquisition was observed in all participants. Even though training was administered for vowel initiation, duration, and termination first, 4 of the 7 participants managed to produce the words and paragraph successfully before meeting the reaction time criteria of the vowel production. Therefore, the training for vowel initiation and vowel duration seemed to generalize to performance in the production of words, sentences, and the paragraph. This finding also suggests that the EMG-EL reaction time tasks required a skill level that might have exceeded necessary skills for producing tractable conversational speech. As the results of Table 3 indicate, the control of suprathreshold EMG activity to modulate EMG-EL pitch required an amount of training that was somewhere between the amount required for reaction time tasks and the amount required for conversational speech. Table 3 also shows that training on vowel termination required the greatest training effort across all participants, likely stemming from the multiple factoring contributing to VTT, as discussed previously.

In comparison to other EMG-controlled prosthetic devices, EMG-EL training was similar in terms of the amount of training needed for practical use. For example, occupational therapists report that individuals fitted with the EMG-controlled Utah arm (Motion Control Inc.) typically need only about 30 min to start using the device successfully and become increasingly proficient over four to eight sessions of formal training. Similar findings are reported for other EMG-based upper-extremity prostheses (Agnew & Shannon, 1981; Benjuya & Kenney, 1990; Northmore-Ball, Heger, & Hunter, 1980). The present EMG-EL training results, therefore, correspond well with other EMG-prosthesis applications, in which users rapidly acquired basic control skills within the first hour of training and became more proficient with formal training that lasted up to 7 hr over a period of up to 10 days.

Although this investigation showed promising results for EMG control of an electrolarynx, it did not attempt to address all of the potential issues inherent in using the device for real-world communication. For example, although speech intelligibility seemed to be at least as good, or perhaps better, for participants speaking with the EMG-EL compared with a conventional electrolarynx (first author's observation), intelligibility was not formally examined in the present study. Moreover, the pragmatics of daily device setup, maintenance, and durability were not explored beyond a 45-day home trial with 1 participant. Other potential limitations of EMG-EL use include strap muscle fatigue during extended voice use and the overall cosmesis and acceptability of the device by persons with laryngectomy relative to the advantages that it offers. In addition, our participants were relatively skillful in their conventional EL use, so it remains unknown how patients across the normal range of alaryngeal voice use capabilities would perform with the EMG-EL. These aspects will be targets of future EMG-EL research.

Conclusion

A successive-stage, multiple-baseline training protocol showed that 4 normal-voice participants and 3 participants who underwent total laryngectomy learned to produce successfully words and sentences as well as read a paragraph using an EMG-controlled hands-free EL device. In addition, all participants with normal voice and 2 participants who had undergone laryngectomy learned to modulate the pitch of the EMG-EL voice to make intonational distinctions in their speech. These results show that the EMG-EL device is a viable option as an alaryngeal voice source that provides hands-free on/off and dynamic pitch control when neck strap muscles are preserved during laryngectomy. Results indicate that articulatory-coordinated voice onset, maintenance, offset, and semantically appropriate pitch modulation can be achieved after a few hours of simple feedback and instruction, and significant additional gains in device control can be realized over subsequent training sessions. Extended study of 1 participant indicated that these gains may be maintained for at least 6 weeks after training.

Acknowledgments

This research was supported by National Institute on Deafness and Other Communication Disorders Grant R01-DC006449 and U.S. Department of Veterans Affairs Rehabilitation Research and Development Grant V523P-6820, both awarded to Robert E. Hillman. We thank Kevin Kearns for his advice on single-participant experimental design.

References

Agnew, P. J., & Shannon, G. F. (1981). Training program for a myo-electrically controlled prosthesis with sensory feedback system. *American Journal of Occupational Therapy, 35,* 722–727.

Benjuya, N., & Kenney, S. B. (1990). Myoelectric hand orthosis. *Journal of Prosthetics and Orthotics, 2,* 149–153.

Choi, H. S., Park, Y. J., Lee, S. M., & Kim, K. M. (2001). Functional characteristics of a new electrolarynx "Evada" having a force sensing resistor sensor. *Journal of Voice, 15,* 592–599.

Erickson, D. (1993). Laryngeal muscle activity in connection with Thai tones. *Annual Bulletin Research Institute of Logopedics and Phoniatrics, 27,* 135–149.

Fairbanks, G. (1960). *The rainbow passage* (2nd ed.). New York: Harper and Row.

Gandour, J., & Weinberg, B. (1983). Perception of intonational contrasts in alaryngeal speech. *Journal of Speech and Hearing Research, 26,* 142–148.

Gandour, J., & Weinberg, B. (1984). Production of intonation and contrastive stress in electrolaryngeal speech. *Journal of Speech and Hearing Research, 27,* 605–612.

Goldstein, E. A. (1998). *An improved prosthetic electrolarynx controlled by pressure variations and electromyographic signals*. Unpublished S.B. honor's thesis, Harvard University.

Goldstein, E. A. (2003). *Prosthetic voice controlled by muscle electromyographic signals*. Unpublished doctoral dissertation, Harvard University.

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,* 325–332.

Gray, S., & Konrad, H. R. (1976). Laryngectomy: Postsurgical rehabilitation of communication. *Archives of Physical Medicine and Rehabilitation, 57,* 140–142.

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 Otolaryngology, Rhinology, and Laryngology, 113,* 754–764.

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. *Annals of Otology, Rhinology, and Laryngology, 172,* 1–27.

Hirano, M., Koike, Y., & von Leden, H. (1967). The sternohyoid muscle during phonation: Electromyographic studies. *Acta Oto-Laryngologica, 64,* 500–507.

Hirano, M., Ohala, J., & Vennard, W. (1969). The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. *Journal of Speech and Hearing Research, 12,* 616–628.

Hodge, M. M., & Rochet, A. P. (1989). Characteristics of speech breathing in young women. *Journal of Speech and Hearing Research, 32,* 466–480.

Hoit, J. D., & Hixon, T. J. (1987). Age and speech breathing. *Journal of Speech and Hearing Research, 30,* 351–366.

Hoit, J. D., Hixon, T. J., Altman, M. E., & Morgan, W. J. (1989). Speech breathing in women. *Journal of Speech and Hearing Research, 32,* 353–365.

Izdebski, K. (1980). Effects of prestimulus interval on phonation initiation reaction times. *Journal of Speech and Hearing Research, 23,* 485–489.

Jacobsen, S. C., Knutti, D. F., Johnson, R. T., & Sears, H. H. (1982). Development of the Utah artificial arm. *IEEE Transactions on Biomedical Engineering, 29,* 249–269.

Junker, A. (1995). Brain–body actuated system. *United States Patent No. 5474082*.

Kearns, K. P. (2000). Single-subject experimental designs in aphasia. In S. E. Nadeau, L. J. Rothi, & B. Grosson (Eds.), *Aphasia and language: Theory to practice* (pp. 421–441). New York: Guilford Press.

Keith, R. L., Shanks, J. C., & Doyle, P. C. (2005). Historical highlights: Laryngectomy Rehabilitation. In P. Doyle (Ed.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer* (pp. 17–58). Austin, TX: Pro-Ed.

Koike, Y., & Kawato, M. (1995). Estimation of dynamic joint torques and trajectory formation from surface electromyography signals using a neural network model. *Biological Cybernetics, 73,* 291–300.

Latwesen, A., & Patterson, P. E. (1994). Identification of lower arm motions using the EMG signals of shoulder muscles. *Medical Engineering and Physics, 16,* 113–121.

McRae, R. G., & Pillsbury, H. R. (1979). A modified intraoral electrolarynx. *Archives of Otolaryngology, 105,* 360–361.

Meltzner, G., Hillman, R. E., Heaton, J., Houston, K., Kobler, J., & Qi, Y. (2005). Electrolarynx speech: The state-of-the-art and future directions. In P. Doyle (Ed.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer* (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., & Robbins, K. T. (2002). Voice rehabilitation after total laryngectomy and postoperative radiation therapy. *Journal of Clinical Oncology, 20,* 2500–2505.

Min, H., Takahashi, M., Nishizawa, N., Nishizawa, S., Uemi, N., Ifukube, T., & Inuyama, Y. (1994a). Performance of an electrolarynx controlled by the sternohyoid muscle. *IPSJ SIGNotes Human Interface Abstract, No. 056*(005).

Min, H., Takahashi, M., Nishizawa, N., Nishizawa, S., Uemi, N., Ifukube, T., & Inuyama, Y. (1994b). Two pitch control methods of an electrolarynx controlled by the sternohyoid muscle. *IPSJ SIGNotes Human Interface Abstract, No. 056*(006).

Mitchell, H. L., Hoit, J. D., & Watson, P. J. (1996). Cognitive–linguistic demands and speech breathing. *Journal of Speech and Hearing Research, 39,* 93–104.

Morris, H. L., Smith, A. E., Van Demark, D. R., & Maves, M. D. (1992). Communication status following laryngectomy: The Iowa experience 1984–1987. *Annals of Otology, Rhinology & Laryngology, 101,* 503–510.

Northmore-Ball, M. D., Heger, H., & Hunter, G. A. (1980). The below-elbow myo-electric prosthesis: A comparison of the Otto Bock myo-electric prosthesis with the hook and functional hand. *Journal of Bone and Joint Surgery [Brit], 62B,* 363–367.

Roubeau, B., Chevrie-Muller, C., & Lacau Saint Guily, J. (1997). Electromyographic activity of strap and cricothyroid muscles in pitch change. *Acta Oto-Laryngologica, 117,* 459–464.

Saridis, G. N., & Gootee, T. P. (1982). EMG pattern analysis and classification for a prosthetic arm. *IEEE Transactions on Biomedical Engineering, 29,* 403–412.

Scargle, S. (1998). *EMG/EEG head–computer–interface system for computer cursor control*. Unpublished master's thesis, International University of Florida.

Solomon, N. P., & Hixon, T. J. (1993). Speech breathing in Parkinson's disease. *Journal of Speech and Hearing Research, 36,* 294–310.

Takahashi, H., Nakao, M., Okusa, T., Hatamura, Y., Kikuchi, Y., & Kaga, K. (2001a). Pitch control input by finger pressure using electrolarynx or intra-mouth vibrator. *The Japan Journal of Logopedics and Phoniatrics, 42,* 1–8.

Takahashi, H., Nakao, M., Okusa, T., Hatamura, Y., Kikuchi, Y., & Kaga, K. (2001b). Voice generation system using an intra-mouth vibrator. *Journal of Artificial Organs, 4,* 228–294.

Uemi, I., & Takahashi, M. (1994). Design of a new EL having a pitch control function. *IEEE International Workshop on Robot and Human Communications,* 378–383.

Watson, B. C. (1994). Foreperiod duration, range, and ordering effects on acoustic LRT in normal speakers. *Journal of Voice, 8,* 248–254.

Yamada, M., Niwa, N., & Uchiyama, A. (1983). Evaluation of a multifunctional hand prosthesis system using EMG controlled animation. *IEEE Transactions on Biomedical Engineering, 30,* 759–763.

Yorkston, K. M., & Beukelman, D. R. (1981). Communication efficiency of dysarthric speakers as measured by sentence intelligibility and speaking rate. *Journal of Speech and Hearing Disorders, 46,* 296–301.

Received July 19, 2005

Accepted June 27, 2006

DOI: 10.1044/1092-4388(2007/024)

Contact author: James T. Heaton, Center for Laryngeal Surgery and Voice Rehabilitation, Massachusetts General Hospital, One Bowdoin Square, 11th Floor, Boston, MA 02114. E-mail: jheaton@partners.org.

Appendix A. Instructions for initial probe experiments.

You are participating in a study investigating the use of a new device aimed at offering laryngectomy patients a better method of oral communication. In order to simulate the conditions of a laryngectomy participant, you are asked NOT to use your normal voice, but rather hold your breath and "mouth" the vowels, words, and sentences you are asked to read.

In order to fit you with the device, we have placed an electrode on your neck surface to record the activity from your muscles. This activity determines whether the buzzer you are wearing around your neck will come on or not, and it will change its 'tone' according to the level of that activity as well. We ask you to make sure the buzzer is in tight contact with your neck, so that the buzzing sound travels into your neck and out of your mouth as much as possible.

You will go through five short exercises:

1. Watch the screen in front of you and follow the commands: Relax, Ready, Say /a/, Stop, and Relax again. The background color of the screen will change according to the command. The sequence will repeat 10 times. Be ready and alert as much as possible in order to respond as quickly as you can.
2. Read/mouth a set of 10 words while holding your breath, attempt to read each word as it appears on the screen once only.
3. Read/mouth a set of 10 sentences while holding your breath, you should only attempt to read the sentence once, if you fail, do not try to repeat.
4. Read a paragraph with pauses set at specific points. Please read along until a pause is reached, then stop, take a breath, then hold your breath again and continue reading until the next pause is encountered.
5. Read/mouth five sentences in two ways: first as a statement, then as a question. Again, these sentences are to be read while holding your breath, taking a breath between sentences.

When reading/mouthing utterances, please be as clear as possible in your pronunciation and articulation, and move your mouth as if you are actually speaking with your normal voice. If the device happens to come on unexpectedly, or fails to come on when expected, just carry on with the task without interruption or repetition. Thank you.

Appendix B. EMG-EL training scoring instructions and criteria.

- You are asked to judge the performance of 5 participants in producing words, sentences, a paragraph, and questions, using an EMG-controlled EL.
- Material:
  - Words: 10 words.
  - Sentences: 10 sentences.
  - Paragraph: The first 10 sentences in the rainbow passage.
  - Questions: 10 statement/question pairs using the same words.
- Judging: you will be shown a video tape of each of the 5 participants going through each of the four exercises, you will also be given a printed sheet with the material on it. For each token, you are asked to judge in a binary fashion whether or not the participant was able to use the EMG-EL device correctly, based on the following criteria:
  - Words: The device was triggered at beginning of articulation and produced sound throughout the whole word. If the participant mouthed the word without triggering the device, then the token is a failure. If the device started correctly but did not stay on throughout the whole word, then it is also a failure. Due to the design of the device, the EMG-EL tends to stay on a bit more than desired, therefore the participant is not penalized for the device not turning off promptly at the end of a word.
  - Sentences: The same rules apply as to words, except that the participant may insert pauses at the appropriate places within the sentence (i.e., between words, or at a punctuation).
  - Paragraph: This criterion is identical to sentences, except that the participant is allowed to read multiple sentences without stopping the device.
  - Questions: In addition to reading the sentences with the correct triggering, the participant is expected to raise the pitch toward the end of a question, while ending with a falling pitch at the end of the statement. If the fundamental frequency rise at the end of the question can be distinguished from the falling pitch at the end of a statement, then the token is a success. Otherwise, it is a failure.
- Examples: You will have access to a demo tape that contains samples of 10 failing tokens and 10 passing tokens from each exercise at all times. You will be shown this demo tape before judging the randomized tokens. You can replay that demo tape at any time. You may also replay any of the tokens being judged as many times as you wish.

Thank you for your participation!!