



Elements of Clinical Training with the Electrolarynx

9

Kathleen F. Nagle

Elements of Clinical Training with the Electrolarynx

Laryngectomees have three main options for postsurgery speech communication: esophageal speech (ES) and tracheoesophageal (TE) speech, which have internal sound sources, and artificial speech produced using an external sound source. ES is a hands-free mode of postlaryngectomy speech that requires no prosthesis, device, additional surgery, or any particular maintenance. Successful ES speakers inject air into the esophagus and then control its release to create a “pseudo-voice” through vibration of pharyngoesophageal tissue (Globlek, Štajner-Katušić, Mušura, Horga, & Liker, 2004; Robbins, Fisher, Blom, & Singer, 1984; Štajner-Katušić, Horga, Mušura, & Globlek, 2006). TE speech, which is driven by pulmonary air, requires the creation of a fistula in the common tracheoesophageal wall, either at the time of laryngectomy or after the site of laryngeal reconstruction has healed (Brown, Hilgers, Irish, & Balm, 2003; Singer & Blom, 1980). A small, valved TE puncture prosthesis is then placed into the puncture site to maintain the link between the trachea and esophagus. This TE prosthesis provides unidirectional airflow from the lungs

to the vocal tract. By most accounts, TE speech is judged by listeners to be more acceptable and intelligible than other types of alaryngeal speech (Hillman, Walsh, Wolf, Fisher, & Hong, 1998; Most, Tobin, & Mimran, 2000; Pindzola & Cain, 1988; Williams & Watson, 1987). It is not uncommon, however, for laryngectomees to abandon TE speech in favor of an artificial larynx, presumably due to complications related to leaking or repeated extrusion of the prosthesis (Mendenhall et al., 2002; Singer et al., 2013).

The artificial larynx is an external sound source driven by pneumatic or electromechanical vibration. Pneumatic, reed-based devices (e.g., “Tokyo” larynx) provide an external alaryngeal voice source via a vibrating reed similar to that used in wind instruments (Nelson, Parkin, & Potter, 1975). Pulmonary air is directed into the oral cavity via a small-diameter tube that is coupled to a housing placed on the tracheostoma. Users can therefore use preoperative respiratory speech patterns to create natural-sounding durational speech characteristics (Weinberg & Riekens, 1973). The sound produced by the reed-based devices is also quite similar to the human voice because the reed vibrates in the same way as the vocal folds. Most importantly, reed-based devices lack the radiated noise that is an integral component of electronic sound sources. Pneumatic devices are rarely used in the United States today, presumably because coupling the device to the tracheostoma and managing the oral tube can be

K. F. Nagle (✉)
Department of Speech-Language Pathology, Seton
Hall University, South Orange, NJ, USA
e-mail: naglekat@shu.edu

awkward and unacceptably unsanitary for the user (Barney, 1958; Blom, 2000; Nelson et al., 1975).

The electrolarynx (EL) is an electromechanically driven device that supplies an entirely external, electronic sound source; no pulmonary driving air pressure is required. Because EL devices are comparatively simple and easy to use, they are used both in the early stages of postoperative care and ultimately chosen as the primary, backup, or emergency mode of speech by most laryngectomees (Doyle, 2005; Graham, 2006; Hillman et al., 1998; Meltzner & Hillman, 2005). This chapter will focus on EL devices and training in their optimal use.

Electromechanical Speech

EL devices provide a “voice” source with a range of potential options. The location and method of sound source transmission is a major determinant of these options (Meltzner et al., 2005). Transcervical or *neck-type* devices produce an external sound source that is transmitted to the oral cavity via vibration of neck, chin, or cheek tissue. The neck-type EL is the most commonly used artificial larynx among those who have a choice (Koike, Kobayashi, Hirose, & Hara, 2002). The neck-type EL contains a vibrating element and an electromagnetically driven vibrating membrane housed within a plastic or metal cylinder; several examples are shown in Fig. 9.1. Vibration of the EL membrane is initiated using a power button controlled by thumb pressure, with



Fig. 9.1 Electrolarynges with quarter for scale. From left: (a) NuVois III Digital™; (b) TruTone™; (c) Servox® Inton; (d) Romet® R210. (Photo by K.F. Nagle)

the amplitude of vibration controlled via a second button or dial. Successful use of this type of EL requires adequate coupling of the vibrating head of the device with the skin of the neck, chin, or cheek at what has been referred to clinically as the “sweet spot.” The tissue in the sweet spot(s) must retain sufficient elasticity to transmit the maximum amount of vibratory energy to the oropharyngeal cavity for speech production.

Transoral or *oral-type* ELs deliver acoustic energy directly to the oral cavity via a small-diameter plastic tube inserted into the mouth. The Cooper-Rand Electronic Speech Aid is the best known oral-type device made specifically for this purpose (Fig. 9.2); however, most neck-type ELs can easily be converted into oral-type devices using an oral adaptor that can be attached to the vibrating head of the EL. Because oral-type devices bypass radiated or reconstructed neck tissue, they are particularly attractive in certain contexts. For example, immediately following laryngectomy, tenderness in the healing pharyngoesophageal segment makes ES and TE speech difficult and painful. Effects of neck dissection or radiation may reduce the vibratory capability of tissue typically used in neck-type EL speech production; similarly, placement of the vibrating head of the neck-type EL directly



Fig. 9.2 Cooper-Rand Electronic Speech Aid. (Photo by K.F. Nagle)

on neck tissue may also cause substantial discomfort. Oral-type ELs are, therefore, a good choice for alaryngeal speech if adequate and comfortable coupling of the vibrating head of a neck-type EL cannot be achieved in the long term (Doyle, 2005; Hillman et al., 1998; Ward, Koh, Frisby, & Hodge, 2003).

As noted, most neck-type ELs are sold with an oral adaptor, which can be fitted over the vibrating head to provide the benefits of an oral-type EL. New laryngectomees who plan to use a neck-type EL may acclimate to EL use in the postsurgery period by using an oral adaptor while they heal. If the effects of radiation or surgery prohibit identification of an adequate sweet spot, they may choose to use the oral adaptor for the longer term. The adaptor consists of a rubber, plastic, or silicone cover for the head of the EL, with an opening in the center (shown fitted to an EL in Fig. 9.3a), and a small-diameter plastic or silicone tube, often with a fitted tip at one end, as in Fig. 9.3b. One end of the tube is placed into the EL cover opening, and the tip, which acts as a filter for saliva and food particles, is placed in the mouth.

Use of an oral adaptor requires learning to control the placement of the tube within the oral cavity. The tip of the long-tubed adaptor is ide-

ally placed off of midline by the upper teeth and inside of the cheek, with the open end facing the roof of the mouth. Figure 9.3d shows an EL fitted with long-tubed oral adaptor. Placement high in the mouth reduces interference from the tongue during articulation. High placement may also keep the tube cleaner for longer by reducing direct contact with the saliva that pools in the lower mouth. The open end of the short-tubed adaptor, shown fitted to an EL in Fig. 9.3c, is also placed between the teeth and cheek.

The hygiene issues related to constant use of a tube while speaking may be unpleasant and unacceptable for some users, and the presence of the tube can affect the user's intelligibility (Weinberg & Riekema, 1973). The voice produced with oral-type ELs also obviously lacks some of the spectral qualities of voice produced with the normal voice at the glottis (i.e., pharyngeal and nasal resonance); this further reduces the naturalness and intelligibility of speech already affected by the presence of the oral tube and its artificial source (Barney, 1958). Therefore, if they can use neck-type ELs, most laryngectomees opt for them over the oral-type (Koike et al., 2002).

Intraoral devices (also referred to as "palate devices") are variants of the oral-type EL. These devices mount to a dental plate or orthodontic retainer, generating a sound source directly within the oral cavity (e.g., Ultra Voice™). Vibration onset is controlled by the user with a remote switch that may be held in the hand, worn on the body, or kept in a pocket. Intraoral devices must be fitted to the individual user, making them comparatively expensive; however, once fitted, they are reportedly easy to use (Takahashi, Nakao, Kikuchi, & Kaga, 2005). There is limited empirical evidence of the popularity of intraoral ELs, but their use is likely restricted to laryngectomees who are unable to generate intelligible speech using other EL devices.

Users of EL devices must become physically and psychologically comfortable with the sound and feeling of the device, and the choice to make it a primary mode of communication depends on several factors. That is, not all laryngectomees are physically or cognitively able to use an EL. Surgery, chemotherapy, and radiation may



Fig. 9.3 Neck-type electrolarynges, with quarter for scale. From left: (a) NuVois III Digital™ fitted with head cover; (b) long oral tube; (c) Servox® Inton™ fitted with short-tube adaptor; (d) Romet® R210 fitted with long-tube adaptor. (Photo by K.F. Nagle)

impose temporary or permanent physiological and anatomical limitations that will affect successful use of the EL (e.g., fibrosis of neck tissue). For example, pain and impaired range of motion in the shoulder may lead to difficulty raising the arm or maintaining the position necessary for manual EL use (Doyle, 1994, 2005). Unlike ES and TE speech, EL speech has limited hands-free options.¹ Manual operation is generally required to initiate voice and adjust the amplitude of vibration for even the most advanced ELs currently available. Gripping the device while simultaneously manipulating one or more of its buttons may be particularly challenging. For some individuals, the need to operate an EL manually may make it an option of last resort, but most EL users are able to overcome this limitation.

Additional considerations may also create substantial challenges for some speakers. Some may initially reject the use of an EL because of the uniquely robotic, “buzzing” sound and the radiated noise associated with electromechanical speech (Doyle, 2005; Espy-Wilson, Chari, MacAuslan, Huang, & Walsh, 1998; Meltzner & Hillman, 2005; Qi & Weinberg, 1991). Artificial speech of any type is distracting, and even perfect coupling between EL and skin tissue cannot eliminate all non-speech noise radiated by the device. Moreover, most EL speech notoriously lacks the prosodic characteristics that make speech interesting and intelligible (Bien et al., 2008; Gandour & Weinberg, 1983; Gandour, Weinberg, & Garziona, 1983; Hillman et al., 1998). The sound of EL speech is especially problematic for female laryngectomees, who may be socially penalized for the perceived low pitch of their EL voices (Cox, Theurer, Spaulding, & Doyle, 2015; Nagle, Eadie, Wright, & Sumida, 2012). The monopitch quality of most ELs may also affect women to a greater degree than men, as women tend to have greater speaking fundamental frequency (f_0) ranges (Goy, Fernandes,

Pichora-Fuller, & van Lieshout, 2013; Pepiot, 2014). The consequence for female laryngectomees is frequent misidentification as male when listeners lack visual information (e.g., when speaking on the telephone; Smithwick, Davis, Dancer, Hicks, & Montague, 2002). Anecdotally, some EL users have reported that they or their family members or friends just do not like the sound of the EL (Eadie et al., 2016).

Given that all laryngectomees are likely to use an EL at some point, it is essential for SLPs to be aware of the basic and advanced features of ELs. The choice of a particular EL model depends on several design factors described in the next section.

Design Features of the Electrolarynx

The “perfect” artificial larynx would mimic a natural voice source. It would be unobtrusive, reliable, hygienically acceptable to the user, inexpensive, and simple to operate; its output would match that of a natural voice in volume, quality, and pitch inflection (Barney, 1958; pp. 558–559). To varying degrees, most of these criteria have been met by currently available ELs. The size and shape of ELs has changed over time so that some neck-type ELs can now be nearly hidden in a man’s fist as he speaks. Technological advances have increased the EL’s reliability, ease of use, and range of options for mimicking natural laryngeal speech (Searl, 2006; Meltzner et al., 2005). The relative cost has concurrently dropped. For users who are uncomfortable with the hygienic drawbacks of using an oral adaptor or intraoral device, neck-type ELs provide a relatively clean and user-friendly alternative. A discussion of ongoing developments in EL technology is beyond this chapter, but the current focus is on energy-efficient, wireless, hands-free activation and dynamic pitch modulation as a means of refining the naturalness of EL speech (Guo, Nagle, & Heaton, 2016; Heaton, Robertson, & Griffin, 2011; Matsui, Kimura, Nakatoh, & Kato, 2013; Nakamura, Toda, Saruwatari, & Shikano, 2012; Stepp, Heaton, Rolland, & Hillman, 2009;

¹The HandsFree™ Electrolarynx Holder (Griffin Labs, Temecula, CA) is worn around the neck and activated with a chin press; it is not appropriate for all users, however, as optimal use may occur only within a specific and limited area of the user’s neck space.

Wan, Wu, Wu, Wang, & Wan, 2012; Wu, Wan, Xiao, Wang, & Wan, 2014).

All ELs are designed with three features that the user can control directly: activation of vibration, volume adjustment, and base pitch setting. Even a basic analog EL model allows the user to control activation and amplitude of vibration using a button, dial, or “rocker switch.” Volume range is preset by the manufacturer and intuitive for the user. Volume is usually adjusted by rotating a small thumb dial. The volume dial is frequently placed next to the activation button, allowing users to adjust the loudness of the device with the thumb as well. Despite the proximity of these mechanisms, however, two fingers would be needed to permit simultaneous control of activation and amplitude of vibration. Typically, thumb-button users have to stop “vocalizing,” at least momentarily, in an effort to effectively adjust the amplitude of vibration. In such circumstances, one may subsequently observe limitations in real-time prosodic variation of loudness by the user.

Pitch capabilities vary significantly from device to device, and base pitch settings can be complicated to alter. Most ELs offer a range of potential base pitch settings. Users can adjust the EL to suit their preference as a base “speaking” pitch, but it will vibrate at a single f_0 . For newer digital devices, pitch adjustments can be accomplished electronically, but analog ELs require manual adjustment. Adjusting the base pitch of an analog device may be as simple as dialing a wheel on the side of the device, or it may require opening the device to access its mechanical workings. For individuals with reduced dexterity or visual acuity, these adjustments may be an intractable challenge.

Several EL models offer two base pitch settings, with one button assigned to each. The advantage of two settings is that two preset f_0 s can be achieved; each button produces its own pitch. The user is, therefore, able to alternate monotone base pitches. This ability to change pitch as a signal of paralinguistic features (e.g., “yelling” at the higher pitch, using the lower pitch only to sound authoritative, etc.) may be attractive to some users despite the monotone

quality of speech output. In some models, speech volume can also be adjusted separately for each pitch setting, which offers further flexibility of use, particularly across communication settings.

Dynamic (real-time) pitch modulation is currently available in a single EL model (TruTone™, Griffin Labs, Temecula, CA). This device allows users to produce more natural prosody by adjusting the degree of finger pressure placed on the activation button/tone sensor. Once the pitch range is set, users can modify pitch while speaking by altering finger or thumb pressure on the activation button located on the exterior of the device. Increased finger or thumb pressure on the activation button results in increased pitch; as the button is released, pitch drops to the baseline level. Setting a relatively wide pitch range accommodates more natural-sounding changes in prosody; however, as the set pitch range increases, it becomes more difficult to control pitch with finger pressure. It is relatively easy to maintain the maximum pitch with maximal thumb pressure. It can be difficult to sustain the minimum pitch, however, because of the necessity of keeping thumb pressure at a level that is “just detectible” by the device. Before operating such a device, users must set its pitch range by adjusting two tiny actuating dials inside the device housing. Very small adjustments to these dials can lead to rather large changes in fundamental frequency, so the process of setting a pitch range can be time-consuming and a bit frustrating, particularly if dexterity or visual problems exist.

In practice, many TruTone™ users do not take advantage of pitch variability inherent to the device (Nagle & Heaton, 2016, 2017). Clinical observation of EL users’ behaviors with the device suggests three related reasons for this. First, there is not an intuitive link between subconscious prelaryngectomy pitch regulation (i.e., using laryngeal musculature) and conscious postlaryngectomy pitch modulation using the thumb. Attempting to execute real-time pitch changes may require an unusual degree of attention to speech output. Practically speaking, the cognitive load connected with using dynamic, thumb-button pitch modulation may be too much for many users. Second, the quick, precise muscular changes associated with

intonation in laryngeal speech are quite difficult to match with thumb pressure alone. The dexterity needed to capitalize on thumb-button pitch modulation also may be too great for some EL users. Recent work by Al-Zanoon, Parsa, Lin, and Doyle (2017) has revealed that despite the mechanical capacity of an EL device to produce pitch changes, the ability for the user to effectively convey such changes is challenging. Third, although the thumb is used to control activation and adjustment of vibration in the TruTone™, it also may be needed to stabilize placement of the device against the skin for optimized signal transmission. That is, when maintaining coupling of the EL with the sweet spot, it is easier to apply consistent thumb pressure, than to attempt to vary it. In terms of intelligibility, it is arguably better to produce some pitch variability than none at all (Bunton, Kent, Kent, & Duffy, 2001). Nonetheless, both intelligibility and naturalness could likely be improved if TruTone™ users manipulated the device to match natural pitch contours (Watson & Schlauch, 2009). Ultimately, most TruTone™ users seem to produce speech that is perceived as nearly monopitch, despite the capability of the device to do more.

Using the Artificial Electrolarynx

It is rare for individuals to pick up an EL for the first time and immediately produce intelligible speech with it. To reduce the amount of radiated noise from the device, users must first be instructed to identify the location at which the device couples best to the tissue. In addition to finding the sweet spot, it is necessary to master the features of the device itself and to modulate the articulators to accommodate and maximize the quality of the artificial voice. With good instruction, most who are laryngectomized can learn to use either the neck-type or intraoral EL effectively.

Basic Operation of the Electrolarynx

When providing options to the new user, the SLP should be able to model the use of any ELs being considered, and familiarity with several EL mod-

els/types is advisable. If possible, users should initially be trained to hold and activate the device with their non-dominant hand. Although a simple button push activates the EL, most users need to be trained in how best to manage it during connected speech (Doyle, 1994, 2005). For example, many users' instinct is to deactivate the vibration between each word, producing a staccato-sounding speech quality that may reduce the intelligibility and naturalness of their speech. Modifying this behavior may require a discussion about pausing, phrasing, and the voicing characteristics of running speech. Conversely, new users may initially fail to deactivate vibration between utterances, producing one long buzz of noise. Fortunately, this latter tendency is easy to correct once it is pointed out and instructions for modification provided.

Laryngectomees need guidance as they find their electrolaryngeal "voice"; once they have identified and begun to use it, they are unlikely to want to make changes. Adaptability of in different communication contexts is a particular strength of ELs, and SLPs are uniquely qualified to instruct users in how to exploit this flexibility. For example, when setting the habitual volume for their devices, users may benefit from the practiced ear of the SLP to guide them. Some users set EL volume to a lower than optimal level in an apparent attempt to reduce its noise, unnaturalness, or robotic sound. These users can be trained to adjust the volume to suit their environment, including considering the effects of EL speech on potential communication partners with hearing loss. (Older EL users may also benefit from evaluation of their own hearing acuity as they contemplate maximizing communicative effectiveness.)

Clinical observation likewise suggests that new users may need assistance in choosing a base pitch setting. There are several rules of thumb regarding the pitch of EL speech. Although typical male laryngeal speaking f_0 ranges between 100 and 146 Hz (Baken & Orlikoff, 2000), evidence suggests that setting the EL below 100 Hz may provide a better outcome for most men, as intelligibility is relatively enhanced at lower speaking f_0 (Nagle et al., 2012; Watson &

Schlauch, 2008). The gender-neutral range of laryngeal speech seems to be wider than that of laryngeal speech (roughly 145–165 Hz; Gelfer & Bennett, 2013); that is, there appears to be a bias toward perceiving EL speech at even higher f_0 s as male (Nagle et al., 2012). In fact, setting the EL within the typical female speaking f_0 range of 188–221 Hz may be counterproductive (Baken & Orlikoff, 2000; Nagle et al., 2012). It is clear that female laryngectomees may have to choose between intelligibility and sounding female. This is a potential trade-off that must be considered by individual users, and SLPs are uniquely qualified to provide education and counseling to new laryngectomees.

Optimizing Electromechanical Speech

SLPs training new users should help them become comfortable with EL features, but they may also need to boost the user's knowledge of the features of speech that contribute most to comprehensibility and naturalness. Successful EL users have the metalinguistic awareness to maximize intelligibility and minimize distractions to communication partners that accompany the use of an artificial larynx.

One way for new EL users to think about their speech output is to imagine how it is perceived by potential communication partners. The sound of EL speech affects not only the user's comfort with the device but also the ability of listeners to understand what is said. Specifically, the speech signal should be perceptually separable from the accompanying non-speech noise emitted by an EL. Listeners may have trouble parsing signal from noise, however, if the user fails to couple the device properly to neck or cheek tissue or is unable to filter the EL voice source appropriately within the oral cavity. Likewise, paralinguistic aspects of speech such as pitch and loudness contours are lost in typical EL speech (Gandour & Weinberg, 1983; Gandour et al., 1983). These reductions in the complexity of signal quality can affect speech intelligibility (Doyle, 2012). Speech *intelligibility* is the degree to which the acoustic

signal is understood, without context. *Comprehensibility*, in contrast, encompasses acoustic, visual, gestural, and proxemic information, along with other contextual factors. Comprehensibility is the extent to which a listener understands utterances produced by a speaker in a communication context (Yorkston, Strand, & Kennedy, 1996). Speakers can improve their comprehensibility by improving their intelligibility and by making the most of gestural and other nonverbal cues to communication. Enhanced EL communication is aided by considering multiple factors that improve comprehensibility for which SLPs have particular expertise. These include optimizing perceptual quality, emphasizing salient suprasegmental cues, and attending to nonverbal communicative signals.

Perceptual aspects of voice quality beyond comprehensibility are particularly important for laryngeal speech because of the potential effects of its atypical sound source on the success of communicative interactions (Doyle & Eadie, 2005; Meltzner & Hillman, 2005). Despite the similarity of electromechanical sources, the quality of EL speech can vary quite a bit among users, given the unique characteristics of an individual's oronasopharyngeal cavity (i.e., the speech filter) before and after laryngectomy. Dynamic aspects of speech production may be affected by additional medical or surgical procedures (e.g., glossectomy, radiation), reducing the accuracy of speech sound production for some laryngectomees. To compensate for a reduction in segmental accuracy, EL users must attend to suprasegmental factors such as pitch, duration, and loudness. Although they may not increase intelligibility per se, adjustments to suprasegmental aspects of EL speech may improve communicative success by enhancing speech and voice quality.

Alaryngeal voices are frequently described on the basis of their speech *acceptability* or perceived *naturalness* (Bennett & Weinberg, 1973; Eadie & Doyle, 2002, 2005). Speech acceptability is a multidimensional descriptor including naturalness, pleasantness, and the degree to which the voice is not distracting (Bennett & Weinberg, 1973; Most et al., 2000). Perceived

naturalness typically addresses the degree to which the rate, intonation, rhythm, and stress pattern of disordered speech resemble normal speech (Eadie & Doyle, 2002; Meltzner & Hillman, 2005). Pitch, loudness, rate, and rhythm changes may affect not only comprehensibility but also overall communicative success if communication partners are “put off” by the sound of speech or feel that they have to expend too much effort to listen to it. Ultimately, if speech sounds unnatural enough, the EL user’s quality of life may suffer (Eadie et al., 2016; Law, Ma, & Yiu, 2009).

Training Goals

Because of the unusual sound sources, alaryngeal speech provokes different perceptual expectations from other types of disordered speech or voice. ES and TE speeches are characterized by highly aperiodic sources of relatively low signal amplitude, whereas EL speech generally features flat or near flat intonation accompanied by radiated noise. EL speech introduces an external noise source that competes with the very speech signal it is designed to enhance. For example, differences between voiced and voiceless speech sounds are generally not perceptible in connected EL speech given the necessarily constant vibration of the device source. Overall goals of maximizing user comfort, comprehensibility, and naturalness depend heavily on minimizing the effects of radiated noise (Doyle, 1994; Graham, 2006). Finding a sweet spot where the EL produces the least rattle and the most oral resonance is the first goal of learning how to use a neck-type EL. The sweet spot should be a location where the tissue is most elastic and close enough to the oral cavity to maximally amplify the vibration of this tissue. If contact between the head of the EL and the skin is incomplete or lost, noise will radiate directly from the EL, and the capacity to produce speech will be lost until adequate contact is regained. Placement of the EL head must be also comfortable enough for the user to maintain during speech and reachable by the user every time he/she wants to speak. Likewise, when using an

intraoral-type EL, the optimal placement of the oral tube must be maintained.

As mentioned above, another way to improve the perceived naturalness of EL speech is to exaggerate its prosodic characteristics. For example, lexical and prosodic stresses are generally marked in normal laryngeal speech by longer duration, higher pitch, and increased volume. In contrast, EL users tend to intuitively mark lexical or syntactic stress using duration and by making stressed syllables relatively longer than unstressed syllables (Gandour & Weinberg, 1984). If they do not make such adjustments, instructing them to do so may increase the naturalness (and potentially comprehensibility) of their connected speech.

Maximizing intelligibility often increases comprehensibility and naturalness. As mentioned previously, perception and distinction of speech sounds that differ only in voicing are affected for EL speech because when the device is activated, its “voice” is always on. Turning off the EL during production of unvoiced cognate sounds (e.g., /p, t, k, s/) is not feasible during running speech and not advisable even for short phrases as a rule. Although on-off control serves as an important EL skill to enhance communication, the onset or termination of the signal must fall at points within a given utterance where such changes would also appear for a normal speaker. During speech, the EL should be silenced only at grammatically appropriate points in an utterance (i.e., between breath groups). Repeatedly turning the device on and off at very brief intervals creates an unpleasant staccato effect that is likely to negate any intelligibility gained by producing voiceless consonants without a voice. Simply put, the phonetic features of running speech change too quickly for this type of adjustment.

A more top-down approach that may maximize intelligibility for EL speech involves *clear speech* (Cox, 2016). Clear speech is a speaking style adopted to increase intelligibility in difficult listening situations (Krause & Braida, 2003). Speakers instructed to use clear speech make subconscious changes to enhance speech clarity. Initially it may be helpful to simply instruct the user to imagine speaking to someone who is hard

of hearing, as this often prompts intuitive use of clear speech. Clear speech has specific properties that maximize intelligibility, such as over-articulation (Smiljanić & Bradlow, 2009). Using clear speech tends to cause individuals to reduce speech rate as well, often by taking longer and more frequent pauses (Smiljanić & Bradlow, 2009). Although reducing speech rate by too much can decrease its naturalness, existing data suggests that reducing speech rate is generally beneficial to intelligibility (Yorkston, Hammen, Beukelman, & Traynor, 1990).

Comprehensibility is also increased by capitalizing on nonverbal cues, and it may be necessary for SLPs to attend specifically to gesture, body language, and proxemics when training an EL user. Reduced paralinguistic information in EL speech heightens the relevance of these nonverbal communicative cues. Visual cues in particular can reinforce the delivery of an intended message to a communication partner. Because most ELs require manual voice activation, however, at least one of the user's hands may not be free for simultaneous gestural communication. Users of oral-type ELs may display unusual facial expressions while attempting to manage the oral tube. In addition to the potentially decreased intelligibility caused by the articulatory limitations of the EL, the user's comprehensibility may be further limited by reduced access to these types of nonverbal cues. Consequently, the next section outlines several specific tasks that can be used to train laryngectomees in the successful use of EL devices.

Training Targets

Certain speech tasks are difficult for all laryngectomees, but EL users face particular challenges. Because EL speech is continuously voiced and lacks driving air pressure, voicing and manner cues may be lost unless specific attention is paid to emphasizing them. Fortunately, phonetic cues such as duration can be exaggerated to influence what listeners perceive. Therefore, EL users should learn to make the most of segmental, suprasegmental, and nonverbal features that

complement what they are initially able to produce with an EL.

Production of specific segments Speech production is similar for all types of ELs, although the oral tube may complicate articulation of certain segments. For example, placement of the tube may interfere with lip closure and tongue movement for labial and lingual consonants. New users may want to take a hierarchical approach to learning to use the EL. That is, they might begin to practice by producing simple consonant-vowel or vowel-consonant syllable and then move to multisyllable words, phrases, and beyond (Graham, 2006). Given relatively intact articulators, EL users who have located their sweet spot should be able to produce vowels with little training. It may be most instructive to new EL users to start with production of diphthongs. Placing the EL at the sweet spot and producing oversized vowel combinations such as "ow" and "aye" will immediately give the user a sense of what to expect from EL speech. It may also be necessary at first to make big oral gestures for both consonants and vowels to compensate for the abnormal acoustic qualities of EL speech (Wu, Wan, Wang, & Wan, 2013). A trial-and-error approach is generally adequate for learning to differentiate vowels.

Skilled EL users can capitalize on the redundancy of speech cues, such as the influence of vowel features on the consonant perception. Duration can be strategically modified to hint at consonant features not fully articulated in alaryngeal speech. Vowels preceding voiceless consonants are perceived as shorter in duration than those preceding voiced consonants (Peterson & Lehiste, 1960). To create a voicing distinction in the absence of voicing cues, vowels preceding voiced consonants should be strategically lengthened (Weiss & Basili, 1985; Weiss, Yeni-Komshian, & Heinz, 1979). Likewise, vowels following /h/ may be perceived as longer than syllable-initial vowels. As an unvoiced glottal fricative, /h/ is problematic for EL speakers who lack access to pulmonary air or a glottis. Even if it were feasible to turn off an EL during running speech to produce /h/, laryngectomees would

have difficulty generating adequate frication for the sound. To influence the listener to perceive /h/, a speaker can strategically lengthen vowels meant to follow /h/.

To differentiate consonant segments, EL users should be instructed exaggerate the phonetic features that remain available to them. Without voicing contrasts, they must over-articulate both place and manner characteristics. To the degree possible, EL users should approximate the burst “plosion” that accompanies stop consonants. Laryngectomees lacking a TE puncture cannot build up high intraoral pressures using pulmonary air. If the tongue, lips, and velum are intact and functioning, however, they can exploit ambient air pressure to overdo the release of stop consonants. Some refer to this as “popping” a consonant, a maneuver that is common to traditional esophageal speech training (Doyle, 1994). Similarly, with guidance, laryngectomees should work to lengthen the duration of fricative and nasal consonants, to exaggerate both their manner and place of articulation. They may also choose to produce voiceless targets for a relatively shorter duration than their voiced cognates.

For users of oral-type ELs, lingual consonants (e.g., /t, d, s, z/) can also be challenging because the oral tube may impede natural tongue movements. Individuals will have to experiment with the device to find the most practical way to produce these sounds clearly. For these and all speech sounds, it is critical that the user keep the oral tube relatively high and out of the way of the tongue. The use of a short-tubed adaptor may in some instances alleviate the interference of the tube with lingual movements, but because the EL itself must consequently be placed closer to the mouth, the tube adaptor may get in the way of labial movements. Thus, careful monitoring of a larger set of speech behaviors is essential in the treatment process.

Production of suprasegmentals Although the lack of a voicing distinction can affect the intelligibility of EL speech, the absence of prosodic features arguably affects it even more (Watson & Schlauch, 2008). The markers of stressed syllables,

a major feature of prosody, are increased loudness, duration, and pitch. Because few users actually manipulate ELs with dynamic pitch and amplitude modulation to align with natural speech contours, the only prosodic feature available to most EL users is duration. Research suggests EL users tend to lengthen syllables to signal stress; they also lengthen pauses following a final stressed syllable or preceding an initial stressed syllable (Gandour & Weinberg, 1984). Those users who do not automatically differentiate stressed from unstressed syllables with duration differences may need to be instructed to do so. A directive to use clear speech may instigate an immediate change (Smiljanić & Bradlow, 2009).

As with the production of single segments and syllables, inter-syllable duration can be used strategically to modulate EL speech. For example, EL users can strategically lengthen the pause preceding or following a stressed syllable. Depending on the device, they may be able to adjust pitch or volume in real time as well (although no currently available EL device offers the option of adjusting both of them simultaneously). If pitch modulation is offered, it is sensible for the EL user to take advantage of it. At the very least, exploration of pitch modulation may provide the speaker with a better understanding of the complexity of speech. The two general options for pitch modulation are described in the following section.

Optimizing two-button pitch modulation As described previously, a few EL models allow the user access to two base pitches. The differences come down to the simplicity of pitch adjustment and the ease of button and switch activation. In addition to volume control, two-pitch-button ELs have upper and lower buttons on the upper side of the external casing. For example, the Servox@ Inton™ (Servona, Troisdorf, Germany) has what the manufacturer calls a “base tone” button and an “accentuation tone” button (see Fig. 9.1c). The user is directed to use the upper button to produce a base pitch while speaking. To emphasize certain words, the user must press both buttons at the same time. Although the accentuation button is directly under the tone button, it is necessary for

most speakers to use two fingers (rather than a thumb) to accomplish this task. As an option, the Inton™ EL can also be programmed by the SLP to drop slightly in pitch as the base pitch button is released. This pitch drop is meant to replicate the natural pitch drop at the end of declarative utterances. Because it must be adjusted via computer, this feature cannot be toggled on and off during use. Although the Inton™ is digital, pitch must be adjusted manually, by opening the device and toggling the “dip switch” for the given button.²

The NuVois III Digital™ (NuVois, Meridian, ID) has two pitch buttons and two volume buttons (Fig. 9.1a). Pitch is adjusted by toggling one of the pitch buttons and pressing the volume buttons until a pitch is chosen. Pitch can be changed without opening the device, but not in real time; the user must stop and make adjustments to the device each time he/she wishes to change its pitch. As with the Inton™, this device can produce two distinct pitches, one of which can be used to indicate emphasis. The pitch and volume buttons are roughly on the same horizontal plane, but the volume buttons are smaller; users can learn to differentiate them by their size.

The bottom line for two-button pitch modulation is that a maximum of two base pitches are available to the user. The user cannot produce the full range of pitch between the two settings and can only hint at intonation with the second pitch button. It is, therefore, important to decide how and when the second button will be used. If the pitch difference between the button settings is too great, any use of it will suggest great excitement or anger. If it is too small, it may not be registered to the listener as a linguistic difference. The user must also decide how frequently to use the second pitch. For example, will it indicate any stressed syllable or just the stressed syllable in the word with the most prominence in the utterance? Will it be used only to indicate a question? Once prosodic patterns have been used for a period of time, they will become part of the user’s new voice identity. Attempts to modify how an

individual uses the EL, even in the interest of increasing comprehensibility or naturalness, may confuse listeners who have become used to the user’s postlaryngectomy EL voice. Again the SLP can play an important role in identifying targets and providing practice for these changes.

Optimizing dynamic pitch modulation The TruTone™ EL has a single button for activation and pitch modulation, along with a volume wheel that must be operated separately. As described previously, modulating pitch for this device requires the user to gauge the degree to which the button is depressed using haptic feedback. There are limits to the range that can be accurately distinguished using thumb pressure. Assuming that an appropriate pitch range has been set, users will need to figure out their aptitude by experimenting with the device – an experimental approach is appropriate. Targeting natural pitch contours requires recognition of the characteristics of normal intonation; that is, users of ELs with dynamic pitch modulation may need to spend time just listening to natural laryngeal speech. Specifically, they will have to learn to where the greatest stress is placed in an utterance. Unlike users of two-button type ELs, however, they will have the ability to produce gradations of pitch change. To make the most of this feature, users will have to think about aspects of intonation beyond emphasizing the word with the most stress in an utterance. SLPs, who are trained to attend to and identify perceptual characteristics that increase intelligibility, can be invaluable during this phase.

Optimal use of dynamic pitch requires some attention to the role of prosody in laryngeal speech. One fairly easy way to capitalize on dynamic pitch modulation is to mark “WH” questions with a rising/falling set of terminal tones. Yes/no questions are similarly marked by rising terminal tones in North American English. A skilled EL user will drop the pitch slightly at the end of a declarative sentence. An even more skilled user will use pitch shifts when listing items, providing a string of numbers, or maintaining a turn before pausing within an utterance. Users should listen to themselves and decide what sounds most natural to them. They should

²There are videos online (e.g., <https://www.youtube.com/watch?v=iV3rP%2D%2DrcTA>) that demonstrate this process.

also get feedback not only from a SLP but also from familiar conversation partners. A systematic and integrated approach to enhancing communication is of extreme importance clinically. Such an approach necessarily includes attending to nonverbal communication.

Nonverbal Communication

Nonverbal communicative cues include vocal non-speech sounds, gesture, and body language. The loss of natural non-speech vocalizations can have a dramatic effect on the quality of communication interactions. Laughter, for example, is a spontaneous vocal reaction. EL users have to decide whether their new laugh will be silent or artificial or coded through respiratory sounds (Doyle, 2008). Some adopt a strategy of smiling and saying “ah-ah-ah” with the EL. It is similarly worth the effort for EL users to come up with replacements for conversational fillers, such as the grunting assent or the heavy sigh. Some EL users choose the silent nod or head shake to move conversation along; others may choose to activate the EL briefly. Appropriate decisions about these aspects of speech pay dividends in naturalness and overall comprehensibility, even if overall intelligibility may have been compromised.

As a rule, it is easy enough to raise the volume on most devices if the user wants to communicate anger or excitement. Similarly, speaking in short bursts of one or two words may signify that the user is upset. EL speech is currently incapable of conveying much else in the way of emotion, however, and EL users may have to rely on facial expressions or gesture to get their feelings across. Likewise, beyond lengthening its duration, the ability to emphasize a given word or utterance is lost to many EL users. To maximize comprehensibility in the absence of pitch modulation, EL users may want to use or exaggerate facial expressions or gestures that complement their speech. Even with only one hand available, the options listed in Table 9.1 can provide valuable paralinguistic cues.

Table 9.1 Gestures and facial expressions that complement EL speech

Communicative intent	Action
Questions	Upturned hand Raised eyebrows Tilted head
Emphasis	Raised, open hand Widened eyes
Approval/Disapproval	Thumbs up/thumbs down
Impatience	Closed eyes Pressing lips together

The specific nonverbal cues that an EL user provides, intentionally or otherwise, may vary, but the topic of nonverbal communication is important for SLPs to discuss with EL users. Comprehensibility may not seem like an issue for some EL users until they find themselves in a noisy environment or on the telephone, where nonverbal cues are not available. The comprehensibility and naturalness of EL speech may decrease if nonverbal cues are not intuitive to the user but may increase if these cues can be exploited appropriately. SLPs can help by role-playing situations in which nonverbal communication makes a difference.

Conclusions

This chapter has addressed optimal use of currently available ELs. Although EL use is not appropriate for all laryngectomees, most use it as a primary or backup mode of communication. SLPs should be familiar with the variety of EL features available and should be able to model appropriate use of numerous devices. Beyond choosing the suitable EL, the role of the SLP is to maximize the user’s comfort with the chosen device and to assist in making EL speech as natural and comprehensible as possible. Specific EL training targets should target segmental, suprasegmental, and nonverbal aspects of communication unique to EL speech production. In particular, this means capitalizing on redundant phonetic cues that aid in perception of EL speech; attending to available options for dynamic pitch modulation; and heightening awareness of nonverbal communicative cues. Clinical attention to factors such as these will boost communicative success for EL users.

References

- Al-Zanoon, N., Parsa, V., Lin, X., & Doyle, P.C. (2017, November). Using visual feedback to enhance intonation control of electrolarynx speakers. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association, Los Angeles, CA.
- Baken, R., & Orlikoff, R. (2000). *Clinical measurement of speech and voice* (2nd ed.). San Diego, CA: Singular Pub. Group.
- Barney, H. L. (1958). A discussion of some technical aspects of speech aids for post-laryngectomized patients. *The Annals of Otolaryngology, Rhinology, and Laryngology*, 67(2), 558–570.
- Bennett, S., & Weinberg, B. (1973). Acceptability ratings of normal, esophageal, and artificial larynx speech. *Journal of Speech and Hearing Research*, 16, 608–615.
- Bien, S., Rinaldo, A., Silver, C. E., Fagan, J., Pratt, L., Tarnowska, C., ... Ferlito, A. (2008). History of voice rehabilitation following laryngectomy. *The Laryngoscope*, 118(3), 453–458. <https://doi.org/10.1097/MLG.0b013e31815db4a2>
- Blom, E. C. (2000). Current status of voice restoration following total laryngectomy...including commentary by Shah, J.P. *Oncology (08909091)*, 14(6), 915–931.
- Brown, D. H., Hilgers, F. J. M., Irish, J. C., & Balm, A. J. M. (2003). Postlaryngectomy voice rehabilitation: State of the art at the millennium. *World Journal of Surgery*, 27(7), 824–831. <https://doi.org/10.1007/s00268-003-7107-4>
- Bunton, K., Kent, R. D., Kent, J. F., & Duffy, J. R. (2001). The effects of flattening fundamental frequency contours on sentence intelligibility in speakers with dysarthria. *Clinical Linguistics & Phonetics*, 15(3), 181–193.
- Cox, S. R. (2016). *The application of clear speech in electrolaryngeal speakers* (Dissertation). The University of Western Ontario, London, Ontario.
- Cox, S. R., Theurer, J., Spaulding, S., & Doyle, P. C. (2015). The multidimensional impact of total laryngectomy on women. *Journal of Communication Disorders*, 56, 59–75. <https://doi.org/10.1016/j.jcomdis.2015.06.008>
- Doyle, P. C. (1994). *Foundations of voice and speech rehabilitation following laryngeal cancer*. San Diego, CA: Singular Pub. Group.
- Doyle, P. C. (2005). Clinical procedures for training use of the electronic artificial larynx. In *Contemporary considerations in the treatment and rehabilitation of head and neck Cancer* (pp. 545–570). Austin, TX: Pro-Ed.
- Doyle, P. C. (2008). Conveyance of emotion in postlaryngectomy communication. In K. Izdebski (Ed.), *Emotions in the Human Voice, Volume II: Clinical Evidence* (Vol. 2, pp. 49–67). San Diego, CA: Plural Publishing.
- Doyle, P. C. (2012). Seeking to better understand factors that influence postlaryngectomy speech rehabilitation. *Perspectives on Voice and Voice Disorders*, 22(1), 45–51. <https://doi.org/10.1044/vvd22.1.45>
- Doyle, P. C., & Eadie, T. L. (2005). The perceptual nature of alaryngeal voice and speech. In *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing* (pp. 113–140). Austin, TX: Pro-Ed.
- Eadie, T. L., & Doyle, P. C. (2002). Direct magnitude estimation and interval scaling of naturalness and severity in tracheoesophageal (TE) speakers. *Journal of Speech, Language, and Hearing Research*, 45(6), 1088–1096. [https://doi.org/10.1044/1092-4388\(2002\)087](https://doi.org/10.1044/1092-4388(2002)087)
- Eadie, T. L., & Doyle, P. C. (2005). Scaling of voice pleasantness and acceptability in tracheoesophageal speakers. *Journal of Voice*, 19(3), 373–383. [https://doi.org/S0892-1997\(04\)00071-2](https://doi.org/S0892-1997(04)00071-2) [pii]. <https://doi.org/10.1016/j.jvoice.2004.04.004>
- Eadie, T. L., Otero, D., Cox, S., Johnson, J., Baylor, C. R., Yorkston, K. M., & Doyle, P. C. (2016). The relationship between communicative participation and postlaryngectomy speech outcomes. *Head & Neck*, 38(Suppl 1), E1955–E1961. <https://doi.org/10.1002/hed.24353>
- Espy-Wilson, C. Y., Chari, V. R., MacAuslan, J. M., Huang, C. B., & Walsh, M. J. (1998). Enhancement of electrolaryngeal speech by adaptive filtering. *Journal of Speech, Language & Hearing Research*, 41(6), 1253–1264.
- 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.
- Gandour, J., Weinberg, B., & Garziona, B. (1983). Perception of lexical stress in alaryngeal speech. *Journal of Speech and Hearing Research*, 26, 418–424.
- Gelfer, M. P., & Bennett, Q. E. (2013). Speaking fundamental frequency and vowel formant frequencies: effects on perception of gender. *Journal of Voice*, 27(5), 556–566. <https://doi.org/10.1016/j.jvoice.2012.11.008>
- Globelek, D., Štajner-Katušić, S., Mušura, M., Horga, D., & Liker, M. (2004). Comparison of alaryngeal voice and speech. *Logopedics Phoniatrics Vocology*, 29(2), 87–91.
- Goy, H., Fernandes, D. N., Pichora-Fuller, M. K., & van Lieshout, P. (2013). Normative voice data for younger and older adults. *Journal of Voice: Official Journal of the Voice Foundation*, 27(5), 545–555. <https://doi.org/10.1016/j.jvoice.2013.03.002>
- Graham, M. S. (2006). Strategies for excelling with alaryngeal speech methods. *SIG 3 Perspectives on Voice and Voice Disorders*, 16(2), 25–32. <https://doi.org/10.1044/vvd16.2.25>
- Guo, L., Nagle, K. F., & Heaton, J. T. (2016). Generating tonal distinctions in Mandarin Chinese using an electrolarynx with preprogrammed tone patterns. *Speech*

- Communication*, 78, 34–41. <https://doi.org/10.1016/j.specom.2016.01.002>
- Heaton, J. T., Robertson, M., & Griffin, C. (2011). Development of a wireless electromyographically controlled electrolarynx voice prosthesis. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (pp. 5352–5355).
- Hillman, R. E., Walsh, M. J., Wolf, G. T., Fisher, S. G., & Hong, W. K. (1998). Functional outcomes following treatment for advanced laryngeal cancer. Part I – voice preservation in advanced laryngeal cancer. Part II – laryngectomy rehabilitation: The state of the art in the VA system. Research speech-language pathologists. Department of Veterans Affairs Laryngeal Cancer Study Group. *Annals of Otolaryngology, Rhinology & Laryngology, Supplement*, 172, 1–27.
- Koike, M., Kobayashi, N., Hirose, H., & Hara, Y. (2002). Speech rehabilitation after total laryngectomy. *Acta Oto-Laryngologica*, 122(4), 107–112. <https://doi.org/10.1080/000164802760057716>
- Krause, J. C., & Braida, L. D. (2003). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *Journal of the Acoustical Society of America*, 112(5), 2165–2172. <https://doi.org/10.1121/1.1509432>
- Law, I., Ma, E. P.-M., & Yiu, E. M.-L. (2009). Speech intelligibility, acceptability, and communication-related quality of life in Chinese alaryngeal speakers. *Archives of Otolaryngology - Head and Neck Surgery*, 135(7), 704–711.
- Matsui, K., Kimura, K., Nakatoh, Y., & Kato, Y. (2013). Development of electrolarynx with hands-free prosody control. In *Proceedings of the 8th ISCA Speech Synthesis Workshop* (pp. 273–278). Barcelona, Spain: International Speech Communication Association.
- Meltzner, G. S., & Hillman, R. E. (2005). Impact of aberrant acoustic properties on the perception of sound quality in electrolarynx speech. *Journal of Speech, Language & Hearing Research*, 48(4), 766–779.
- Meltzner, G. S., Hillman, R. E., Heaton, J. T., Houston, K. M., Kobler, J. B., & Qi, Y. (2005). Electrolaryngeal speech: The state of the art and future directions for development. In *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing* (pp. 571–590). Austin, TX: Pro-Ed.
- Mendenhall, W. M., Morris, C. G., Stringer, S. P., Amdur, R. J., Hinerman, R. W., Villaret, D. B., & Robbins, K. T. (2002). Voice rehabilitation after total laryngectomy and postoperative radiation therapy. *Journal of Clinical Oncology*, 20(10), 2500–2505. <https://doi.org/10.1200/JCO.2002.07.047>
- Most, T., Tobin, Y., & Mimran, R. C. (2000). Acoustic and perceptual characteristics of esophageal and tracheoesophageal speech production. *Journal of Communication Disorders*, 33(2), 165–181. [https://doi.org/10.1016/S0021-9924\(99\)00030-1](https://doi.org/10.1016/S0021-9924(99)00030-1)
- Nagle, K., & Heaton, J. (2016). Perceived naturalness of electrolaryngeal speech produced using sEMG controlled vs. manual pitch modulation. In *Proceedings of the 17th Annual Conference of the International Speech Communication Association* (pp. 238–242). San Francisco, CA. <https://doi.org/10.21437/Interspeech.2016-1476>
- Nagle, K. F., Eadie, T. L., Wright, D. R., & Sumida, Y. A. (2012). Effect of fundamental frequency on judgments of electrolaryngeal speech. *American Journal of Speech-Language Pathology*, 21(2), 154–166. [https://doi.org/10.1044/1058-0360\(2012\)11-0050](https://doi.org/10.1044/1058-0360(2012)11-0050)
- Nagle, K. F., & Heaton, J. T. (2017). Comparison of thumb-pressure vs. electromyographic modes of frequency modulation for electrolaryngeal speech. Poster presented at the 173rd Meeting of the Acoustical Society of America, Boston.
- Nakamura, K., Toda, T., Saruwatari, H., & Shikano, K. (2012). Speaking-aid systems using GMM-based voice conversion for electrolaryngeal speech. *Speech Communication*, 54, 134–146. <https://doi.org/10.1016/j.specom.2011.07.007>
- Nelson, I., Parkin, J. L., & Potter, J. F. (1975). The modified Tokyo larynx: An improved pneumatic speech aid. *Archives of Otolaryngology*, 101(2), 107–108. <https://doi.org/10.1001/archotol.1975.00780310029008>
- Pepiot, E. (2014). Male and female speech: A study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers. In *Speech Prosody 2014* (pp. 305–309). Dublin, Ireland. <https://doi.org/10.1016/j.prospe.2014.05.032>
- Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *J Acoust Soc Am*, 32, 693–703. <https://doi.org/10.1121/1.1908183>
- Pindzola, R., & Cain, B. (1988). Acceptability ratings of tracheoesophageal speech. *Laryngoscope*, 98(4), 394–397.
- Qi, Y. Y., & Weinberg, B. (1991). Low-frequency energy deficit in electrolaryngeal speech. *Journal of Speech & Hearing Research*, 34(6), 1250–1256.
- Robbins, J., Fisher, H. B., Blom, E. C., & Singer, M. I. (1984). Acoustic differentiation of laryngeal, esophageal, and tracheoesophageal speech. *Journal of Speech & Hearing Research*, 27(4), 577–585.
- Searl, J. (2006). Technological advances in alaryngeal speech rehabilitation. *SIG 3 Perspectives on Voice and Voice Disorders*, 16(July), 12–18. <https://doi.org/10.1044/vvd16.2.12>
- Singer, M. I., & Blom, E. D. (1980). An endoscopic technique for restoration of voice after laryngectomy. *Annals of Otolaryngology, Rhinology & Laryngology*, 89(6), 529–533. <https://doi.org/10.1177/000348948008900608>
- Singer, S., Wollbrück, D., Dietz, A., Schock, J., Pabst, F., Vogel, H.-J., ... Meuret, S. (2013). Speech rehabilitation during the first year after total laryngectomy. *Head & Neck*, 35(11), 1583–1590. <https://doi.org/10.1002/hed.23183>
- Smiljanić, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and*

- Linguistics Compass*, 3(1), 236–264. <https://doi.org/10.1111/j.1749-818X.2008.00112.x>
- Smithwick, L., Davis, P., Dancer, J., Hicks, G., & Montague, J. (2002). Female laryngectomees' satisfaction with communication methods and speech-language pathology services. *Perceptual and Motor Skills*, 94(1), 204–206.
- Štajner-Katušić, S., Horga, D., Mušura, M., & Globlek, D. (2006). Voice and speech after laryngectomy. *Clinical Linguistics & Phonetics*, 20(2–3), 195–203. <https://doi.org/10.1080/02699200400026975>
- Stepp, C. E., Heaton, J. T., Rolland, R. G., & Hillman, R. E. (2009). Neck and face surface electromyography for prosthetic voice control after total laryngectomy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 17(2), 146–155. <https://doi.org/10.1109/TNSRE.2009.2017805>
- Takahashi, H., Nakao, M., Kikuchi, Y., & Kaga, K. (2005). Alaryngeal speech aid using an intra-oral electrolarynx and a miniature fingertip switch. *Auris, Nasus, Larynx*, 32(2), 157–162. <https://doi.org/10.1016/j.anl.2005.01.007>
- Wan, C., Wu, L., Wu, H., Wang, S., & Wan, M. (2012). Assessment of a method for the automatic on/off control of an electrolarynx via lip deformation. *Journal of Voice*, 26(5), 674.e21–674.e30. <https://doi.org/10.1016/j.jvoice.2012.03.002>
- Ward, E. C., Koh, S. K., Frisby, J., & Hodge, R. (2003). Differential modes of alaryngeal communication and long-term voice outcomes following pharyngolaryngectomy and laryngectomy. *Folia Phoniatrica et Logopaedica*, 55(1), 39–49.
- Watson, P. J., & Schlauch, R. S. (2008). The effect of fundamental frequency on the intelligibility of speech with flattened intonation contours. *American Journal of Speech-Language Pathology*, 17(4), 348–355. [https://doi.org/10.1044/1058-0360\(2008/07-0048\)](https://doi.org/10.1044/1058-0360(2008/07-0048)) [pii] 10.1044/1058-0360(2008/07-0048).
- Watson, P. J., & Schlauch, R. S. (2009). Fundamental frequency variation with an electrolarynx improves speech understanding: A case study. *American Journal of Speech-Language Pathology*, 18(2), 162–167.
- Weinberg, B., & Riekens, A. (1973). Speech produced with the Tokyo artificial larynx. *Journal of Speech and Hearing Disorders*, 38(3), 383–389.
- Weiss, M. S., & Basili, A. G. (1985). Electrolaryngeal speech produced by laryngectomized subjects: Perceptual characteristics. *Journal of Speech & Hearing Research*, 28(2), 294–300.
- Weiss, M. S., Yeni-Komshian, G. H., & Heinz, J. M. (1979). Acoustical and perceptual characteristics of speech produced with an electronic artificial larynx. *The Journal of the Acoustical Society of America*, 65(5), 1298–1308. <https://doi.org/10.1121/1.382697>
- Williams, S. E., & Watson, J. B. (1987). Speaking proficiency variations according to method of alaryngeal voicing. *Laryngoscope*, 97(6), 737–739.
- Wu, L., Wan, C., Wang, S., & Wan, M. (2013). Improvement of electrolaryngeal speech quality using a supraglottal voice source with compensation of vocal tract characteristics. *IEEE Transactions on Biomedical Engineering*, 60(7), 1965–1974. <https://doi.org/10.1109/TBME.2013.2246789>
- Wu, L., Wan, C., Xiao, K., Wang, S., & Wan, M. (2014). Evaluation of a method for vowel-specific voice source control of an electrolarynx using visual information. *Speech Communication*, 57, 39–49. <https://doi.org/10.1016/j.specom.2013.09.006>
- Yorkston, K., Strand, E., & Kennedy, M. R. T. (1996). Comprehensibility of dysarthric speech: Implications for assessment and treatment planning. *American Journal of Speech-Language Pathology*, 5(1), 55–66.
- Yorkston, K. M., Hammen, V. L., Beukelman, D. R., & Traynor, C. D. (1990). The effect of rate control on the intelligibility and naturalness of dysarthric speech. *Journal of Speech and Hearing Disorders*, 55(3), 550–560. <https://doi.org/10.1044/jshd.5503.550>