The electrolarynx: voice restoration after total laryngectomy

Rachel Kaye¹
Christopher G Tang²
Catherine F Sinclair³

¹Department of Otolaryngology, New York Center for Voice and Swallowing Disorders, New York, NY, ²Department of Otolaryngology, Kaiser Permanente San Francisco Medical Center, San Francisco, CA, ³Department of Otolaryngology, Mount Sinai Icahn School of Medicine, New York, NY, USA

Abstract: The ability to speak and communicate with one’s voice is a unique human characteristic and is fundamental to many activities of daily living, such as talking on the phone and speaking to loved ones. When the larynx is removed during a total laryngectomy (TL), loss of voice can lead to a devastating decrease in a patient’s quality of life, and precipitate significant frustration over their inability to communicate with others effectively. Over the past 50 years there have been many advances in techniques of voice restoration after TL. Currently, there are three main methods of voice restoration: the electrolarynx, esophageal speech, and tracheoesophageal speech through a tracheoesophageal puncture (TEP) with voice prosthesis. Although TEP voice is the current gold standard for vocal rehabilitation, a significant minority of patients cannot use or obtain TEP speech for various reasons. As such, the electrolarynx is a viable and useful alternative for these patients. This article will focus on voice restoration using an electrolarynx with the following objectives: 1) To provide an understanding of the importance of voice restoration after total laryngectomy. 2) To discuss how the electrolarynx may be used to restore voice following total laryngectomy. 3) To outline some of the current electrolarynx devices available, including their mechanism of action and limitations. 4) To compare pros and cons of electrolaryngeal speech to TEP and esophageal speech.

Keywords: total laryngectomy, voice restoration, electrolarynx, esophageal speech, tracheoesophageal puncture, silent speech, electrolarynx mechanics

History of voice restoration
Voice restoration in the absence of a functional larynx has been reported for over 150 years.¹ The first known description of a laryngeal prosthesis was by Czermak in 1859 who reported voice production after diverting airflow from a tracheal cannula through a reed-containing tube to the mouth in a patient with complete laryngeal stenosis.² Billroth performed his first laryngectomy in 1873 and his assistant, Gussenbauer, outfitted the patient with a custom speech apparatus; a reed-like pneumatic device was mounted onto a double lumen tracheostomy tube with a port extending into the pharyngostome.¹,³ The successful result was reported at the Third Congress of the German Company of Surgeons in Berlin in 1874.⁴ Advancements in voice restoration quickly expanded as reports of esophageal speech were published in the mid-19th century,⁵,⁶ mechanical vibration at the turn of the 19th century,⁷ fistula creation that directed air to the upper esophagus and pharynx in the mid-20th century, and finally unidirectional prosthetic valves that cannulated the trachea and upper esophagus called tracheoesophageal puncture (TEP) speech in the mid-20th century.¹
Overview of voice restoration

The definition of human voice is sound produced by means of the lungs and larynx or the faculty of utterance. Voice production with an anatomically normal vocal tract requires three distinct elements (Figure 1A):

1. Air generator: the lungs generate an air stream that flows through the larynx.
2. Vibrating apparatus: apposition of the paired vocal folds when supplied by the air stream generates undulating vibrations to create speech sound.
3. Articulating tract: sound is modulated by the pharynx and oral cavity to produce phonetic voice.

Patients who have undergone total laryngectomy (TL) have had their vibrating apparatus removed and although the air generator and articulating tract remain, the air stream is diverted and does not pass through the articulating tract. It is, by this mechanism, that they lose the ability to produce sound (Figure 1B). Furthermore, the articulating tract may be modified as well during the surgical excision depending on the extent of concomitant pharyngeal or tongue base disease involvement. Voice restoration aims to artificially create a sound source by reintroducing a vibrating air column that is then modified by the articulating apparatus.

The physiology of the three principal options in alaryngeal voice differs mainly in the approach to reconstruction of the vibrating apparatus and the air generator as the articulating tract is similar for all three approaches. For electrolaryngeal voice production, the vibrating apparatus is an external electromechanical device that produces either pharyngeal or oral cavity vibrations. The air generator is not used to induce vibration and thus the mucosal vibrations differ from those produced by waves of air stream (as in tracheoesophageal and esophageal methods). In fact, with electrolaryngeal voicing there is absence of airflow through the mouth during phonation and aerodynamic studies have suggested a decoupling or decreased coupling of the respiratory system to speech demands in electrolaryngeal users.

Novice electrolarynx users were shown to continue to mimic the presurgical speech mechanics by exhaling during speech production. However, veteran electrolarynx users decouple these functions and instead produce speech while breath holding, thereby garnering the advantage of improved speech acceptability. This is important as there is no physiologic need to maintain the behavior of exhalation during speech production as there is no pulmonary requirement needed to produce speech. In contrast, both TEP and esophageal methods still rely upon the pulmonary air generator; however, they have different modifications of its use. The TEP utilizes a one-way tracheostoma valve to shunt pulmonary air into the esophagus while esophageal speech utilizes air captured in the mouth and pharynx that is delivered to and stored in the esophagus. Both rely on subsequent air egress from the esophagus coupled with the vibrating apparatus of

Figure 1  Voice production.

Notes: (A) Sound production for the intact vocal tract is dependent upon the presence of three elements: air generator, vibrating apparatus, and articulating tract; (B) total laryngectomy removes the vibrating apparatus and redirects the air stream, significantly impacting vocal abilities. Copyright©2015. Elsevier. Adapted from Tang CG, Sinclair CF. Voice restoration after total laryngectomy. Otolaryngol Clin North Am. 2015;48(4):687–702.
the neoglottis (pharynx). The produced mucosal wave is in direct response to an air stream and thus not dissimilar to the presurgical laryngeal mucosal wave, thereby highlighting a difference in voicing mechanics between these methods and electrolaryngeal voicing.

Perceived quality of life is dramatically impacted by TL with many reports showing a dramatic decrease in psychosocial quality of life. Furthermore, voice deprivation limits social relationships and can induce social withdrawal. Reduced sexual enjoyment and libido are also common after laryngeal and hypopharyngeal cancer surgery. Thus, a primary focus of rehabilitation following TL is rapid, effective voice restoration as this is pivotal in preventing deleterious psychosocial and economic consequences. Voice restoration is certainly achievable for some; there are reports of equivalent post-treatment quality of life values when comparing patients who underwent TL with successful vocal rehabilitation to those who underwent laryngeal salvage with (chemotherapy) radiation techniques.

We recommend that all patients who plan to undergo TL should have a preoperative speech therapy assessment. Speech language pathologists have a crucial role in voice rehabilitation as they direct patients through the learning curve for rehabilitation techniques. Patients need to learn how to use an electrolarynx, perform esophageal speech, or utilize their TEP. There are financial considerations as well; different vocal rehabilitation methods have variable costs, especially in the third world. Longitudinal vocal dysfunction should be documented both pre- and postoperatively using voice rating scales. Popular scales include the Voice Handicap Index 10 and the University of Washington Quality of Life scale.

By and large, the ability to communicate with others is indispensable to one’s quality of life and activities of daily living and voice restoration after TL can restore this. Although there are few studies looking solely at quality of life measures with electrolaryngeal speech, both ongoing microtechnology developments and upgraded user interfaces have enriched electrolaryngeal speech with enhanced sound quality, leading to improved quality of life among electrolarynx users.

The electrolarynx

The electrolarynx was the initial method devised for the goal of restoring voice after TL. It has obvious advantages over synthesized speech and text to speech options as it allows continuance of certain natural nuances of speech including utilization of a patient's oral cavity and preserved articulatory abilities. As such, there are several types of electrolarynx devices, both historical and current. An overview of some of the more widely popularized ones is given below.

History and evolution

Prior to the early 20th century, experimental nonelectric vibrating pneumatic devices were available in select academic settings. However, due to their impracticality and poor usability, these devices were not widely utilized until the late-1920s, when the first recognized electrolarynx was developed. The Western Electric 5A (Number 5 Type Electronic Artificial Larynx, Western Electric, New York City, NY, USA) was the fifth installation and consisted of an electrically powered device with an attached vibrating diaphragm that was placed on the postlaryngectomy patient's neck. The design of the diaphragm portion of this device subsequently became the prototype for many of the currently available electrolarynges, albeit with significant modifications to overall device size and portability.

During World War II, the electric revolution came underway and a more advanced electrolarynx was developed by the Aurex Corporation in Chicago, IL, USA. The Aurex Neovox was placed on the postlaryngectomy patient’s neck. 20,21 The design of the diaphragm portion of this device subsequently became the prototype for many of the currently available electrolarynges, albeit with significant modifications to overall device size and portability.

Mechanism of action

The electrolarynx is an external device that induces vibration of oral or pharyngeal mucosa, generally, at a constant fundamental frequency (Figure 2). Notably, since its inception over 100 years ago, very little development has occurred with regards to improving the sound quality and fundamental frequency contour. The electrolarynx can function either indirectly by contacting the skin, which induces pharyngeal vibrations or directly through intraoral contact, which induces oral cavity vibrations. Muscles of articulation (e.g., tongue...
and lips) are, generally, intact after TL and thus amenable to shaping the supplied vibration noise into understandable speech.

The intraoral or direct technique utilizes a small vibrating tube that is introduced through the mouth and is laid against buccal mucosa, pharyngeal mucosa, or base of tongue. The placement site depends on the technique and preference of the patient. By transmitting sound directly into the mouth there is less leakage of energy than with the transcervical devices, resulting in louder sound and better speech quality. The intraoral tube is, particularly, useful for people whose neck anatomy precludes adequate sound transmission through neck tissues or those who cannot tolerate manual pressure on the neck (as would be required with the transcervical device). However, the intraoral device can be inconvenient for some patients and the tube can interfere with pronunciation of sounds made by putting the tongue to the roof of the mouth (e.g., “d”, “t”, “k”, “g”), leading to decreased speech intelligibility. In addition, the intraoral tubes can have sanitary issues associated with them. To combat these design weaknesses, a new device has been developed that allows for hands free vocalization by utilizing an electromyographic (EMG) transducer attached to the strap muscles in order to activate a vibratory source. However, this device is not in widespread use and often the strap musculature is removed during surgical excision, thereby precluding use of this device.

**Acoustic characteristics of electrolaryngeal speech**

As mentioned earlier, the electrolarynx produces voice by emitting constant vibrations that are transmitted to the pharynx through cervical skin or directly to the intraoral mucosa. Monotonic or robotic speech quality is produced when certain acoustic deficits are present including a flattened fundamental frequency (Fo), radiating noise, and an improper source spectrum. Radiating noise represents the mechanical hum from the electrolarynx that is not filtered by the vocal tract and instead is perceived directly by the listener. Various noise-reduction techniques and filtering systems have been invented that serve to lessen this radiating noise significantly and thus currently, monotonous Fo remains the main constraint to greater speech intelligibility of the electrolarynx. Intonation (Fo variation over time during phonation) results in improved speech intelligibility and is thought to be a greater contributor to intelligibility than the exact value of the Fo. As such, numerous Fo adjustment methods have been proposed that typically modulate the frequency either by electrophysiological signals or by pressure signals generated by force-sensing resistors or air pressure sensors. Additional prototypes have been created with manual Fo adjustments with finger movements. This research is vital due to the significant difficulties utilizing an electrolarynx in tonal languages (i.e., Mandarin) where tone variation is a result of Fo variation and imparts important
meaning to monosyllables. In, predominantly, nontonal Western languages, Fo holds less importance, although intonation does convey emotional quality and can provide various linguistic cues and aid the listener in identifying the gender and age of the speaker. Furthermore, some believe that the lack of Fo variation, and not the exact value of the Fo, contributes the most to decreased speech intelligibility. As technological advances continue to improve the automation ofFo variation, we can expect speech intelligibility to improve.

Commercially available electrolarynges

The four most widely used commercially available electrolarynges are discussed below. Of note, there is considerable price variation among the different models (almost a sixfold discrepancy) that can impact decision making for some, especially in developing countries.

The basic electrolarynx is the Nu-Voice (Mountain Precision Mfg. Ltd. Co.; Boise, ID, USA). It has two buttons – one that controls intensity and another that controls pitch settings. It is applied transcervically and has an oral adaptor tube. Additional information can be found in http://www.nu-vois.com.

The Servox (Servox Medizintechnik, Troisdorf, Germany) electrolarynx is very similar to the Nu-Voice. The standard model comes with two buttons, both programmable for two settings – one that controls loudness and another that controls intonation. The electrolarynx is applied transcervically and comes with an oral adaptor tube to be used transorally. Additional information on the product can be obtained at http://www.servona.de/.

The third commercially available electrolarynx is the TruTone electrolarynx (Griffin Laboratories, Temecula, CA, USA). Similar to the servox and the Nu-Voice, it is a transcervical handheld device with an oral adaptor. However, the TruTone allows for pitch control with direct finger pressure with pitch controlled by increasing or decreasing pressure placed on a single button – the greater the pressure, the higher the pitch. Additional information on the TruTone electrolarynx can be found at http://www.griffinlab.com.

The Ultravoice electrolarynx (The Ultravoice Plus Inc.) is different from the previous three in that it does not have a transcervical component. This electrolarynx consists of a denture mounted vibratory device with a handheld transmitter allowing for transoral transduction. This allows activation of the device without having to physically lift the vibratory source to place it on the neck. After the denture is placed in the mouth, the patient has control of the device with a wireless miniature fingertip switch on a controller. Additional information can be obtained at http://www.ultravoicelm.com.

Electrolarynx compared to other forms of postlaryngectomy vocal rehabilitation

As previously mentioned, there are three methods of TL vocal rehabilitation – the electrolarynx, esophageal voice, and TEP. The electrolarynx has certain advantages and disadvantages over these other methods as discussed below.

Esophageal speech is achieved through a process of esophageal insufflation with swallowed air. The air is released from the esophagus in a controlled fashion, thereby allowing for vibration of the mucosa of the upper esophagus/neopharynx. The air column is then funneled through the residual articulatory apparatus, where it is adapted to form intelligible voice. Limited air reserve within the esophagus may restrict esophageal voice quality and limit the duration of speech prior to another ingestion of air, both of which result in a shorter “staccato” phonation time, voicing at low pitch and volume, and greater labor intensiveness for the patient. Moreover, esophageal speech is more difficult to learn than electrolaryngeal speech and anatomical patient factors such as tight cricopharyngeal musculature and esophageal or pharyngeal stenosis can diminish airflow strength and thus make esophageal speech difficult. Furthermore, this technique is resource intensive for the speech/language therapist when educating new users. Patients may find esophageal voicing to be labor-intensive with limited air reserve. Advantages of esophageal voicing over the electrolarynx include its low cost as there is no need to purchase an electrolarynx. In fact, it is used most commonly in developing countries as it is relatively inexpensive.

TEP with voice prosthesis insertion allows patients to direct tracheal air through a puncture site in the posterior tracheal wall and drive it into the esophagus. This air column then is released from the esophagus and enters the pharynx and oral cavity where it is transformed into speech by articulatory mechanisms. As first described by Singer and Blom in 1980, the TEP allows for excellent speech quality and sound. Patients need to have favorable anatomy for the placement of a TEP as stomas that are situated too inferior or have too small of an aperture result in difficulties with proper prosthesis placement and use. The prosthesis may be placed at the time of TL or as a secondary procedure and one advantage of the electrolarynx over TEP speech is the avoidance of additional surgical procedures. Similar to the
electrolarynx where poor manual dexterity can make using a handheld device difficult, patients with musculoskeletal disorders, amputations, history of cerebrovascular accidents, or other impairments of manual dexterity may have difficulty manipulating the stoma to use a TEP prosthesis. Furthermore, patients without ready access to a speech-language pathologist will have difficulty learning to use and care for the TEP prosthesis. TEP valves need to be replaced at various time points due to obstruction, leakage, inadequate size of the prosthesis, and granulation or leakage around the fistula.\(^\text{36,37}\) The replacement of TEP valves adds additional costs as does regular appointments with the speech-language pathologist for education, changes, and troubleshooting.

An innate and distinct advantage of electrolaryngeal voicing is its intuitiveness and ease of use; in general, it is not difficult to learn how to use the electrolarynx although optimal use is attained through guidance from a speech-language pathologist. Due to the decreased requirement for extensive training prior to initial use and ease of commercial availability (although cost can vary widely), it can be utilized as a secondary or backup means of communication when the primary means of communication is unavailable. This frequently occurs as patients progress along their treatment pathway, including immediately postoperatively (i.e., while the TEP and primary resection sites are healing), if the patient undergoes radiotherapy (some delay TEP placement to allow for primary wound healing), and as an adjuvant therapy while mastering either TEP or esophageal voice methods.

As mentioned above, a major disadvantage of electrolaryngeal voicing is the mechanical quality of the voice which can result in greater perceived vocal handicap when compared to TEP, especially for tonal languages.\(^\text{16,18,38,39}\) Furthermore, reduced speech intelligibility is most pronounced for voiceless consonants as the electrolarynx supplies a continuous mechanical sound source.\(^\text{40}\) As outlined below, multiple studies have reported that the electrolarynx results in superior voicing when compared to esophageal speech but inferior results when compared to TEP restoration. Acoustic analysis studies have confirmed that tracheoesophageal speech has a more normalized fundamental frequency, greater intensity, and longer maximum phonation time than esophageal and electrolaryngeal speech.\(^\text{41}\) Finizia and Bergman reported that voice restoration with an electrolarynx resulted in inferior Sickness Impact Profile (SIP) and Hospital Anxiety and Depression Scale scores compared to both TEP and also laryngeal preservation with primary radiation therapy. Furthermore, the electrolarynx group was noted to have more patients with clinically important dysfunction (SIP score >10) for most categories.\(^\text{16}\) Clements et al found that patient satisfaction with their voice was superior with TEPs compared to the electrolarynx and esophageal speech.\(^\text{38}\) They found that TL patients with TEP reported significantly superior speech satisfaction, improved speech quality, increased ability to communicate through the telephone, less interaction limitations with others, and improved overall quality of life.\(^\text{38}\) Ward et al compared TEP to electrolarynx use and found significantly lower levels of disability, handicap, and distress for patients using tracheoesophageal speech.\(^\text{39}\)

In summary, the main disadvantages of the electrolarynx include (1) financial considerations in purchase and maintenance of the device which may be too fiscally burdensome for some, especially those in low income populations and developing countries;\(^\text{33,34}\) (2) the mechanical voice quality; (3) need for reasonable manual dexterity to use a hand held device (although the new EMG transducer device allows for hands-free voicing and makes poor manual dexterity less of a contraindication); and (4) the requirement for functioning articulatory musculature which can limit its usability in patients with extended laryngectomy resections. The electrolarynx is particularly suited to patients who do not wish further surgery or the possible complications associated with a TEP, and patients who either cannot vocalize with esophageal speech or choose not to.

**Silent speech innovations**

Digital speech recognition systems have been developed that measure the spoken sound and identify the output through various methods. In general, the recognition rates are 90% for continuous speech in patients with a functional larynx, but this is an imperfect system especially when there is considerable ambient noise. Newer innovations in projecting “silent speech” focus on tracking the articulatory tract (i.e., mouth, tongue, palate, etc.) movements as patients attempt to speak, with computer analysis then interpreting these movements and generating corresponding speech.\(^\text{42}\) There are a variety of techniques used to achieve this, the most common of which include palatometers or magnetic implants with sensors which focus on measuring the key articulators (i.e., velum, tip and body of tongue, and lips). Preliminary results suggest that this could become a successful method of voice generation\(^\text{42}\) and that development of a low-cost and portable articulography speech recognition program with high accuracy (77.2%) is feasible. However, this system remains to be validated in alaryngeal patients.\(^\text{44}\) Another avenue of parallel research involves noninvasive electroencephalogram techniques or implanted electrodes to generate speech; however, this novel technique is currently in its infancy.\(^\text{42,45}\)
Innovations and future of the electrolarynx

Although TEP voice is the current gold standard for voice rehabilitation postlaryngectomy as noted above, there are and will remain a subset of patients who will be better suited to electrolarynx use over TEP voicing. For these patients, future developments in the realm of the electrolarynx are necessary to enhance its usability and nuances of audibility. Such developments will likely allow for improved control of the electrolarynx in a hands-free manner, thus enabling its use by patients who have impaired manual dexterity, and improved range of intonation, thus enhancing its intelligibility and comparability to natural voice. There are several experimental electrolarynx devices that are currently not commercially available. As described earlier, in 2007 Goldstein et al developed an electromyographically controlled electrolarynx that is attached to the strap muscles, thereby allowing transcervical vibration when the strap muscles are activated and allowing for hand’s free control of the device.22 Choi et al described an “Evada” force sensing resistor type of experimental electrolarynx. Their model included a transcervical transducer with a sensor that can modulate frequency and intensity, allowing for changes in frequency and intensity depending on the spoken phrases.28 As mentioned above, the surgical extent of laryngectomy does impact an individual’s ability to use certain types of electrolarynges (e.g., strap musculature resection may prohibit use of an EMG transducer device, extensive tongue base resection may prohibit effective articulation with any devices). Thus, in addition to technical electrolarynx developments, pre-operative recognition of a patient’s need or desire to use an electrolarynx by the treating multidisciplinary surgical team could help to guide surgical reconstruction of laryngectomy defects in order to optimize use of individual devices for individual patients.

Conclusion

The electrolarynx is a viable and useful method of voice rehabilitation for patients who have had laryngectomies, particularly benefiting those patients for whom TEP use is not practical. Intraoral and transcervical electrolarynges are available with choice dependent on anatomical factors and patient preference. New developments allow the electrolarynx to be utilized with greater ease in patients with poor manual dexterity as well as improve its ability to produce voice with varied intonations.

Key points

• Voice impairment following TL can significantly detract from patients’ quality of life; voice restoration aims to reverse this deterioration.
• Voice restoration is achieved using three speech modalities: electrolaryngeal, esophageal, and tracheoesophageal. The electrolarynx is more straightforward to learn when compared to esophageal speech, but requires patients to purchase and maintain the device.
• The electrolarynx works by inducing vibrations of oral or pharyngeal mucosa by an external device, generally, at a constant fundamental frequency. The choice of device is dependent on anatomical factors and patient preference.
• New technological developments will allow the electrolarynx to be utilized with increased ease in patients with poor manual dexterity and improve the ability to speak with more varied intonations.

Acknowledgment

The authors wish to thank Dr Jason Gilde for his illustrations as used in this review.

Disclosure

The authors report no conflicts of interest in this work.

References


