Objectives: The goal of this study was to determine whether speech breathing changes over time in laryngectomy patients who use an electrolarynx, to explore the potential of using respiratory signals to control an artificial voice source.

Methods: Respiratory patterns during serial speech tasks (counting, days of the week) with an electrolarynx were prospectively studied by inductance plethysmography in 6 individuals across their first 1 to 2 years after total laryngectomy, as well as in an additional 8 individuals who had had a laryngectomy at least 1 year earlier.

Results: In contrast to normal speech that is only produced during exhalation, all individuals were found to engage in inhalation during speech production, and those studied longitudinally displayed increased occurrences of inhalation during speech production with time after laryngectomy. These trends appear to be stronger for individuals who used an electrolarynx as their primary means of oral communication rather than tracheoesophageal speech, possibly because of continued dependence on respiratory support for the production of tracheoesophageal speech.

Conclusions: Our results indicate that there are post-laryngectomy changes in the speech breathing behaviors of electrolarynx users. This has implications for designing improved electrolarynx communication systems, which could use signals derived from respiratory function as one of many potential physiologically based sources for more natural control of electrolarynx speech.

Key Words: alaryngeal speech, breathing, electrolarynx, inductance plethysmography, tracheoesophageal speech.

INTRODUCTION

A majority of laryngectomy patients in the United States use an electrolarynx (EL) to communicate, 1-4 but EL devices are cumbersome to use (ie, requiring the use of one hand) and EL speech has reduced intelligibility and a poor or unnatural quality. 5 For several years our group has been investigating multiple approaches for improving EL speech communication, 2,5-11 including the use of signals that are related to speech physiology as a basis for providing more natural control of an EL device. 6-9 One potential source of such control signals is the respiratory system, because of well-defined relationships between respiratory kinematics and speech production in normal speakers (eg, Hixon12). Thus, the purpose of this study was to determine the potential for using respiratory-related signals to control an EL in individuals after laryngectomy. This was done by examining respiratory function during EL speech in laryngectomy patients who use either an EL or an aerodynamically driven form of alaryngeal speech (esophageal or tracheoesophageal [TE]) as their primary means of communication.

In individuals with normal voice and speech, respiration is coordinated with the intent to speak. Although the kinematics of speech breathing have been studied extensively for different speech tasks and in differing groups, 12-15 the most obvious requirement of speech respiration in most populations is generating an adequate exhalation to create sound at points of articulatory constriction and drive the vocal folds into motion, thereby valving the air stream for the production of voice. After laryngectomy, the airway is separated from the pharynx and an alaryngeal voice source is needed to speak. In EL users, voice and speech production are no longer dependent on the respiratory system. Conversely, individuals who use TE or, to a lesser degree, esophageal speech still maintain some physical relationship between the respiratory system and the production of voice and speech.

Given the drastic changes in aerodigestive tract anatomy experienced by laryngectomy patients, speech breathing patterns may be expected to change over time, particularly for individuals who exclusively use an EL to speak. However, Liu et al 16 reported that although individuals who had had a laryngectomy 2 to 3 months earlier (N = 8) were...
PARTICIPANT DEMOGRAPHIC INFORMATION AND RECORDING SESSIONS

<table>
<thead>
<tr>
<th>Participant</th>
<th>Research Group</th>
<th>Age at Last Recording (y)</th>
<th>Sex</th>
<th>Sessions Since Surgery (Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EL1</td>
<td>Longitudinal</td>
<td>64</td>
<td>M</td>
<td>4, 6, 8, 12</td>
</tr>
<tr>
<td>EL2</td>
<td>Longitudinal</td>
<td>41</td>
<td>M</td>
<td>2, 4, 6, 8, 10, 12</td>
</tr>
<tr>
<td>EL3</td>
<td>Longitudinal</td>
<td>71</td>
<td>M</td>
<td>2, 4, 6, 8</td>
</tr>
<tr>
<td>TE1</td>
<td>Longitudinal</td>
<td>62</td>
<td>M</td>
<td>4, 6, 8, 12</td>
</tr>
<tr>
<td>TE2</td>
<td>Longitudinal</td>
<td>61</td>
<td>M</td>
<td>1, 5, 3, 8, 31</td>
</tr>
<tr>
<td>ES1</td>
<td>Longitudinal</td>
<td>71</td>
<td>M</td>
<td>5, 11, 15</td>
</tr>
<tr>
<td>EL4</td>
<td>Single time point</td>
<td>80</td>
<td>M</td>
<td>204</td>
</tr>
<tr>
<td>EL5</td>
<td>Single time point</td>
<td>60</td>
<td>F</td>
<td>156</td>
</tr>
<tr>
<td>EL6</td>
<td>Single time point</td>
<td>62</td>
<td>M</td>
<td>15</td>
</tr>
<tr>
<td>EL7</td>
<td>Single time point</td>
<td>60</td>
<td>M</td>
<td>30</td>
</tr>
<tr>
<td>EL8</td>
<td>Single time point</td>
<td>61</td>
<td>M</td>
<td>26</td>
</tr>
<tr>
<td>EL9</td>
<td>Single time point</td>
<td>54</td>
<td>M</td>
<td>18</td>
</tr>
<tr>
<td>TE3</td>
<td>Single time point</td>
<td>48</td>
<td>F</td>
<td>14</td>
</tr>
<tr>
<td>TE4</td>
<td>Single time point</td>
<td>64</td>
<td>M</td>
<td>53</td>
</tr>
</tbody>
</table>

EL — electrolarynx; TE — tracheoesophageal; ES — esophageal.

still likely to exhale during speech, individuals who had had a laryngectomy 5 to 11 years earlier (N = 4) were likely to hold their breath during speech. Surprisingly, they reported that it was very rare for their participants to inhale during speech tasks, regardless of the time since surgery, and even when asked to do so. This finding conflicted with our previous informal observations that breathing patterns are often independent of EL speech.

Speech breathing patterns in laryngectomy patients have implications for the development of EL technology based on respiratory control, motivating the present investigation. If the respiratory kinematics after laryngectomy prove to be adaptable, the design of EL technology utilizing respiratory signals is less limited with respect to control schemes; however, if breathing patterns are stable after laryngectomy, the control scheme design will be dependent upon these specific respiratory kinematics. The breathing patterns of 6 laryngectomy patients across their first 1 to 2 years after total laryngectomy and those of 8 individuals at least 1 year past total laryngectomy were recorded by respiratory inductance plethysmography during serial EL speech tasks (ie, counting from 1 to 10 and saying the days of the week) to quantify their speech breathing patterns in relation to time since laryngectomy.

MATERIALS AND METHODS

The participants were 14 laryngectomy patients (2 female, 12 male) between the ages of 41 and 80 years at the time of their last (or only) recording session. From April 2000 to April 2006, a prospective assessment of speech breathing was conducted on 6 individuals who underwent total laryngectomy (longitudinal group). Based on the trends seen in these original data, from April 2007 to September 2007, 8 individuals at least 1 year past total laryngectomy were also recorded (single time-point group). These study protocols were approved by the human studies institutional review boards of the Massachusetts General Hospital and the Massachusetts Eye and Ear Infirmary.

The demographic information, recording sessions, and primary mode of communication are detailed across participants in the Table. The 6 individuals in the longitudinal group were recorded across their first year after total laryngectomy with time periods that varied by individual, but ranged from 1 month after laryngectomy to 31 months after laryngectomy. All of these individuals were recorded at 3 or more different time points, with an average sampling of 4.5 visits, and all were recorded at least as early as 5 months after surgery. Two of these 6 individuals became proficient TE speakers during the recording period (TE1 by month 8 and TE2 by month 5). The individual ES1 was a proficient user of esophageal speech by month 11, but communicated with a combination of esophageal and EL speech. Individuals in the single time-point group were recorded at least 1 year after laryngectomy. Of these individuals, 2 were proficient users of TE speech and had used TE speech as their primary mode of communication for at least 1 year (while maintaining an EL for backup communication). The remaining 6 individuals used EL speech as their primary mode of communication. All 14 individuals reported themselves as nonsmokers during the time of the experiment, without a history of language, hearing, or speech disorders (before laryngectomy), and were proficient users of EL speech, even if it was not their primary mode of communication.

Simultaneous acoustic signals and thorax and abdomen movement signals were filtered and recorded digitally (20,000 Hz sampling rate) with Axon Instruments (Foster City, California) hardware (Cyberamp 380, Digidata 1200) and software (Axoscope). The acoustic signal was obtained from a lavalier microphone (ECM50PSW, Sony, New York,
New York) or a headset microphone (C 420 PP, AKG Acoustics, Vienna, Austria). Thoracic and abdominal wall movement was recorded with 2 flexible inductance bands that fit around the chest and abdomen (Respitrace Systems, Ambulatory Monitoring, Inc, Ardsley, New York). During setup, kinematic data were checked to ensure that the full breathing range was within the amplitude range of the recorded signals and that increases in signal strength corresponded to chest and abdominal band expansion. Attempts were made in 3 individuals to use a fixed-volume calibration bag (910-0185-000, VivoMetrics, Ventura, California) to calibrate the respiratory signals, but air leakage around the calibration bag’s tube placement against the tracheostoma prevented consistent measurement of a fixed volume exchange in all but 1 individual. These calibration attempts did, however, validate the assumed relationship between respiratory signal deflection and direction of airflow from the stoma (inhalation or exhalation). Isometric respiratory calibration maneuvers were not performed because of the difficulties associated with achieving airway closure in neck-breathing individuals (ie, air leakage around stomal occlusion and air passage through TE valves when present). Therefore, kinematic data were used to observe the direction of chest wall motion only, with the goal of identifying regions of inspiration during speech, as discussed below.

All participants were able to complete serial speech tasks using an EL without difficulty, regardless of their primary alaryngeal voice source. These speech tasks consisted of saying the days of the week and counting to 10. The participants were instructed to pause (turn off the EL) after each word. Although the instruction to pause between words elicits speech behavior that is perhaps less natural than running speech, it facilitated accurate analysis of respiration patterns for vocal production versus pauses that are otherwise often obscured by continuous EL activation on the part of the user. However, in order to compare the behavior quantified during serial tasks with that in the more natural context of running speech, we also recorded the individuals’ reading of the first paragraph of the Rainbow Passage.

The extent to which inspiration occurred during speech was calculated as a percent of total EL voiced speech time for each individual at each recording session. The abdominal and thoracic kinematic data were analyzed with respect to slope, with positive slopes indicating inspiratory movements. Because of the possible occurrence of paradoxical movement during breathing, inspiration was marked at those times when both the abdominal and thoracic bands displayed positive slopes. This method likely resulted in an underestimation of inspiration time, but ensured that behavior identified as inspiration was unequivocally inspiration.

Speech time during the serial speech tasks was determined by visual and acoustic identification of EL activation in the microphone signal by means of custom software written in Matlab (Mathworks Inc, Natick, Massachusetts) without simultaneous observation of respiration signals. Speech time during the Rainbow Passage was determined by adjusting an audio amplitude threshold of the recorded passage, which removed low-amplitude (subthreshold) sounds while maintaining suprathreshold sounds. The threshold was adjusted for each recording (by the first author) until the resulting speech was intelligible but lacked the low-amplitude EL buzzing associated with a closed mouth. This approach erred toward removing some low-amplitude speech components, but it ensured that nonspeech periods were not included in the speech time.

To screen for obvious evidence of compromised respiratory function, we made recordings of all subjects “at rest” to obtain their tidal breathing rates. The average breaths per minute (BPM) were calculated for each subject by visual selection of isolated inspirations during tidal breathing.

RESULTS

All participants deviated from natural speech breathing patterns, although their speech breathing patterns varied qualitatively. Some individuals displayed breathing that was obviously modulated by the intent to speak, but was some degree out of phase with natural behavior. For example, they were noted to begin speaking before their peak in inspiration, during what would otherwise look like a normal pre-phonatory inspiration. In other individuals, speech breathing patterns approximated those of tidal breathing or were extremely atypical, including relatively long periods of breath holding interspersed with randomly placed large inspiratory and expiratory excursions. Examining the speech breathing behavior more quantitatively, Fig 1 shows the percent of EL speech time during which all 14 participants inhaled during serial speech tasks as a function of their time since laryngectomy. All participants spent some portion of their speech time inhaling, with individual values ranging from 8% to 64%. Analyses of respiration during reading of the Rainbow Passage confirmed the presence of inhalation during running speech, with individual values ranging from 10% to 44%. The remaining results presented
below pertain exclusively to the serial speech tasks (isolated words) because of the higher confidence in differentiating speech versus nonspeech time.

**Longitudinal Group Results.** For the 6 individuals of the longitudinal group, an average change in the percent of EL speech time during which the participant inhaled was calculated for the final recording session (8 to 31 months) versus the point closest to 4 months after laryngectomy. Although there were data available for some individuals earlier than 4 months, the 4-month time point was chosen because it was the most frequently occurring sample point in the early post-surgical time period (ie, less than 6 months). All of these individuals were found to engage in inspiration during EL speech production to an increasing degree as their time since laryngectomy progressed. The average increase in the percent of EL speech time spent inhaling across all participants was 19%, which was statistically significant ($p < .01$, 1-tailed $t$-test). Figure 2 illustrates the difference seen in the raw data for an EL participant (EL3) at 4 months and 12 months after laryngectomy. When the longitudinal group participants are grouped by their primary alaryngeal communication mode of EL (3 subjects) or TE or esophageal (3 subjects) speech, the data indicate that the individuals who primarily used TE or esophageal speech may have maintained more natural speech breathing compared to their EL-using counterparts when all participants used an EL. The average increase in the percent of EL speech time spent inhaling for the 3 TE or esophageal speech participants was 12% (SD, 8%); whereas the 3 EL users had more than twice that degree of change, with an average increase of 27% (SD, 15%) over time. However, a statistical comparison between these average subgroup changes over time (3 subjects per subgroup) failed to find a significant difference ($p = .11$, 1-tailed $t$-test).

**Longitudinal and Single Time-Point Group Results Combined.** The last post-surgical data point collected for all 14 participants is shown in Fig 3. The average percent of speech time spent inhaling...
for all participants was 28% (SD, 17%), which is statistically greater than zero (p < .001, 1-tailed t-test). These data also support the trend that TE and esophageal speakers display less inhalation during speech than EL speakers. The TE and esophageal speakers spent an average of 19% (SD, 9%; N = 4) of speech time inhaling, whereas EL speakers spent an average of 33% (SD, 19%; N = 10) of speech time inhaling (with all subjects using an EL), producing a statistically significant difference (p = .04, 1-tailed t-test). For the last post-surgical data point collected for all participants, the BPM was compared to the percent of speech time spent inhaling. However, no correlation was found between these measures (R² = 0.03). Further, no difference in BPM was found between EL and TE or esophageal groups (p = .42, 1-tailed t-test), with EL users displaying an average tidal breathing rate of 18.5 BPM (SD, 6.2 BPM) and TE users 19.3 BPM (SD, 7.4 BPM).

**DISCUSSION**

This study provides insight into changes in speech breathing associated with post-laryngectomy use of an electrolarynx. In contrast to normal speech that is only produced during exhalation, the results indicate that the relationship of respiratory phase (ie, inhalation versus exhalation) to volitional EL speech becomes variable after laryngectomy. Moreover, although all participants performed speech tasks with an EL in this study, the amount of time they spent inhaling during speech depended somewhat on their typical (primary) mode of alaryngeal communication.

The results of the present study differ considerably from those of Liu et al., who did not observe inspiration during the speech of the laryngectomy patients in their study. Even with our conservative estimate of the time spent inhaling during speech (simultaneous positive slopes on both abdominal and thoracic bands), it was obvious that our participants spent a sizeable percentage of their time speaking on inhalation. However, the finding by Liu et al that 2 to 3 months after laryngectomy individuals found it difficult to speak with an EL while independently inhaling and exhaling is not contradicted by our data. Our longitudinal data display a lower incidence of inhalation during speech at 2 or 3 months after surgery than at later time points, perhaps reflecting that speech breathing patterns learned over a lifetime are still active and that individuals may resist performing “unnatural” speech breathing tasks. However, we do not have an explanation for the differences between our data and the findings of Liu et al in their 4 individuals who had had laryngectomy 5 to 11 years earlier. Although their participants (full-time EL users) were not seen to inhale during natural speech and even found it impossible to learn to inhale and exhale independently while speaking with an EL, our EL participants did so frequently by their final data point (8 to 204 months; see Table) without having received any instruction regarding their breathing patterns.

The present findings indicate not only that speech breathing changes after laryngectomy, but that it does so to a degree that generally relates to speech breathing requirements in the post-surgical period. The laryngectomy patients who relied exclusively on an EL as their alaryngeal voice source experienced a total loss of physical coupling between alaryngeal voice production and exhalation, and also spent the greatest percentage of speech time inhaling, especially at 8 months or more after laryngectomy. In contrast, the laryngectomy patients who initially used an EL but came to produce TE (4 subjects) or esophageal (1 subject) speech needed to coordinate breathing with their primary means of ala-
ryngeal speech, and consequently either maintained or relearned a more normal speech breathing pattern by their last recording session. This was demonstrated by the fact that although the use of an EL during the speech tasks in this study permitted indiscriminant breathing patterns, the TE or esophageal speakers (on average) demonstrated less inhalation during EL speech than did subjects who typically spoke with an EL (p = .04, 1-tailed t-test). This finding suggests that an artificial voice control strategy that depends on normal speech breathing patterns may work best if initiated early in the post-surgical period, or that it may require a period of training to extinguish instances of inhalation during speech production.

One possible explanation for altered breathing patterns in EL users versus TE or esophageal speakers is inferior pulmonary function in the EL user population. Inadequate pulmonary function is one of several reasons why laryngectomy patients are sometimes unable to adopt TE speech,18-20 potentially increasing the likelihood that our EL group had more individuals with compromised pulmonary function that could lead to their increased inhalation during speech. However, no correlation was found between these measures, and no difference in BPM was found between EL and TE or esophageal groups. Further, the average BPM found in both of these groups is less than 1 SD of the average tidal breathing rate of a group of 50-year-old healthy speakers (13.7 BPM; SD, 4.7; N = 30), indicating a lack of obvious evidence that poor respiratory health (ie, a need to inhale more frequently) had a role in the tendency of some of the subjects to inhale during speech production.21

The increased appearance of inhalation during EL speech production over time after laryngectomy in a communication mode–dependent manner is not surprising given the flexibility that anatomic-ally normal individuals have in speech-related respiration. As an extreme example, some individuals learn to habitually speak on inhalation (ie, “reverse phonation”) as a rehabilitation technique for laryngeal focal dystonia.22 This ability is consistent with the known importance of auditory targets for speech production23-25 and changes in speech breathing that result from abnormal hearing26 suggest that the targets for speech breathing are auditory as well. Thus, it is not unexpected that the strict coordination between exhalation and speech is abandoned once exhalation is not necessary to produce the required auditory target.

Evidence of the primacy of auditory feedback in speech breathing behavior supports the potential for developing EL control systems that are respiration-based. Perhaps if the respiration is once more linked with audible speech production, the inherent flexibility of the respiratory system in supporting speech will facilitate the adoption of breathing patterns advantageous for more naturally controlling an artificial voice source. However, prior research by Uemi et al27 indicates that the dissociation between breathing and speaking may be resistant to change. They studied the ability of 16 laryngectomy patients with various primary alaryngeal voice sources to control pitch and timing of an EL with respiration. They found that after 1 week of training (30 minutes a day), only 6 of the 14 participants were able to control the on-off timing of the device within 500 ms, and only 3 of the 14 participants were able to control the pitch of the device in order to match a given 2-pitch stimulation signal. Although the time since laryngectomy of these participants was not reported, there were 6 individuals studied who used TE speech as their primary alaryngeal voice source. The performance of these individuals was somewhat better than that of the entire cohort studied, with 4 of these individuals able to control the on-off timing within 500 ms, and 2 able to control the pitch of the device. Thus, the present results indicate that speech breathing patterns after laryngectomy relate to speech breathing demands (ie, EL versus TE speakers speech breathing patterns), suggesting that further studies regarding specific respiratory-driven control schemes are needed to fully explore the possibilities of a respiration-controlled EL.

CONCLUSIONS

Our results suggest that speech breathing associated with EL use after laryngectomy changes over time and that perhaps new speech breathing patterns could be learned. Furthermore, the observation that breathing patterns differed during EL speech according to a person’s primary means of alaryngeal speech, and presumably their typical speech breathing demands, has implications for the design of, and training needs for, physiologically controlled EL technology, which could use signals derived from respiration as a basis for more naturally controlling EL speech.

REFERENCES


