14 Intelligibility in Postlaryngectomy Speech

Lindsay E. Sleeth and Philip C. Doyle

Introduction

A diagnosis of laryngeal cancer has far-reaching effects that will impact all areas of an individual's life including physical, emotional, psychological, economic, and social well-being (Bornbaum & Doyle, Chap. 5; Doyle, 1994, 2005; Doyle & MacDonald, Chap. 27; Eadie & Doyle, 2004, 2005; Meyer et al., 2004). Distinctive to a diagnosis of laryngeal cancer is the potential need to surgically remove the entire larynx leading to the loss of the individual's normal vocal mechanism and, subsequently, a loss of normal verbal communication. While cancer itself carries substantial disease burden, the loss of voice at the time of serious illness will create an added distress for the individual (Bornbaum et al., 2012; Doyle, 1994). Loss of verbal communication at the time of a health crisis is not typically experienced with other sites of cancer. For this reason, changes in verbal communication secondary to treatment for laryngeal cancer have long been of critical importance in postlaryngectomy rehabilitation.

Multiple studies have shown verbal communication to be one of the greatest predictors of quality of life (QOL) in individuals with laryngeal cancer (Eadie & Doyle, 2004; Karnell, Funk, & Hoffman, 2000; Meyer et al., 2004; Terrell et al., 2004; and others). The notion of QOL encompasses the areas of an individual's life within the physical, psychological, social, and spiritual domains of functioning. When expressed using the World Health Organization's (WHO, 2001) International Classification of Functioning, Disability and Health (ICF), issues secondary to the diagnosis and treatment of laryngeal cancer encompass all components of the ICF framework (body functions and structures, activities and participation, environmental factors, and personal factors). Eadie (2003) was the first to contextualize laryngeal cancer within the ICF framework, and she described the dramatic interactions that may emerge in one's postlaryngectomy functioning. Therefore, the ability to effectively restore an individual's verbal communication following removal of the larynx has the ability to positively impact a person's QOL. However, reacquisition of a new alaryngeal method of verbal communication does not in and of itself offer the sole index of postlaryngectomy rehabilitation success.

L. E. Sleeth (✉)
Voice Production and Perception Laboratory, Rehabilitation Sciences, Western University, London, ON, Canada

South West Local Health Integration Network, London, ON, Canada

P. C. Doyle
Voice Production and Perception Laboratory & Laboratory for Well-Being and Quality of Life, Department of Otolaryngology – Head and Neck Surgery and School of Communication Sciences and Disorders, Western University, Elborn College, London, ON, Canada

P. C. Doyle (ed.), *Clinical Care and Rehabilitation in Head and Neck Cancer*,
https://doi.org/10.1007/978-3-030-04702-3_14

Understanding the loss of speech and its restoration through rehabilitative efforts raises numerous questions on the resultant effectiveness of postlaryngectomy communication. If social capacity is to be enhanced in the postlaryngectomy period, it cannot be achieved without at least a "good" level of SI – that is, good speech will not place excessive demands on the listener during communication (Evitts, Chap. 28). SI forms a core element underlying effective communication. For this reason, it is important that continued clinical efforts be directed at assessing and documenting postlaryngectomy speech rehabilitation outcomes. This includes that direct attention is paid to a variety of factors underlying its composite product, namely, intelligibility. Regardless of the method of alaryngeal speech acquired, whether it be the use of the artificial electrolarynx or esophageal or tracheoesophageal (TE) speech, a clinical focus on optimizing SI will form one of the foundational aspects of all head and neck cancer rehabilitation. Consequently, this chapter presents information related to SI with a specific focus on those are undergo total laryngectomy.

The concepts to be addressed herein have broad applications to all modes of postlaryngectomy voice and speech rehabilitation. In many respects today, information on TE speech is much more available in the literature, thus, TE speech has some prominence in the discussion to follow. This prominence is primarily based on the fact that TE speech is widely used; however, an additional factor also exists. That is, while considerable research on intelligibility related to TE speech was conducted in the first 25 years after its introduction (Singer & Blom, 1980), in recent years work in this area has been relatively sparse. It is also of value to note that while the historical literature on alaryngeal SI addressed comparative performance between methods (i.e., electrolaryngeal, esophageal, and TE), more recent comparative data are lacking. Thus, information on SI in postlaryngectomy speakers is the specific focus of subsequent sections of this chapter.

Postlaryngectomy Voice and Speech Rehabilitation

Alaryngeal Speech

When a total laryngectomy is required, an alternate method of postlaryngectomy "alaryngeal" voice and speech will need to be learned. Without doing so, the individual will be unable to communicate verbally and will be required to use writing or alternative or augmentative methods of communication (Childs, Palmer, & Fried-Oken, Chap. 15). At present, there are three primary methods of alaryngeal speech employed by laryngectomized individuals: (1) use of an artificial electrolarynx, (2) esophageal speech, and (3) tracheoesophageal (TE) speech. While multiple methods may be used by some speakers (e.g., use of esophageal speech and the electrolarynx), one method will almost certainly be identified by the individual as being their primary method of communication.

In 1980, the tracheoesophageal (TE) puncture voice restoration method and the first TE puncture voice prosthesis (Singer & Blom, 1980) was introduced as a new alaryngeal speech option. Briefly, the procedure involves creating a small, controlled midline puncture through the posterior wall of the trachea into the esophagus (Singer & Blom, 1980). A one-way, valved voice prosthesis is then inserted into the puncture to prevent closure of the site and to allow one-way flow of air from the trachea into the esophageal reservoir below the pharyngoesophageal (PE) segment (Blom, 1998). Upon exhalation, and when the tracheostoma is occluded by the individual's thumb or another "hands-free" device, pulmonary air is shunted into the esophagus, setting the PE segment into vibration and allowing for sound generation. Thus, while the alaryngeal tissue source is the same for both the esophageal and TE speech methods (the PE segment), it is the manner in which the system is placed into vibration and the amount of air available to continuously modulate that tissue prior to re-insufflation of the esophageal reservoir that distinguishes these two methods.

Since its introduction, TE voice restoration has become widely used as a postlaryngectomy speech rehabilitation method. In the early years following its introduction, the puncture was completed as a secondary procedure, that is, at some point following laryngectomy and full healing and postsurgical recovery. However, in the years to follow, use of the method as a primary procedure performed at the same time as the laryngectomy was increasingly pursued (Kao, Mohr, Kimmel, Getch, & Silverman, 1994; Singer, Blom, & Hamaker, 1983; Yoshida, Hamaker, Singer, Blom, & Charles, 1989). The larger influence, impact, and clinical implications of these approaches on postlaryngectomy voice and speech rehabilitation are addressed in greater detail elsewhere in this volume (see Graville, Palmer, & Bolognone, Chap. 11; Knott, Chap. 12). Consequently, in the section to follow, factors that may influence postlaryngectomy SI and the unique relationship of these factors to specific alaryngeal methods will be outlined.

Factors Influencing Speech Intelligibility: Preliminary Issues

SI is influenced by multiple factors. Normal SI is a result of a complex and highly coordinated interaction of physiologic systems under finely tuned neurological control. In the normal speech production system, intelligibility will be influenced by the power supply or driving source (the lungs), the vibratory element (the vocal folds), and a system of valves and filters (structures of the vocal tract including the oral cavity and its structures). The interaction of these systems provides for a maximal degree of flexibility that permits a wide range of acoustic changes which cross the frequency, intensity, and temporal domains. Postlaryngectomy voice and speech production will, therefore, present with alterations in the nature and interaction of all of these systems.

A breakdown in one component of the speech production system may create changes both upstream and downstream, with a net result on the final speech product. However, the loss of one's natural voice and decreases in the understandability of a new voicing source and the speech produced will result in substantial psychosocial changes. The impact of such changes in SI extend beyond the communication process itself to have a broader, negative influence on perceived QOL (Meyer et al., 2004). Yet each alaryngeal method needs to be considered independently in an effort to further understand the more refined aspects of why SI decreases.

Electrolaryngeal Speech For the electrolaryngeal speaker, the power supply and voicing source is now non-biologic (electronic) and external; this signal will be directed into the vocal tract (either via transcervical or intraoral application) where the sound source will be articulated into speech (see Nagle, Chap. 9). The electrolaryngeal voice source will be modified by the method of its transmission through neck tissues or through direct introduction into the oral cavity. The use of an electrolarynx voicing source also will be characterized by a relatively narrow range of frequencies (Nagle, Eadie, Wright, & Sumida, 2012), as well as a continuous "all voiced" signal source. Because of this continuous sound activation, perceptual challenges related to the listener's ability to make distinctions between voiced and voiceless cognate sounds (e.g., /p/ vs. /b/) will be observed (Weiss & Basili, 1985). Further, the unique nature of the electrolarynx with its robotic and monotone quality will influence the listener's perception of speech due to concerns related to its overall acceptability (Bennett & Weinberg, 1973). Thus, the interaction and impact of all three components of the physical analog system of speech production – the power supply, the voicing source, and the valves and filters – must be considered collectively in seeking to understand the intelligibility of the signal.

Esophageal speech For an esophageal speaker, the esophagus now becomes the driving source, but the capacity of this reservoir is limited, and it must be regularly replenished with air in order to continue producing speech (see Doyle & Finchem, Chap. 10). Additionally, esophageal speech will be

generated by an anatomical voicing structure comprised of lower pharyngeal and upper esophageal tissues (the PE segment), one which does not have active adductory or abductory capabilities. Similar to the use of an electrolarynx, voiced-voiceless distinctions may be problematic because the PE segment cannot rapidly turn on and off; however, in esophageal speech, this also may be the result of limitations in both the power supply and the speaker's subsequent inability to generate adequate sound intensity, as well as the "all voiced" nature of the esophageal sound source (Christensen, Weinberg, & Alphonso, 1978; Connor, Hamlet, & Joyce, 1985). If the esophageal vibratory source is not fully powered by the air passing through it, a result of limited access to air within the esophagus that drives this tissue and then tissue oscillation for voicing will be incomplete and of reduced temporal duration. The reduced amplitude (intensity) of the esophageal voicing signal also has been shown to have a direct influence on phonetic quality (Blood, 1981). Again, interactions between the three analog systems of speech production cannot be underestimated in the context of understanding reductions in esophageal SI.

TE Speech Finally, the typical TE speaker has the capacity to exploit a very large volume of lung air to replenish the esophageal reservoir in a relatively continuous manner (see Bohnenkamp, Chap. 7; Searl, Chap. 13). Because the TE speaker has access to a pulmonary air supply, research data have most prominently documented increases in overall speech and syllable rates for TE speech when compared to esophageal speakers. In fact, word and syllable per minute measures for TE speakers have been shown to be at values approximating those of normal laryngeal speakers (Robbins, Fisher, Blom, & Singer, 1984). TE speakers typically are able to produce conversational speech without any unusual breaks in the flow of their speech.[1] This is in clear contrast to even the most proficient esophageal speakers who will exhibit momentary stoppages in the flow of speech in order to re-insufflate the esophageal reservoir (Snidecor & Curry, 1960).

Further, because of the increased volume and pressure of air moving through the TE voice prosthesis into the esophageal reservoir, and the subsequent propagation of this signal into the vocal tract, the speaker may have the ability to produce voiceless sounds despite the continuous vibration of the PE voicing source (Doyle, Danhauer, & Reed, 1988). Doyle and colleagues hypothesized that this ability was secondary to the exploitation of air pressure within the upper vocal tract during TE speech production; that is, it was believed that some level of vocal tract turbulence[2] could be achieved during the TE speech process by manipulating the signal within the oral cavity during articulation. Results from studies by Doyle et al. (1988) and Searl and colleagues (Searl & Carpenter, 2002; Searl, Carpenter, & Banta, 2001) have shown that intelligibility issues commonly arise in the areas of voiced-voiceless distinctions of consonants (Gomyo & Doyle, 1989), as well as for the general consonant manner classes of stops, fricatives, and affricates.

The issue surrounding the voiced-voiceless distinction in alaryngeal speech in general (Jongmans, Hilgers, Pols, & van As-Brooks, 2006) involves confusing voiceless phonemes for voiced phonemes (e.g., perception of a /b/ when its voiceless cognate /p/ was intended). Doyle et al. (1988) hypothesized this to be a result of shortened voice onset time (VOT) in TE speakers, as well as a lag in the termination of voicing by the PE segment (Robbins, Christensen, & Kempster, 1986). As well, a study conducted by Doyle and Haaf (1989) found postvocalic consonants to be more intelligible than their prevocalic counterparts, suggesting that onset and offset phenomena must be considered. Doyle and Haaf (1989) also found voiced-voiceless confusions

[1]Electrolaryngeal speakers will also have the ability to generate a full phrase length that approximates that observed for the normal speaker. However, the signal itself may be monotone, and its mechanical quality may to some extent distract the listener.

[2]If air pressures and flows are of sufficient magnitude within the vocal tract, the speaker may be able to further compress that air during the act of articulation. If this does in fact occur, the compression of this air may be perceived by the listener as a voiceless sound despite the fact that the PE source has provided an energized, voiced signal.

and a manner of production intelligibility hierarchy similar to that found by Doyle et al. (1988). More recently, Searl et al. (2001) evaluated the intelligibility of stops and fricatives in TE speech; their findings were consistent with the two studies previously mentioned (Doyle et al., 1998; Doyle & Haaf, 1989) in that the most common errors emerged from the listeners' confusion of voiced for voiceless phonemes.

Speech Intelligibility

Kent, Weismer, Kent, and Rosenbek (1989) have defined SI as "the degree to which the speaker's intended message is recovered by the listener" (p. 483). Employing a more procedural definition, Hillman, Walsh, and Heaton (2005) have indicated that SI represents the percentage of speech items correctly identified by the listener. Regardless of the underlying etiology, reductions in SI have been a critical and longstanding concern in the area of speech disorders. When there is a breakdown in a speaker's ability to be easily understood by his or her communicative partner, many challenges will be experienced at multiple levels of communication functioning (Ackerstaff, Hilgers, Aaronson, & Balm, 1994). Reductions in SI will result in increased burden to both the speaker and the listener with a direct impact on social well-being. This concern is of particular importance to those who have undergone total laryngectomy and will be trained to use any alternative alaryngeal method of verbal communication (Doyle, 1994). However, numerous factors will influence any measure of SI.

Over the years, the intelligibility of alaryngeal speech has been studied by numerous researchers with varying populations of speakers, under a variety of conditions, and with a range of assessment of stimuli (Amster et al., 1972; Bridges, 1991; Clark & Stemple, 1982; Doyle et al., 1988; Filter & Hyman, 1975; Hillman, Walsh, Wolf, Fisher, & Hong, 1998; Hyman, 1955; Kalb & Carpenter, 1981; Miralles & Cervera, 1995; Tardy-Mitzell, Andrews, & Bowman, 1985; Weiss & Basili, 1985; and others). This includes multiple studies that have addressed individual methods of alaryngeal speech (esophageal, electrolaryngeal, and TE). A number of studies also have evaluated SI from a comparative perspective. The collective results of these studies have shown that SI has the potential to increase or decrease based on a range of factors such as the experience and training of speakers, the experience of the listeners, the type of stimuli, background noise and environmental conditions, speaker gender, type of postlaryngectomy speech mode, etc. (McColl, Fucci, Petrosino, Martin, & McCaffrey, 1998). Therefore, when considering the findings and implications of SI research in postlaryngectomy populations, it is important to consider the potential interaction of multiple factors (Yorkston & Beukelman, 1980). Doing so will allow one to contextualize the results of intelligibility testing in those who use alaryngeal speech. Interestingly, however, over the past 20 years, there has been an increasing paucity of information specific to SI issues that characterize postlaryngectomy speakers.

Interrelationships in Alaryngeal Speech Production

The previously outlined analog system and the inherent differences specific to each alaryngeal method will have a direct impact on SI. Perhaps the only level where intelligibility concerns do not exist in any substantial manner for postlaryngectomy speakers is related to vowel production.[3] The reason for this is that regardless of alaryngeal mode, the new voice source will always be voiced which is a fundamental requirement of vowels. However, changes in the amplitude and duration of vowels may impact the accuracy of their perception by the listener. For example, if the vowel signal is underpowered (e.g., during esophageal speech), it may carry insufficient information relative to its formant structure; similarly, if the temporal duration of the vowel is altered in either direction (reduced or extended),

[3]The exception here would be any laryngectomy procedure that includes the removal of any portion of the tongue or changes that occur secondary to resections at the base of the tongue.

intelligibility changes may occur (Sisty & Weinberg, 1972). Thus, the flow and continuity of TE speech that occurs due to access to pulmonary air has been shown to offer considerable perceptual advantages to listeners. When temporal components of alaryngeal speech production are optimized, durational aspects of sound production will also benefit (Doyle et al., 1988; Searl & Carpenter, 2002; Searl et al., 2001; and others).

While temporal speech advantages are well documented in the literature, TE speakers will also find some advantage in the frequency (pitch) and intensity (loudness) domains of the speech produced; however, these changes do in fact vary considerable from normal expectation. First, because of the TE speaker's access to a substantial volume of air from the lungs, the ability to "drive" pulmonary air through the PE segment at a greater pressure and rate of flow will result in the creation of an increased "duty cycle" specific to the vibratory source. As the duty cycle or rate of tissue vibration increases, the perceived pitch of TE voice will also be greater relative to esophageal speech despite use of the same vibratory (source tissues of the PE segment) for both methods. Even though the fundamental frequency of TE speech exceeds that of esophageal speakers, it will remain reduced from normal expectation. For this reason, gender considerations related to frequency for both esophageal and TE speakers must be considered (Bellandese, Lerman, & Gilbert, 2001; Eadie, Doyle, Hansen, & Beaudin, 2008). Secondly, the relative intensity of TE speech also will be increased because of greater short-term volumes of air that are available to the speaker.[4]

As a fundamental factor, the TE speaker's access to a large volume driving source (the lungs) to power the esophagus does have considerable advantages across a variety of acoustic dimensions (Baggs & Pine, 1983; Qi & Weinberg, 1995; Robbins et al., 1984). The TE system is also able to be actively exploited by muscles of respiration which may then serve to fine-tune airflows through the vibratory source (Bohnenkamp, Chap. 7; Bohnenkamp, Forrest, Klaben, & Stager, 2012). Research has revealed that an esophageal speaker produces voice that is 10 dB-SPL lower for sustained vowels than that of a normal speaker (Weinberg, Horii, & Smith, 1980); this finding is a direct consequence of esophageal speakers not having access to vast volumes of air to drive the PE segment. Recall that an esophageal speaker will need to recharge the esophageal reservoir regularly as its volumetric capacity is only in the range of 60 cc (Diedrich, 1968).

Interestingly, Robbins et al. (1984) found that TE speakers produce voicing signals that are approximately 10 dB-SPL and 10 dB-A greater than normal speakers for vowels and conversational speech, respectively. From a simple acoustic perspective, this increase of "above normal" loudness may be interpreted as advantageous, yet it does carry with it limitations relative to real-life conversational interaction. More specifically, the increased vocal loudness associated with TE speech may limit a speaker's ability to maintain privacy when communicating with others. For that reason, a clinician's comprehensive knowledge related to the process of TE speech production, dynamic interactions between the new vibratory source and the vocal tract, and the resultant acoustic consequences (both negative and positive) are essential components underlying one's understanding of changes in SI. As stated by Weinberg, Horii, Blom, and Singer (1982, p. 1982) in relation to observed differences between esophageal and TE speakers, "*Since the voicing sources used to produce esophageal voices are regarded as surgical residue (Weinberg, 1980), we suggest that the operation of these residue sources has been maximized or optimized by alterations in respiratory drive state.*" Thus, clinicians must acknowledge that isolated components of the new alaryngeal speech process, regardless of mode, are likely to have unique consequences

[4]Although it is beyond the scope of the present chapter, increases in vocal loudness whether in a normal or esophageal-based alaryngeal system are directly correlated with one's ability to increase pressure below the point of vibration. Thus, the ability to modulate a relatively large volume of pulmonary air during TE voicing provides the speaker with an increase potential for achieving a wider range of vocal intensities.

that may directly influence SI. These changes may most often be detected in specific ways dependent upon the manner in which SI is evaluated (i.e., are stimuli comprised of words, sentences, etc.).

Does Alaryngeal Voice Quality Influence Intelligibility?

Despite the fact that the TE voicing process may mimic normal acoustic values for some dimensions (e.g., speech rate, increased pitch levels, etc.), it is critical to note that the "optimization" of voice/speech that Weinberg et al. (1980) identified *does not* result in a normal voice signal. All alaryngeal voices will be identified as being abnormal in regard to the overall perceived quality of the voice signal (Eadie & Doyle, 2002, 2005; Nagle & Eadie, 2012; McDonald, et al., 2010). And, data would suggest that the quality of the signal is the most significant factor relative to how a listener may judge the proficiency of a given speaker. Additionally, alaryngeal SI and the potential consequences that noise features which are simultaneously present in the signal may be quite variable both within and across alaryngeal methods. While clear distinctions in quality and performance will be observed between extrinsic (electrolaryngeal) and intrinsic (esophageal and TE) speech methods, it is important to note that each speech mode is highly variable. Even for individual speakers, ongoing signal variability will result in greater perceptual challenges for the listener who must work more to extract speech from noise (Doyle, 2017a). The importance of individual speaker variability across modes of speech was elegantly reported by Kalb and Carpenter (1981), and their work continues to hold merit today.

Even the most proficient esophageal or TE speaker will exhibit a voice quality that is almost certainly characterized by aperiodicity as part of the composite nature of the signal (Maryn, Dick, Vandenbruaene, Vauterin, & Jacobs, 2009; Smith, Weinberg, Feth, & Horii, 1978). This aperiodicity is a product of the variability of the tissue that comprises the PE segment (differences in location and mass) and its subsequent response to airflows and pressures, factors that carry their own inherent degree(s) of variability (Doyle & Eadie, 2005; Moon & Weinberg, 1987). Thus, expectations of what any given alaryngeal speaker will sound like or how their intelligibility will manifest must be made with caution. Finally, the literature is rich with studies comparing a variety of features of the three alaryngeal methods, both to each other and to normal speech (Blom, Singer, & Hamaker, 1986; Clements, Rassekh, Seikaly, Hokanson, & Calhoun, 1997; Cullinan, Brown, & Blalock, 1986; Doyle et al., 1988; Robbins, 1984; Robbins et al., 1984; Tardy-Mitzell et al., 1985). When viewed together, the findings from these studies have often found TE speech to be judged as superior to the other alaryngeal modes in areas such as acceptability, overall intelligibility, pitch, intensity, and patient satisfaction, with ratings approaching those of normal speech in some instances. Nevertheless, it is also of importance to note that TE speech is not without substantial limitations.

As mentioned previously, esophageal and TE speakers will produce speech that is judged by listeners to be of lowered pitch. In fact, female esophageal and TE speakers tend to have pitch values similar to those of males, resulting in a voice that sounds more masculine (Bellandese et al., 2001; Trudeau, 1994). The loss of gender identity for a woman who undergoes total laryngectomy and uses alaryngeal speech may have a significantly negative impact on social performance and interaction and ultimately, one's judgment of their QOL. In addition, although TE speech often has shown to be highly acceptable when compared to other methods, it is clearly judged as less acceptable than laryngeal speech (Clark & Stemple, 1982; Finizia, Dotevall, Lundstrom, & Lindstrom, 1999; van As, Hilgers, Verdonck-de Leeuw, & Koopmans-van Beinam, 1998). Finally, even in the presence of several comparative advantages, TE speech has consistently been reported to be reduced in its intelligibility (Blom et al., 1986; Doyle et al., 1988; Pindzola & Cain, 1988; Robbins, 1984; Williams & Watson, 1985; and others). Thus, no one alaryngeal method is free of limitations; as a general rule, there will always be specific advantages and disadvantages to each method.

Clinical efforts that focus directly on increasing SI beyond basic "functional" levels are often disregarded in the contemporary rehabilitation setting. This is often a result of the cost associated with extra treatment sessions, limitations in the time available to skilled personnel, and sometimes, limitations related to general access to high quality and comprehensive alaryngeal speech rehabilitation services. SI is an area of inquiry that received generous attention when TE speech was first introduced, but unfortunately, it has been somewhat overlooked for the past 20 years. This in part may be a consequence of the fact that unless some unexpected complication occurs with TE voice restoration and its acquisition, most individuals will quickly acquire their new voice. Because voice and speech restoration in many instances is reacquired rather rapidly, efforts directed at refining speech may be pursued less often. It is, however, essential to reiterate that reduced SI has the potential to negatively impact a person's participation in society and it should remain an area of continued interest and exploration (Eadie et al., 2016). This problem is also applicable to those who use esophageal and electrolaryngeal methods.

Measuring Speech Intelligibility

As noted, a wide range of factors will potentially impact findings from SI assessments.[5] It is, therefore, important to understand and consider these factors before pursuing the clinical evaluation of alaryngeal speech, as well as when conducting research in this area. As stated by Subtelny (1977, p. 183) "Intelligibility is considered the most practical single index to apply in assessing competence in oral communication." Throughout history, intelligibility measurement has largely been obtained through two separate methods: scaling procedures and word identification (Schiavetti, 1992). Scaling procedures, such as the use of equal-appearing interval (EAI) scales which allow the listener to make judgments about a speaker's intelligibility, were historically used more frequently due to their ease of application and scoring (Schiavetti, 1992).

Scaling Procedures Briefly, when intelligibility is assessed using the EAI scaling method, the listener judges intelligibility by selecting a discrete number that falls between two extreme anchors that represent the range of potential performance (e.g., "fully understandable" speech to "unable to understand"). Scales may range in numerical representation, for example, from 1 to 5 or from 1 to 7 or greater. The key to understanding the EAI method is that it ultimately asks the listener to assign a numeric rating that best represents their impression of where on the scale a speaker falls given the anchors provided. The simplicity of this type of rating task is beneficial, but it also allows for error on several fronts.

Recently, as intelligibility testing has continued to grow in many disordered speech populations, perceptual scaling procedures have received considerable attention and criticism. Although timely and efficient, scaling procedures often lack the ability to pinpoint specific areas of increased or decreased intelligibility. Intelligibility measures and the findings gathered from their application are, at times, also subject to misinterpretation and misrepresentation. As such, any approach to scaled assessments using the EAI method may have limited strength in accurately estimating an intelligibility score for each individual without also obtaining percentage values for the accurate retrieval of stimuli produced (Schiavetti, 1992). This suggests that in most instances, single approaches to evaluating SI must carry one or more caveats related to interpretation and use of the data.

A categorical judgment of intelligibility based on EAI scaling, regardless of the number that is selected from a given scale, may ultimately present a considerable range of performance. In considering any EAI scale, the question that often arises is that pertaining to, "what distinguishes or separates one number from the next?" Using an EAI scale, the question of what lies between any

[5]It should be noted that the factors which have been identified to influence measures of speech intelligibility apply to all communication disorders, not just in relationship to postlaryngectomy speakers.

two numbers on the scale in the context of the anchors provided is not easily discerned (see Stevens, 1975). Further, any scaled numeric assignment may not be consistent with measures obtained at the "word identification level." This potential problem in turn decreases the generalizability of findings gathered when using EAI scales to other studies conducted on intelligibility, as well as making it more difficult for both lay listeners and those who are experienced to rectify differences in their judgments. Thus, word identification testing procedures have increasingly become the method of choice when conducting intelligibility assessments and/or research, especially for alaryngeal speech. However, with only one exception (Weiss & Basili, 1985), there have not been specific measures developed for use with alaryngeal speakers. Given the types of factors identified herein that have the potential to influence speech production, it would seem that some consideration of developing specific types of evaluation instruments for alaryngeal speakers would be of value.

Direct Identification of Stimuli With application of the direct identification assessment method, listeners are required to transcribe each word, sentence, or phrase uttered by the speaker. Listener responses are then compared to the list of target stimuli produced by the speakers; these data are then subsequently converted into a percentage of incorrect and correct responses, resulting in an overall intelligibility score (Schiavetti, 1992. This measurement method has the clear advantage of being easily interpretable to not only clinicians but also naïve individuals and, perhaps more importantly, in conveying such information to those who use alaryngeal speech. If the stimuli and measurement procedure used are assessed in a consistent manner, changes over time or subsequent to therapy can be easily documented. Lastly, the measure is "objective" in nature which offers the potential for identification of what specific type(s) of intelligibility deficit(s) exist for each individual (Schiavetti, 1992). If, for example, an objective measure of intelligibility indicates that stop-plosive consonants are problematic (Gomyo & Doyle, 1989; Doyle & Haaf, 1989; Searl et al., 2001), then efforts to remedy those deficits can be actively pursued. It is, however, important to acknowledge that despite the objective nature of such measures, the score obtained will always be contextually bound and may not easily be generalized to other types of stimuli or evaluation settings.

The determination of the loci or "where" intelligibility deficits exist (e.g., word-initial vs. word-final phonemes, relationships to vocalic elements, etc.) also must be considered with great care given the number of factors that can influence intelligibility judgments. Recent work by Doyle (2017b, unpublished data) has suggested that objective intelligibility scores can vary widely depending upon the construction of the test stimuli used, even if the word list is well-established and regularly used at the clinical level for intelligibility assessment. Nevertheless, the assumed sensitivity of word identification procedures and the ability to gain information solely from such measures has made this approach an obvious choice for many intelligibility investigations (Blom et al., 1986; Doyle et al., 1988; Pindzola & Cain, 1988; Smith & Calhoun, 1994; Tardy-Mitzell et al., 1985; and others).

Listener Experience Another area of intelligibility testing that may impact findings pertains to the influence of listener experience on the SI results obtained. Previous studies have employed the use of either naïve listeners (no prior educational experience with or formal exposure to the speaker population of interest) or experienced listeners typically, speech-language pathologists (SLPs) or physician/surgeons. Multiple studies with a variety of speaker populations have shown that intelligibility may be influenced by the sophistication of listeners (Beukelman & Yorkston, 1980; Doyle, Swift, & Haaf, 1989; Williams & Watson, 1985). These studies all suggest that assessment scores provided by SLPs reflect better speaker intelligibility than those made by inexperienced (naïve) listeners. Based on this observation, it has been suggested that the experienced listeners' prior exposure to the speaker population, and most likely the stimuli being evaluated as well, may potentially inflate

their intelligibility scores. This makes the information less generalizable and possibly, less representative of the general listening population, a listener group that may have the most interaction with the speaker (Doyle et al., 1989). Yet the use of naïve listeners can influence findings as well.

First, since naïve listeners typically have had little exposure to alaryngeal speech (or other disorders for that matter), they may focus on the unnatural quality of the voice instead of the words or sounds being produced, a potentially confounding factor in the interpretation of the data. The quality of the vocal signal may in some way distract the listener from the stimuli they have been asked to assess. While all potential sources of distraction which may confound "pure" assessments of intelligibility cannot reasonably be excluded, recognition of the potential influence of this factor on listener judgments is of value. This issue has been raised in prior auditory-perceptual works associated with the dysarthrias (Darley, Aronson, & Brown, 1969, and others).

As well, naïve listeners may not be challenged by the task itself, and an internal desire to perform the task accurately, leading to "second guessing" confusions or errors, rather than lack of speaker intelligibility. Hence, the demands of the task that the listener is asked to perform must always be considered. This requires that efforts directed toward assessing intelligibility in alaryngeal speakers must weigh numerous factors and understand the potential strengths and weaknesses of any given approach to assessment. Within the dysarthria literature, the listener's familiarity with the stimuli has also been raised (Beukelman & Yorkston, 1980; Tjaden & Liss, 1995; Yorkston & Beukelman, 1978). There is not, however, a perfect method for intelligibility assessment in those who have undergone laryngectomy; similarly, and as noted there is no measure dedicated to the assessment of postlaryngectomy speech regardless of mode. Thus, when conducting intelligibility research or evaluating the validity of previous research on alaryngeal speakers, it is important to consider external factors that have the potential to influence results. In reference back to prior sections of this chapter, intelligibility assessments of alaryngeal speech must also carefully consider the nuances of each method and the potential impact that such alterations will have on measures obtained.

Findings from Alaryngeal Speech Intelligibility Research

Many studies have compared speaker performance across the three modes of alaryngeal speech, and results have indicated that TE speech is generally judged to be more intelligible than esophageal or electrolaryngeal speech (Blom et al., 1986; Doyle et al., 1988; Pindzola & Cain, 1988; Robbins, 1984; Robbins et al., 1984; Tardy-Mitzell et al., 1985; Williams & Watson, 1985). Blom et al. (1986) conducted a study assessing the intelligibility of individuals both before undergoing the TE puncture procedure and after. Prior to the procedure, these speakers were using either esophageal speech or an electrolarynx as their primary mode of communication. Following the TE voice restoration procedure, all individuals used TE speech to communicate. Intelligibility was determined by calculating the percentage of correct responses found using a multiple-choice response format test, the *Modified Rhyme Test* (House, Williams, Hecker, & Kryter, 1965).

Blom et al. (1986) reported a statistically significant improvement in the SI by the group following TE puncture, with preoperative mean intelligibility reported to be 78.15%, versus 91.51% mean intelligibility postoperatively (Blom et al., 1986). This not only illustrates the high intelligibility levels of those using TE speech but also the *potential advantages* of TE speech when compared to esophageal and artificial electrolaryngeal communication. In this context, it is important to note that the use of a forced-choice, closed set identification auditory-perceptual paradigm may influence results, thereby leading to higher intelligibility scores. That is, because the content of stimuli included in the test list was designed to assess particular types of perceptual errors, as well as the request for listeners to select a choice from a set of

perceptual options, some "chance" occurrence of a correct response even when the signal has not been accurately detected may occur. This requires careful consideration of the data obtained (some margin of error must be acknowledged), and this will place greater importance on simultaneous assessments of within-listener agreement when they are asked to rate a subset of stimuli a second time.

Despite the early reports of Blom et al. (1986), TE speech is still less intelligible than speech produced by an individual with an intact larynx (Hillman et al., 2005). Studies have reported TE SI to range from 65% to 93% (Doyle et al., 1988; Pindzola & Cain, 1988; Tardy-Mitzell et al., 1985) dependent upon the procedures employed. Doyle et al. (1988) determined intelligibility through the assessment of consonant-vowel-consonant-vowel-consonant (CVCVC) nonsense syllables that were phonetically transcribed by naïve listeners using an open-response paradigm. This resulted in an average intelligibility of 65% (range 59–72%). The use of nonsense constructions as stimuli as well as the use of an open set response format clearly provided a more challenging task to the listeners, which in turn may have resulted in lower scores. However, while this more restrictive assessment process may reveal poorer intelligibility scores, it is in fact context stripped and may not accurately represent how a listener might perceive stimuli within a conversation or similar interaction with the speaker. In those circumstances, the listener can utilize context (grammatical) and employ what is equivalent to a Cloz procedure as part of a predictive process (Duffy & Giolas, 1974; Epstein, Giolas, & Owens, 1968; Giolas, Cooker, & Duffy, 1970). Thus, the choice of stimuli used to assess intelligibility is critical.

Pindzola and Cain (1988) used an entirely different method of intelligibility assessment by asking TE speakers to record monosyllabic English words from the *Multiple Choice Intelligibility Test* (Black & Haagen, 1963). Naïve listeners then identified their response a set of four options using a forced-choice paradigm. Their study reported an overall intelligibility of 93.20% across speakers. Tardy-Mitzell et al. (1985) used a method similar to that of Pindzola and Cain (1988) with intelligibility judged from monosyllabic word lists (House et al., 1965). Once again, this study employed the forced-choice method with six possible response options for each stimulus word. Comparable intelligibility values were found with an average score of 93% (range 80.70–97.50%). As demonstrated in the above-cited studies, intelligibility has the potential to vary considerably based on internal and external factors and experimental design (stimuli, response format, listener familiarity, and context).

Continuing and Emerging Issues

Much of the research regarding TE SI was conducted in the mid- to late 1980s when the voice restoration procedure was emerging. As a new postlaryngectomy speech rehabilitation option, comparative data were necessary to assess the potential value and viability of this approach. Yet as TE puncture voice restoration became more popular, explorations of TE, as well as esophageal and electrolaryngeal SI, became less common. Since that time, however, very few new investigations have been conducted in relation to the intelligibility of alaryngeal speech. Over the past 20 years, many changes in the treatment of laryngeal cancer have occurred, improvements have been made to the design of TE puncture voice prostheses, esophageal speech may be more easily learned because of knowledge gained from TE speech failures (i.e., the identification of PE segment spasm), and refinements have been made to a several electrolaryngeal devices.

For the reasons noted, generalizing prior intelligibility research to the current generation of alaryngeal speakers is somewhat precarious. As well, a range of hands-free devices have been made available in recent years, removing the need for manual occlusion of the stoma when speaking (Lewis, Chap. 8; Graville, Palmer, & Bolognone, Chap. 11). These devices may differentially influence listener assessments of both SI and overall proficiency (Pauloski, Fisher, Kempster, & Blom, 1989) and are deserving of

ongoing assessment. Thus, SI remains an important index of postlaryngectomy rehabilitation, and continuing its exploration is recommended.

Another factor that cannot be disregarded relative to intelligibility assessment is the fact that some who will undergo laryngectomy today may be performed following failed chemoradiation therapy. This, as well as other treatment-related factors such as the presence of postlaryngectomy complications (Damrose & Doyle, Chap. 3) and concomitant health comorbidities, may have a direct bearing on speech outcomes. In the current era of head and neck cancer surgery in general, and laryngectomy in specific, the use of more extensive reconstruction methods is increasingly common. As a result, the system that will be utilized for the production of any method of alaryngeal speech may be quite different than what has been reported in the past. It is our belief that continued explorations into SI provide not only a valuable area of clinical inquiry but one that will serve to better educate patients and provide the best opportunity for postlaryngectomy rehabilitation success.

At present, very little new research has been conducted with a focus on specific patterns of increased and decreased intelligibility that may exist across groups of alaryngeal speakers. Individual differences and the potential influence of very aperiodic or unusual quality signals can also serve to distract the listener's attention with a subsequent impact on intelligibility. The lack of current intelligibility research can be attributed, at least in part, to the wide use and relatively spontaneous acquisition of TE speech following puncture. As previously stated, by strict standards, TE speech has shown to be superior to the other alaryngeal methods in relation to the "fluent" nature of the speech produced, its increased overall acceptability, in addition to increased overall intelligibility, mean syllable length, pitch, intensity, and patient satisfaction (Blom et al., 1986; Clements et al., 1997; Cullinan et al., 1986; Doyle et al., 1988; Robbins, 1984; Robbins et al., 1984; Tardy-Mitzell et al., 1985). All of these factors contribute to a belief that intelligibility is relatively intact in all TE speakers. However, with the emerging potential that esophageal speech is becoming much more viable as a non-prosthetic mode of alaryngeal speech, in addition to a number of refinements to electrolaryngeal devices (e.g., active frequency modulation), further research appears necessary.

In summary, it appears that the study of SI associated with alaryngeal speech has been overlooked to some extent in recent years. Accordingly, work specific to this important clinical area must be reignited. This inattention is unfortunate as the dissemination of information regarding intelligibility from the past may limit accurate representations at present. Of particular importance here is a concern that if faulty expectations of intelligibility are made, SLPs may limit their efforts to directly facilitate improvements in intelligibility. Regardless of which speech option any individual pursues, and despite the fact that no measurement "standard" currently exists, the formal assessment of intelligibility may provide the SLP with information that guides their ability to tailor individualized therapy (Christensen & Dwyer, 1990). This may involve tasks that center around targeting known error patterns (Doyle, Danhauer, & Lucks-Mendel, 1990) or contextual influences and the value of communication compensation and adaptations, with the result potentially creating more intelligible speech for each individual. The resultant increase in intelligibility has obvious clinical implications, as well as the potential to influence one's ability to fully participate in a variety of communication situations with benefit to perceived QOL.

Significance of Alaryngeal Speech Intelligibility Research

In past years, research has been conducted showing the potential impact speech and effective verbal communication can have on an individual with laryngeal cancer's QOL (Eadie & Doyle, 2004; Karnell et al., 2000; Meyer et al., 2004; Terrell et al., 2004).

The concept of QOL plays an important and prominent role in laryngeal cancer, particularly in relation to the loss of normal verbal communication. Past research has shown speech

communication to be one of the most important predictors of perceived QOL in individuals with cancers of the head and neck (Terrell et al., 2004). A study conducted by Meyer et al. (2004) looked at the importance of effective communication in head and neck cancer survivors and found that decreased word intelligibility was statistically associated with decreases in survivors' enjoyment across many areas of functioning. This decreased ability to participate in normal daily activities increases the potential for disability among these individuals. Lower SI was also associated with a greater likelihood of altered QOL when compared to their more intelligible counterparts. Karnell et al. (2000) evaluated head and neck cancer survivors and found that speech and eating domains best predicted self-reported QOL scores, further reinforcing the importance of postlaryngectomy speech rehabilitation.

Finally, previous research has shown TE SI and acceptability to be positively correlated with one another, indicating that speech that is highly intelligible also tends to be perceived as highly acceptable to listeners (Pindzola & Cain, 1988). Therefore, highly intelligible speakers are not only more likely to be better understood but better accepted by the general public, in turn leading to a potentially increased QOL. The evidence presented in the studies above show that a relationship between highly intelligible speech and increased QoL exists among laryngeal cancer survivors. This, coupled with the fact that the fundamental objective of verbal communication is to be understood, creates a compelling argument as to why achieving effective and highly intelligible communication is so important for alaryngeal speakers and why continued research in this area is needed.

Conclusions

Loss of verbal communication presents a significant challenge in the presence of a potentially life-threatening disease such as laryngeal cancer. Thus, the ability to provide a functional means of verbal communication is an essential component of postlaryngectomy rehabilitation (Doyle, 1994). There has, however, been limited research conducted on alaryngeal SI over the past decade. This lack of more contemporary information is troublesome when one considers the significant changes that have occurred in the treatment of laryngeal cancer. Given research confirming the impact that communication effectiveness has on an individual's QOL, it is important that updated research on intelligibility be conducted. SLPs and physicians can benefit from more detailed information outlining the intelligibility patterns of those who use alaryngeal methods of speech. Alaryngeal speech continues to be characterized by multiple sound errors and that variability in sound intelligibility also exists specific to whether sounds appear within a word-initial or word-final position; these types of changes are further impacted by potential distractions secondary to the unusual quality of alaryngeal voice and speech, as well as how intelligibility is altered in conditions of competing noise. It is anticipated that clinical intervention can serve each individual to achieve the most intelligible speech possible. Information on various aspects of SI will allow healthcare professionals to better structure their treatment and therapy for each individual, providing them with the best opportunity to achieve the most intelligible speech possible, leading to more effective verbal communication, participation in society, and increased QOL.

References

Ackerstaff, A. H., Hilgers, F. J. M., Aaronson, N. K., & Balm, A. J. M. (1994). Communication, functional disorders and lifestyle changes after total laryngectomy. *Clinical Otolaryngology and Allied Sciences, 19*(4), 295–300.

Amster, W. W., Love, R. J., Menzel, O. J., Sandler, J., Sculthorpe, W. B., & Gross, F. M. (1972). Psychosocial factors and speech after laryngectomy. *Journal of Communication Disorders, 5*, 1–18.

Baggs, T. W., & Pine, S. J. (1983). Acoustic characteristics: Tracheoesophageal speech. *Journal of Communication Disorders, 16*(4), 299–307.

Bellandese, M. H., Lerman, J. W., & Gilbert, H. R. (2001). An acoustic analysis of excellent female esophageal, tracheoesophageal, and laryngeal speakers. *Journal of Speech, Language, and Hearing Research, 44*(6), 1315–1320.

Bennett, S., & Weinberg, B. (1973). Acceptability ratings of normal, esophageal, and artificial larynx speech. *Journal of Speech, Language, and Hearing Research, 16*(4), 608–615.

Beukelman, D. R., & Yorkston, K. M. (1980). Influence of passage familiarity on intelligibility estimates of dysarthric speech. *Journal of Communication Disorders, 13*, 33–41.

Black, J. W., & Haagen, C. H. (1963). Multiple choice intelligibility tests, forms A and B. *Journal of Speech and Hearing Disorders, 28*, 77–86.

Blom, E. D. (1998). Evolution of tracheoesophageal voice prostheses. In E. D. Blom, M. I. Singer, & R. C. Hamaker (Eds.), *Tracheoesophageal voice restoration following total laryngectomy* (pp. 1–8). San Diego, CA: Singular Publishing Group.

Blom, E. D., Singer, M. I., & Hamaker, R. C. (1986). A prospective study of tracheoesophageal speech. *Archives of Otolaryngology - Head and Neck Surgery, 112*, 440–447.

Blood, G. W. (1981). The interactions of amplitude and phonetic quality in esophageal speech. *Journal of Speech and Hearing Research, 24*, 308–312.

Bohnenkamp, T. A., Forrest, K., Klaben, B. K., & Stager, J. (2012). Chest wall kinematics during speech breathing in tracheoesophageal speakers. *Annals of Otology, Rhinology & Laryngology, 121*(1), 28–37.

Bornbaum, C. C., Fung, K., Franklin, J. H., Nichols, A., Yoo, J., & Doyle, P. C. (2012). A descriptive analysis of the relationship between quality of life and distress in individuals with head and neck cancer. *Supportive Care in Cancer, 20*(9), 2157–2165.

Bridges, A. (1991). Acceptability ratings and intelligibility scores of alaryngeal speakers by three listener groups. *British Journal of Disorders of Communication, 26*, 325–335.

Christensen, J. M., & Dwyer, P. E. (1990). Improving alaryngeal speech intelligibility. *Journal of Communication Disorders, 23*, 445–451.

Christensen, J. M., Weinberg, B., & Alfonso, P. J. (1978). Productive voice onset time characteristics of esophageal speech. *Journal of Speech and Hearing Research, 21*, 56–62.

Clark, J., & Stemple, J. (1982). Assessment of three modes of alaryngeal speech with a synthetic sentence identification task in varying message to competition ratios. *Journal of Speech and Hearing Research, 25*, 333–338.

Clements, K. S., Rassekh, C. H., Seikaly, H., Hokanson, J. A., & Calhoun, K. H. (1997). Communication after laryngectomy: An assessment of patient satisfaction. *Archives of Otolaryngology - Head and Neck Surgery, 123*, 493–496.

Connor, N. P., Hamlet, S. L., & Joyce, J. C. (1985). Acoustic and physiologic correlates of the voicing distinction in esophageal speech. *Journal of Speech and Hearing Disorders, 50*, 378–384.

Cullinan, W. L., Brown, C. S., & Blalock, P. D. (1986). Ratings of intelligibility of esophageal and tracheoesophageal speech. *Journal of Communication Disorders, 19*, 185–195.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech, Language, and Hearing Research, 12*(2), 246–269.

Diedrich, W. M. (1968). The mechanism of esophageal speech. *Annals of the New York Academy of Sciences, 155*, 303–317.

Doyle, P. C. (1994). *Foundations of voice and speech rehabilitation following laryngeal cancer*. San Diego, CA: Singular.

Doyle, P. C. (1998). Postlaryngectomy rehabilitation: Clinical care following treatment for laryngeal cancer: The opening of doors. *Perspectives on Voice and Voice Disorders, 8*(3), 4–4.

Doyle, P. C. (2005). Rehabilitation in head and neck cancer: Overview. In P. C. Doyle & R. L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech and swallowing* (pp. 3–15). Austin, TX: PRO-ED, Incorporated.

Doyle, P. C. (2017a). Communication challenges in laryngeal cancer. In J. C. Friberg & L. A. Vinney (Eds.), *Laryngeal cancer: An interdisciplinary resource for practitioners* (pp. 73–89). Thorofare, NJ: Slack Inc.

Doyle, P. C. (2017b, unpublished data). The influence of formal test stimuli construction on assessments of tracheoesophageal speech intelligibility. Voice Production and Perception Laboratory, Western University, London, ON.

Doyle, P. C., Danhauer, J. L., & Lucks-Mendel, L. L. (1990). A SINDSCAL analysis of perceptual features for consonants produced by esophageal and tracheoesophageal talkers. *Journal of Speech and Hearing Disorders, 55*(4), 756–760.

Doyle, P. C., Danhauer, J. L., & Reed, C. G. (1988). Listeners' perceptions of consonants produced by esophageal and tracheoesophageal talkers. *Journal of Speech and Hearing Disorders, 53*, 400–407.

Doyle, P. C., & Eadie, T. L. (2005). Pharyngoesophageal segment function: A review and reconsideration. In P. C. Doyle & R. L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech and swallowing* (pp. 521–543). Austin, TX: PRO-ED, Incorporated.

Doyle, P. C., & Haaf, R. G. (1989). Perception of prevocalic and post-vocalic consonants produced by tracheoesophageal speakers. *The Journal of Otolaryngology, 18*(7), 350–353.

Doyle, P. C., Swift, E. R., & Haaf, R. G. (1989). Effects of listener sophistication on judgments of tracheoesophageal talker intelligibility. *Journal of Communication Disorders, 22*, 105–113.

Duffy, J. R., & Giolas, T. G. (1974). Sentence intelligibility as a function of key word selection. *Journal of Speech, Language, and Hearing Research, 17*(4), 631–637.

Eadie, T. L. (2003). The ICF: A proposed framework for comprehensive rehabilitation of individuals who use alaryngeal speech. *American Journal of Speech-Language Pathology, 12*(2), 189–197.

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.

Eadie, T. L., & Doyle, P. C. (2004). Auditory-perceptual scaling and quality of life in tracheoesophageal speakers. *The Laryngoscope, 114*, 753–759.

Eadie, T. L., & Doyle, P. C. (2005). Quality of life in male tracheoesophageal (TE) speakers. *Journal of Rehabilitation Research and Development, 42*(1), 115–124.

Eadie, T. L., Doyle, P. C., Hansen, K., & Beaudin, P. G. (2008). Influence of speaker gender on listener judgments of tracheoesophageal speech. *Journal of Voice, 22*(1), 43–57.

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*(1), 1955–1961.

Epstein, R., Giolas, T. G., & Owens, E. (1968). Familiarity and intelligibility of monosyllabic word lists. *Journal of Speech and Hearing Research, 11*, 435–438.

Filter, M. D., & Hyman, M. (1975). Relationship of acoustic parameters and perceptual ratings of esophageal speech. *Perceptual and Motor Skills, 40*, 63–68.

Finizia, C., Dotevall, H., Lundstrom, E., & Lindstrom, J. (1999). Acoustic and perceptual evaluation of voice and speech quality. *Archives of Otolaryngology – Head and Neck Surgery, 125*, 157–163.

Giolas, T. G., Cooker, H. S., & Duffy, J. R. (1970). The predictability of words in sentences. *Journal of Auditory Research, 10*, 328–334.

Gomyo, Y., & Doyle, P. C. (1989). Perception of stop consonants produced by esophageal and tracheoesophageal speakers. *The Journal of Otolaryngology, 18*(4), 184–188.

Hillman, R. E., Walsh, M. J., & Heaton, J. T. (2005). Laryngectomy speech rehabilitation - A review of outcomes. In P. C. Doyle & R. L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech and swallowing* (pp. 75–90). Austin, TX: PRO-ED, Incorporated.

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. *The Annals of Otology, Rhinology & Laryngology, 107*(5), 2–21.

House, A. S., Williams, C. E., Hecker, M. H. L., & Kryter, K. D. (1965). Articulation testing methods: Consonantal differentiation with a closed response set. *Journal of the Acoustical Society of America, 37*(1), 158–166.

Hyman, M. (1955). An experimental study of artificial larynx and esophageal speech. *Journal of Speech and Hearing Disorders, 20*, 291–299.

Jongmans, P., Hilgers, F. J. M., Pols, L. C. W., & van As-Brooks, C. J. (2006). The intelligibility of tracheoesophageal speech, with an emphasis on voiced-voiceless distinction. *Logopedics Phoniatrics Vocology, 31*, 172–181.

Kalb, M. B., & Carpenter, M. A. (1981). Individual speaker influence on relative intelligibility of esophageal speech and artificial larynx speech. *Journal of Speech and Hearing Disorders, 46*, 77–80.

Kao, W. W., Mohr, R. M., Kimmel, C. A., Getch, C., & Silverman, C. (1994). The outcome and techniques of primary and secondary tracheoesophageal puncture. *Archives of Otolaryngology–Head & Neck Surgery, 120*(3), 301–307.

Karnell, L. H., Funk, G. F., & Hoffman, H. T. (2000). Assessing head and neck cancer patient outcome domains. *Head and Neck, 22*(1), 6–11.

Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility in dysarthria. *Journal of Speech and Hearing Disorders, 54*, 482–499.

Maryn, Y., Dick, C., Vandenbruaene, C., Vauterin, T., & Jacobs, T. (2009). Spectral, cepstral, and multivariate exploration of tracheoesophageal voice quality in continuous speech and sustained vowels. *The Laryngoscope, 119*(12), 2384–2394.

McColl, D., Fucci, D., Petrosino, L., Martin, D. E., & McCaffrey, P. (1998). Listener ratings of the intelligibility of tracheoesophageal speech in noise. *Journal of Communication Disorders, 31*(4), 279–289.

McDonald, R., Parsa, V., & Doyle, P. C. (2010). Objective estimation of tracheoesophageal speech ratings using an auditory model. *The Journal of the Acoustical Society of America, 127*(2), 1032–1041.

Meyer, T. K., Kuhn, J. C., Campbell, B. H., Marbella, A. M., Myers, K. B., & Layde, P. M. (2004). Speech intelligibility and quality of life in head and neck cancer survivors. *The Laryngoscope, 114*, 1977–1981.

Miralles, J. L., & Cervera, T. (1995). Voice intelligibility in patients who have undergone laryngectomies. *Journal of Speech and Hearing Research, 38*, 564–571.

Moon, J. B., & Weinberg, B. (1987). Aerodynamic and myoelastic contributions to tracheoesophageal voice production. *Journal of Speech, Language, and Hearing Research, 30*(3), 387–395.

Nagle, K. F., & Eadie, T. L. (2012). Listener effort for highly intelligible tracheoesophageal speech. *Journal of Communication Disorders, 45*(3), 235–245.

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.

Pauloski, B. R., Fisher, H. B., Kempster, G. B., & Blom, E. D. (1989). Statistical differentiation of tracheoesophageal speech produced under four prosthetic/occlusion speaking conditions. *Journal of Speech, Language, and Hearing Research, 32*(3), 591–599.

Pindzola, R. H., & Cain, B. H. (1988). Acceptability ratings of tracheoesophageal speech. *Laryngoscope, 98*, 394–397.

Qi, Y., & Weinberg, B. (1995). Characteristics of voicing source waveforms produced by esophageal and tracheoesophageal speakers. *Journal of Speech, Language, and Hearing Research, 38*(3), 536–548.

Robbins, J. (1984). Acoustic differentiation of laryngeal, esophageal, and tracheoesophageal speech. *Journal of Speech and Hearing Research, 27*, 577–585.

Robbins, J., Christensen, J. M., & Kempster, G. (1986). Characteristics of speech production after tracheoesophageal puncture: Voice onset time and vowel duration. *Journal of Speech and Hearing Research, 20*, 499–504.

Robbins, J., Fisher, H. B., Blom, E. C., & Singer, M. I. (1984). A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production. *Journal of Speech and Hearing Disorders, 49*, 202–210.

Schiavetti, N. (1992). Scaling procedures for the measurement of speech intelligibility. In R. D. Kent (Ed.), *Intelligibility in speech disorders* (pp. 11–34). Philadelphia, PA: John Benjamins Publishing Co.

Searl, J. P., & Carpenter, M. A. (2002). Acoustic cues to the voicing feature in tracheoesophageal speech. *Journal of Speech, Language, and Hearing Research, 45*, 282–294.

Searl, J. P., Carpenter, M. A., & Banta, C. L. (2001). Intelligibility of stops and fricatives in tracheoesophageal speech. *Journal of Communication Disorders, 34*, 305–321.

Singer, M. I., & Blom, E. D. (1980). An endoscopic technique for restoration of voice after laryngectomy. *Annals of Otology, Rhinology, and Laryngology, 89*, 529–533.

Singer, M. I., Blom, E. D., & Hamaker, R. C. (1983). Voice rehabilitation after total laryngectomy. *The Journal of Otolaryngology, 12*(5), 329–334.

Sisty, N. L., & Weinberg, B. (1972). Formant frequency characteristics of esophageal speech. *Journal of Speech, Language, and Hearing Research, 15*(2), 439–448.

Smith, B. E., Weinberg, B., Feth, L. L., & Horii, Y. (1978). Vocal roughness and jitter characteristics of vowels produced by esophageal speakers. *Journal of Speech, Language, and Hearing Research, 21*(2), 240–249.

Smith, L. F., & Calhoun, K. H. (1994). Intelligibility of tracheoesophageal speech among naïve listeners. *Southern Medical Journal, 87*(3), 333–341.

Snidecor, J. C., & Curry, E. T. (1960). How effectively can the laryngectomee expect to speak? *The Laryngoscope, 70*(1), 62–67.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual*. New York, NY: Neural and Social Prospects Wiley.

Subtelny, J. (1977). Assessment of speech with implications for training. In F. Bess (Ed.), *Childhood deafness* (pp. 183–194). New York, NY: Grune & Stratton.

Tardy-Mitzell, S., Andrews, M. L., & Bowman, S. A. (1985). Acceptability and intelligibility of tracheoesophageal speech. *Archives of Otolaryngology, 111*(4), 213–215.

Terrell, J. E., Ronis, D. L., Fowler, K. E., Bradford, C. R., Chepeha, D. B., Prince, M. E., … Duffy, S. A. (2004). Clinical predictors of quality of life in patients with head and neck cancer. *Archives of Otolaryngology – Head and Neck Surgery, 130*, 401–408.

Tjaden, K. K., & Liss, J. M. (1995). The role of listener familiarity in the perception of dysarthric speech. *Clinical Linguistics & Phonetics, 9*(2), 139–154.

Trudeau, M. D. (1994). The acoustical variability of tracheoesophageal speech. In R. L. Keith & F. L. Darley (Eds.), *Laryngectomy rehabilitation* (pp. 383–394). Austin, TX: Pro-Ed, Incorporated.

van As, C. J., Hilgers, F. J. M., Verdonck-de Leeuw, I. M., & Koopmans-van Beinam, F. J. (1998). Acoustical analysis and perceptual evaluation of tracheoesophageal prosthetic voice. *Journal of Voice, 12*, 239–248.

Weinberg, B. (1980). *Readings in speech following total laryngectomy*. University Park Press.

Weinberg, B., Horii, Y., Blom, E., & Singer, M. (1982). Airway resistance during esophageal phonation. *Journal of Speech and Hearing Disorders, 47*(2), 194–199.

Weinberg, B., Horii, Y., & Smith, B. E. (1980). Long-time spectral and intensity characteristics of esophageal speech. *The Journal of the Acoustical Society of America, 67*(5), 1781–1784.

Weiss, M. S., & Basili, A. G. (1985). Electrolaryngeal speech produced by laryngectomized subjects: Perceptual characteristics. *Journal of Speech and Hearing Research, 28*, 294–300.

Williams, S., & Watson, J. B. (1985). Differences in speaking proficiencies in three laryngectomy groups. *Archives of Otolaryngology, 111*, 216–219.

World Health Organization. (2001). *ICF: International Classification of Functioning, Disability and Health*. Geneva, Switzerland: Author.

Yorkston, K. M., & Beukelman, D. R. (1978). A comparison of techniques for measuring intelligibility of dysarthric speech. *Journal of Communication Disorders, 11*, 499–512.

Yorkston, K. M., & Beukelman, D. R. (1980). A clinician-judged technique for quantifying dysarthric speech based on single-word intelligibility. *Journal of Communication Disorders, 13*, 15–31.

Yoshida, G. Y., Hamaker, R. C., Singer, M. I., Blom, E. D., & Charles, G. A. (1989). Primary voice restoration at laryngectomy: 1989 update. *The Laryngoscope, 99*(10), 1093–1095.