6-19-2012 12:00 AM

Exploring Intelligibility in Tracheoesophageal Speech: A Descriptive Analysis

Lindsay E. Sleeth
The University of Western Ontariio

Supervisor
Dr. Philip C. Doyle
The University of Western Ontario

Graduate Program in Health and Rehabilitation Sciences
A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

EXPLORING INTELLIGIBILITY IN TRACHEOESOPHAGEAL SPEECH: A DESCRIPTIVE ANALYSIS

(Spine title: Intelligibility in Tracheoesophageal Speech)

(Thesis format: Monograph)

by

Lindsay E. Sleeth BHSc.

Graduate Program in
Health and Rehabilitation Science

A thesis submitted in partial fulfillment
of the requirements for the degree
Master of Science

School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

CERTIFICATE OF EXAMINATION

Supervisor	Examiners
_____	_____
Philip C. Doyle, Ph.D.	Kevin Fung, M.D.
Advisory Committee	_____
	J.B. Orange, Ph.D.
_____	_____
Kevin Fung, M.D.	Jeffrey Tennant, Ph.D.

The thesis by

Lindsay E. Sleeth

entitled:

Exploring Intelligibility in Tracheoesophageal Speech: A Descriptive Analysis

is accepted in partial fulfillment of the
requirements for the degree of
Master of Science
(Rehabilitation Science)

Date_____ _____

Chair of the Thesis Examination Board

ABSTRACT

Despite literature that suggests tracheoesophageal (TE) voice restoration to have the highest intelligibility of the three alaryngeal modes of speech, previous studies have shown that TE speech is less intelligible than "normal" speech. It is important to understand where problems related to intelligibility currently exist in order for members of the rehabilitation team to provide the best therapy/prostheses to each individual using TE speech as a mode of communication. This study evaluated the intelligibility of 15 male and female tracheoesophageal speakers. Eighteen normal-hearing, naive, young adult listeners assessed digital voice samples of 15 adult male and female TE speakers. Listeners made judgments by transcribing the monosyllabic words heard into English orthographics. Confusion matrices were then generated based on the transcriptions. The data were analyzed to determine overall intelligibility and to determine if patterns of increased or decreased intelligibility existed based on manner of classification.

Keywords: laryngeal cancer, tracheoesophageal speech, speech intelligibility, postlaryngectomy rehabilitation, quality of life.

DEDICATION

This work is dedicated to Heather, Ryan and Allie, without whom, none of this would have been possible.

ACKNOWLEDGMENTS

First and foremost, to my supervisor, Dr. Philip Doyle, for your unwavering support and confidence, from start to finish. Your continual guidance, support, and of course, humour, has made this experience a rich and fulfilling one. I consider myself very lucky to have been given the opportunity to work with such a fantastic mentor and supervisor. I have learned more from you than I ever thought possible. Thank you.

To the ENT team at Victoria Hospital. Thank you for always welcoming and supporting me, and for creating an environment from which to learn and grow, as I have.

To the IAL. Firstly, thank you for making this research possible. More importantly, thank you for welcoming me into a truly wonderful community of wonderful people, and for teaching me more in five days than a course ever could. I will cherish the conversations and experiences I had with each and every one of you.

To the labrats. Thank you for being my mentors, confidantes, and most importantly, friends. Thank you also for your countless answered questions, tech support, and never-ending entertainment. You are all truly one of a kind.

To B.M. and J.C. For making me laugh and smile. Thank you for always being prepared to show me the lighter side of life, and most importantly, for being who you are.

To Ryan, thank you for always making me laugh, no matter what. You are irreplaceable.

And lastly, to my mom, the single best person I know. Thank you for your never-ending confidence, support and love. Most importantly, thank you for making me the person I am today. I could never have done this without you. I love you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1
Introduction and Review of the Literature

At present, laryngeal cancer is the most prevalent form of head and neck cancer, and is defined as cancer originating from the larynx, or voice box (Mendenhall et al., 2002). A diagnosis of laryngeal cancer has the potential to impact all areas of an individual's life, including their physical health, emotional, psychological, economic, and social well-being. (Doyle, 1994; Eadie & Doyle, 2004, 2005; Meyer et al., 2004). Distinctive to a diagnosis of laryngeal cancer is the potential need to surgically remove the voice box, leading to the loss of the individual's normal vocal mechanism and subsequently, a loss of normal verbal communication. This creates a unique set of challenges not typically experienced with other sites of cancers. Studies have shown speech and verbal communication to be one of the biggest predictors of quality of life in individuals with laryngeal cancer (Eadie & Doyle, 2004; Karnell, Funk, & Hoffman, 2000; Meyer et al., 2004; Terrell et al., 2004; Theurer & Martin, 2003). The notion of "quality of life" encompasses the potentially impacted areas of an individual's life mentioned above (physical health, emotional, psychological, economic and social well-being). When expressed using the International Classification of Functioning, Disability and Health (ICF), the issues associated with laryngeal cancer encompass all of the four components of the framework (body functions and structures, activities and participation, environmental factors, and personal factors). 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 quality of life. With an improvement of quality of life comes an improvement in all aspects of health and functioning. Therefore, it is important that research be conducted in the area of postlaryngectomy speech rehabilitation in order for

those individuals relying on it as a mode of communication receive the best therapy and devices possible to allow for the most effective and positive communication.

The review to follow will address information regarding laryngeal cancer and its treatments, as well as communication loss and rehabilitation after laryngectomy. Additionally, the evaluation of speech intelligibility and the methods for such assessment with a highlight on the intelligibility of tracheoesophageal speech will be a focus.

Laryngeal Cancer

Etiology

A wide range of risk factors currently exist relative to an increased likelihood of a laryngeal cancer diagnosis. At present, tobacco and alcohol use are the biggest risk factors, with risk directly correlated with usage. According to research conducted by Talamini et al. (2001), current tobacco users have a 20 times higher risk of developing laryngeal cancer than those who have never smoked, with the number of cigarettes smoked, as well as the number of years consuming tobacco positively correlated with laryngeal cancer incidence. Conversely, it has also been shown that risk is greatly reduced after smoking cessation, although the risk will never decrease to levels as low as never-smokers (Talamini et al., 2001).

This same dose-response relationship has been shown in regards to alcohol consumption, although to a smaller extent than that observed with tobacco (Talamini et al., 2001). Unfortunately, one aspect of alcohol consumption found to be dissimilar to tobacco trends is the decrease in risk with cessation of consumption. At this time, there does not appear to be a favourable link between cessation of alcohol consumption and a lowered risk in developing laryngeal cancer (Talamini et al., 2001).

It is imperative to mention that the use of tobacco and alcohol together is said to contribute more strongly to increasing the risk of laryngeal cancer than the use of either alone (Kacker, Wolden, Pfister, & Kraus, 2003; Talamini et al., 2001). According to Hasibe et al. (2007) as many as 75% of all head and neck cancers can be attributed to the synergistic effects of the combined use of tobacco and alcohol. Alcohol is a chemical solvent, and due to the nature of this, it is believed that alcohol has the ability to enhance and prolong the exposure of the carcinogens found in tobacco to the mucous membranes, thereby, creating a synergistic effect (Pai & Westra, 2009).

Another risk factor currently gaining interest in head and neck cancer carcinoma (of which laryngeal cancer is a subset), is the human papilloma virus (HPV). The role of HPV in head and neck cancer was first suggested by Syrjanen and colleagues in 1983 (Syrjanen, 2005). Since then, many studies have further confirmed this link and determined that HPV-associated head and neck cancer is transmitted predominately through sexual behaviours (Marur, D'Souza, Westra, & Forastiere, 2010; McKaig, Baric, & Olshan, 1998; Pai & Westra, 2009; Paz, Cook, Odom-Maryon, Xie, & Wilczynski, 2000; Syrjanen, 2005). Transmission is usually from high-risk HPV's (over 90% attributable to type 16) with increased risk associated with increased number of sexual partners, a history of practicing oral sex, and younger age at culmination of intercourse; all stemming from changes in sexual norms (Gillison & Lowry, 2004; Marur et al., 2010; McKaig et al., 1998; Pai & Westra, 2009). As well, HPV DNA is increasingly being found in non-smoking, younger, individuals, which has the potential to alter the current demographics of the head and neck cancer population (Gillison & Lowry, 2004). Although HPV-DNA is said to be found most often in carcinomas of the oral cavity,

HPV-DNA has also been found in laryngeal carcinomas (Syrjanen, 2005). Fortunately though, individuals with HPV-positive head and neck cancer appear to have better prognosis due to increased sensitivity and response to chemotherapy and radiation treatment (Gillison & Lowry, 2004; Marur et al., 2010).

In addition to tobacco, alcohol and HPV as causal agents, there are a host of other risk factors that may contribute to the development of laryngeal cancer, including, but not limited to, gastroesophageal reflux disease (GERD) (Freije et al., 1996), genetic susceptibility (Kacker, et al., 2003), poor oral hygiene (Pai & Westra, 2009), diets deficient in vitamin A, and fruits and vegetables (Pai & Westra, 2009), contraction of the Epstein-Barr virus (Tyan et al., 1993), marijuana smoke (Pai & Westra, 2009), and occupations exposed to asbestos, chromium, radiation, mustard gas, leatherworking, nickel refinement, textiles, woodworking, and metalworking (Fauci et al., 2008; Pai & Westra, 2009). All of these factors have the potential to contribute to the development of laryngeal cancer, and the subsequent consequences associated with the disease.

Incidence and Mortality

In 2011 alone, 1,150 Canadians will be diagnosed with laryngeal cancer, and 490 individuals will lose their lives as a result (Canadian Cancer Society, 2011). Laryngeal cancer is more prominent in males than females, with a ratio of 4:1 (Canadian Cancer Society, 2011). These current trends are a significant change from the past 30 years, at which time the proportion of men-to-women with laryngeal cancer was approximately 10:1. This indicates a jump in the number of females being diagnosed with laryngeal

cancer over the past few decades with an associated narrowing of the male-to-female ratio for such a diagnosis.

In recent years the Canadian Cancer Society has indicated that the incidence of laryngeal cancer is continuing to decrease in both males and females as a result of decreasing societal trends to engage in heavy drinking and smoking. Despite this fact, one in approximately 33,000 Canadians is diagnosed with cancer of the larynx annually (Canadian Cancer Society, 2011). As well, although the prevalence of smoking in Canada is continuing to see a slow decrease overall, current federal surveys suggest a slight increase in smoking in specific age ranges of females, namely at 15 to 17 years, and over the age of 25 (Smoking in Canada, 2011). If this trend continues, there is a potential to see an increase in the incidence rate of laryngeal cancer in the years to come.

Treatment Options

At present, the treatment modalities available for the treatment of laryngeal cancer are radiation therapy, chemotherapy, surgery, or a combination of some or all of these options (Kacker et al., 2003). The choice of one treatment over another depends on multiple factors including: the size of the tumour, the location, the stage of the cancer, and involvement of surrounding associated structures. However, in all cases the primary determinant of treatment options is based on eradication of the cancer and oncologic safety. Radiation therapy is the most common and preferred treatment for laryngeal cancer, and involves applying high-energy, electromagnetic emissions to the cancerous area (Matthews & Lampe, 2005). Radiation is preferred as surgery is not needed and therefore, less healthy tissue is lost or distorted. Radiation therapy is not without its side

effects though. Individuals treated with radiation therapy may acutely experience skin tenderness, and difficulty swallowing, which can result in weight loss and need for a feeding tube (Matthews & Lampe, 2005). These side effects, as well as decreased muscle strength, muscle fibrosis, and increased susceptibility to infection may become chronic issues that can indefinitely persist, even after completion of radiation therapy (Matthews & Lampe, 2005). Radiation therapy may be ineffective for larger and more advanced tumours and therefore may require the use of chemotherapy and/or surgery in order to remove and manage the cancer most effectively (Treatment for Laryngeal Cancer, 2011).

Chemotherapy is often used in conjunction with radiation therapy in tumours larger in size, and with those that have spread to the surrounding lymphatic system (Treatment for Laryngeal Cancer, 2001). Due to the fact that chemotherapy also damages healthy cells, the side effects associated with this form of treatment are vast and may include: nausea, vomiting, loss of appetite, fatigue, hair loss and increased susceptibility to infection (Treatment for Laryngeal Cancer, 2001).

Surgery may be selected as the sole treatment in early-stage cancers, combined with radiation in advanced tumours, or as a secondary option if other methods fail (Matthews & Lampe, 2005). If surgery is deemed necessary, often times a total laryngectomy will need to be performed in order for the tumor to be fully removed. Total laryngectomy involves removal of the thyroid, cricoid and arytenoid cartilages, the intrinsic membranes and muscles of the neck, the hyoid bone, the four muscles of the infrahyoid (the strap muscles), and one or more tracheal rings (Kacker et al., 2003). Due to the vast array of structures removed during this procedure, the mechanism and structures that create voice and speech (the vocal folds) are also removed, resulting in the

loss of the individual's ability to produce natural voice and speech after surgery. As a result of this loss of communication, individuals with laryngeal cancer are faced with challenges during the rehabilitation process that are unique, and arguably more difficult, than individuals with other types of cancer because of the individual's inability to verbally communicate. Loss of verbal communication presents a significant challenge in the presence of a potentially life-threatening disease such as laryngeal cancer. Hence, the ability to provide a functional means of verbal communication is an essential component of postlaryngectomy rehabilitation (Doyle, 1994).

Postlaryngectomy Voice and Speech Rehabilitation

Alaryngeal Speech

In the event of a total laryngectomy, an alternate method of postlaryngectomy "alaryngeal" speech needs to be acquired to allow the individual to verbally communicate. At present, there are three primary methods of alaryngeal speech employed by laryngectomized individuals: electrolaryngeal, esophageal, and tracheoesophageal (TE) speech. Electrolaryngeal speech can be achieved through two separate methods (transcervical or intra oral), which differ in their placement of the device. The transcervical approach involves placing an electronic device on tissues of the neck. Once activated, the vibratory head of the electronic device generates sound which is then transmitted through the tissues of the pharynx, hypopharynx or into the oral cavity, where it is shaped by the articulators and speech is created (Keith, Shanks, & Doyle, 2005). The intra-oral approach follows along the same mechanism, with the addition of a plastic tube that is attached to the electronic device and inserted into the oral cavity (Farrell, Dietrich-

Burns, & Messing, 2004). Sound is transferred from the electronic device, through the plastic tubing, and into the oral cavity where it is shaped by the articulators to create speech, much like the transcervical approach (Farrell et al., 2004).

Esophageal speech involves insufflating or injecting air into the esophagus, which is then expelled volitionally. As the air travels back up and out the esophagus it vibrates the tissues of the upper esophagus and lower pharynx, or what has been termed the pharyngoesophageal (PE) segment, which creates a new sound source (Diedrich, 1968).

In 1979, Mark I. Singer, MD and Eric D. Blom, PhD developed the tracheoesophageal (TE) puncture voice restoration method and the first TE puncture voice prosthesis, and shortly thereafter TE speech became known as an international standard for voice restoration following total laryngectomy (Singer & Blom, 1980). The TE puncture voice restoration procedure can be performed at the time of total laryngectomy (primary TEP), or as a separate procedure following total laryngectomy (secondary TEP) (Gress & Singer, 2005). The procedure involves creating a small puncture through the posterior wall of the trachea into the esophagus. 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 PE segment. Upon exhalation, and when the tracheostoma is occluded by the individual's thumb or another device, pulmonary air is shunted into the esophagus, setting the PE segment into vibration and allowing for sound generation (Gress & Singer, 2005).

TE speech has come to be the preferred method of choice for many individuals and clinicians, for a variety of reasons. Firstly, it has been identified by multiples sources that TE speech acquisition is fast and simple, with success rates ranging from 80% to

90%; the highest of the three alaryngeal methods (Blom, Singer, & Hamaker, 1986; Gress & Singer, 2005; Hillman, Walsh, & Heaton, 2005) This is especially impressive when compared to esophageal speech, which has a reported failure rate as great as 55% in some cases[1] (Blom et al., 1986). This relative ease of acquisition has been attributed to many factors, including, but not limited to: the simplicity of the surgical procedure, high levels of overall fluency and intelligibility, and the minimal speech therapy needed (Hillman et al., 2005). This, coupled with the potential for the procedure to be completed at the time of total laryngectomy allows for a rapid restoration of voice and verbal communication for the individual postlaryngectomy.

Additionally, the ability to make use of the pulmonary air supply in the TE method allows for increased overall speech rate and syllable rate, with words and syllables per minute capable of reaching values similar to that of normal speech (Robbins, Fisher, Blom, & Singer, 1984). Finally, the literature is rich with studies comparing a variety of features of the three alaryngeal methods to both each other and to normal speech (Blom et al., 1986; Clements, Rassekh, Seikaly, Hokanson, & Calhoun, 1997; Cullinan, Brown, & Blalock, 1986; Doyle, Danhauer, & Reed, 1988; Robbins, 1984; Robbins et al., 1984; Tardy-Mitzell, Andrews, & Bowman, 1985). These studies have found TE speech to be superior to the other alaryngeal modes in areas such as acceptability, overall intelligibility, pitch, intensity, and patient satisfaction, with values approaching those similar to normal speech in some instances.

---

[1] Previously, failure to acquire esophageal speech was attributed to the individual, and a lack of motivation or "laziness". Improved understanding of the PE segment and its functioning has allowed researchers to alter this viewpoint and instead attribute failure rates to the altered anatomy and physiology of these individuals. Please see a detailed discussion by Doyle & Eadie (2005) for further information.

It is also of importance to note that TE speech is not without its limitations. Studies have shown intensity to be higher than laryngeal speech, with greater pauses in between utterances (Robbins et al., 1984). As well, female TE speakers tend to have pitch values similar to those found in males, giving their voice a lower and more masculine sound (Trudeau, 1994). In addition, although TE speech has shown to be highly acceptable, it is clearly judged as less acceptable than laryngeal speakers (Clark & Stemple, 1982; Finizia, Dotevall, Lundstrom, & Lindstrom, 1999; van As, Hilgers, Verdonck-de Leeuw, & Koopmans-van Beinum, 1998). Finally, TE speech is not 100% intelligible and is lower than that of laryngeal speakers (Blom et al., 1986; Doyle et al., 1988; Pindzola & Cain, 1988; Robbins, 1984; Williams & Watson, 1985; and others). Speech intelligibility is an area that received generous attention when TE speech was first introduced but unfortunately, has been somewhat overlooked for the past 20 years. Lowered speech intelligibility has the potential to negatively impact a person's participation in society, and is an issue that should be addressed in further detail. Information on speech intelligibility will be presented with the subsequent section of this review.

Speech Intelligibility

Speech intelligibility has been defined by Hillman, Walsh, and Heaton (2005) as the percentage of speech items correctly identified by the listener. Similarly, Kent, Weismer, Kent, and Rosenbek (1989) define speech intelligibility as "the degree to which the speaker's intended message is recovered by the listener" (p. 483). Over the years, intelligibility of alaryngeal speech has been studied by numerous individuals, with varying populations of speakers, under a variety of conditions, with a variety of stimuli

(Amster, Love, Menzel, Sandler, Sculthorpe, & Gross, 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 et al., 1985; Weiss & Basili, 1985; and others). Results of these studies have shown that speech intelligibility has the potential to increase or decrease based on a range of factors such as: experience of the speakers, experience of the listeners, stimuli, background noise and environment, gender, type of postlaryngectomy speech mode, etc. Therefore, when considering the findings and implications of speech intelligibility research in postlaryngectomy populations, it is important to consider all of these factors.

Speech Intelligibility Testing

A host of factors have the potential to impact the results of speech intelligibility research. It is therefore important to understand and consider these factors before evaluating or conducting research in this area. Throughout history, intelligibility measurement has largely been obtained through two separate methods: scaling procedures and word identification (Shiavetti, 1992). Previously, scaling procedures, such as the use of equal appearing interval scales which allow the listener to make judgments about a speaker's intelligibility, were used more frequently due to their ease of application and scoring (Shiavetti, 1992). Recently, as intelligibility testing has continued to grow in many disordered speech populations, these scaling procedure methods have received attention and criticism. Although timely and efficient, scaling procedures lack the ability to pinpoint specific areas of increased or decreased intelligibility, and have limited strength in estimating an intelligibility score for each individual without obtaining

percentage values (Shiavetti, 1992). This decreases their generalizability to other studies conducted on intelligibility, as well as makes it more difficult for laypeople to interpret. Thus, word identification procedures have increasingly become the method of choice when conducting intelligibility research, especially for alaryngeal speech. With this method, listeners are required to write down each word, sentence, or phrase uttered by the speaker. The listeners responses are then compared to the list of responses intended to be produced by the speakers, which is subsequently converted into a percentage of incorrect and correct responses, resulting in a overall intelligibility score (Shiavetti, 1992). This method has the advantage that it is easily interpretable to not only clinicians, but also naïve individuals. As well, a measure of intelligibility is directly produced and available for immediate dissemination to those seeking the information. Lastly, the measure is objective, and has the potential to identify areas where intelligibility deficits exist in each individual (Shiavetti, 1992). The sensitivity and ability to gain so much information solely from word identification procedures has made it 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 more).

Another area of intelligibility testing with the ability to impact findings is the experience of the listeners. Previous intelligibility studies have employed the use of either naïve listeners (no prior experience with the disordered population) or experienced listeners (typically, speech-language pathologists (SLPs)). Multiple studies have previously shown that intelligibility may be influenced by the sophistication of the listener population (Beukelman & Yorkston, 1980; Doyle, Swift, & Haaf, 1989; Williams & Watson, 1985). These studies all suggest that intelligibility reports made by SLPs are

higher than those made by inexperienced listeners. The experienced listeners' prior exposure to the speaker population, and most likely the stimuli as well, have the ability to inflate the intelligibility scores. This makes the information less generalizable and unrepresentative of the general listening population, which is typically not composed entirely of SLPs. The use of naïve listeners can influence findings as well. Firstly, since naïve listeners have had little exposure to alaryngeal speech they may focus on the unnatural quality of the voice instead of the words or sounds being produced, confounding the data. As well, naïve listeners may not be entirely familiar with the words or passages in the recordings, leading to confusions or errors due to lack of experience with the stimuli, rather than lack of intelligibility from the TE speaker.

When conducting intelligibility research or evaluating the validity of previous research it is important to consider these external factors that have the potential to influence results.

Speech Intelligibility in TE Speech

Many studies have compared the three modes of alaryngeal speech and have found that TE speech is generally 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, Singer, and Hamaker (1986) conducted a study assessing the intelligibility of individuals both before undergoing the TE puncture procedure and after. Prior to the procedure, these individuals were using either esophageal speech or an electrolarynx as their primary mode of communication. Following the procedure, all individuals used TE speech to communicate. Intelligibility was determined through the percentage of correct responses

found on a multiple-choice response form, with the stimulus words taken from the Modified Rhyme Test. This study saw a statistically significant improvement in the group's speech intelligibility following the TE puncture, with preoperative mean intelligibility at 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 superiority of TE speech when compared to the two other alaryngeal modes of communication. In this context, it is important to note though that the use of a forced-choice paradigm in this study had the potential to influence the results, thereby leading to higher intelligibility scores.

With that being said, TE speech is still less intelligible than speech produced by an individual with an intact larynx (Hillman et al., 2005). Studies have reported intelligibility of TE speech ranging from 65% to 93% (Doyle et al., 1988; Pindzola & Cain, 1988; Tardy-Mitzell et al., 1985). Doyle et al. (1988) determined intelligibility through the use of recorded consonant-vowel-consonant-vowel-consonant (CVCVC) nonsense syllables that were transcribed by naïve listeners into IPA format through an open-response paradigm. This resulted in an average intelligibility of 65%, (range 59%-72%). Pindzola and Cain (1988) used an entirely different method to determine intelligibility in their study. TE speakers recorded monosyllabic English words from the Multiple Choice Intelligibility Test (Black & Haagen, 1963). Naïve listeners then chose the correct response from four possible options, in a forced-choice paradigm. This study found an overall intelligibility of 93.20% across speakers. Tardy-Mitzell et al. (1985) used a method similar to that used in the study by Pindzola and Cain (1988). Intelligibility was judged from monosyllabic word lists developed by House, Williams,

and Hecker (1965). This study also employed the forced-choice method with six possible options for each stimulus word. Comparable intelligibility values were also found with an average of 93% (range 80.70%-97.50%). As demonstrated in the above studies, intelligibility has the potential to vary based on internal and external factors pertaining to the study and design (stimuli, response format, listener and passage familiarity, etc). As is shown in references above, much of the research regarding TE intelligibility was conducted in the mid-to-late 1980's, when the procedure was first introduced and gained in popularity. Since that time, very few new investigations have been conducted to determine the intelligibility of the current population of TE speakers. It is important to note that since the mid-to-late 1980's, many changes and improvements have been made to function, shape, and size of the voice prostheses, making it difficult to generalize previous intelligibility research to the current generation of TE speakers.

Until the 1990's, all voice prostheses were exdwelling style devices to be cleaned and removed by the individual or caregiver (Rajashekar et al., 2009). The early 1990's brought the advent of indwelling devices that are inserted and removed by a physician or SLP, making individuals with limited motility and functioning still eligible to use TE speech (Rajashekar et al., 2009). As well, a hands-free device has been made available in recent years, removing the need for manual occlusion of the stoma when speaking (Rajashekar et al., 2009). This valve consists of a pressure-sensitive diaphragm that is open during normal respiration. During speech production, the exhalatory air pressure increases, causing the valve to close and air is diverted through the TEP prosthesis into the esophagus for speech production (Gress & Singer, 2005). These two significant changes, coupled with the continuous increase in new manufacturers entering the voice

prostheses market, has significantly changed the products that are now being used by TE speakers as compared to the devices used when much of the intelligibility research was conducted during the inception of the procedure.

Up to this point, very little research has been conducted with a focus on the specific patterns of increased and decreased intelligibility that may exist across TE speakers. Results from a study by Doyle, Danhauer, and Reed (1988) have shown that intelligibility issues commonly arise in the areas of voiced-voiceless distinctions of consonants, as well as for fricatives and affricates. The issue surrounding the voiced-voiceless distinction 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 shut-off of voicing in the PE segment. As well, a study conducted by Doyle and Haaf (1989) found post-vocalic consonants to be more intelligible than their pre-vocalic counterparts. This study also found voiced-voiceless confusions and a manner of articulation hierarchy similar to that found by Doyle et al. (1988). More recently, Searl, Carpenter, and Banta (2001) evaluated the intelligibility of stops and fricatives in TE speech. Their findings were also consistent with the two studies previously mentioned (Doyle et al., 1998; Doyle & Haaf, 1989) in that the most common errors were confusing voiced phonemes for voiceless phonemes. At this time, these mentioned studies make up the only known research analyzing the specific errors patterns in TE speech.

The lack of current intelligibility research can be attributed, at least in part, to the relatively spontaneous acquisition of TE speech, and the fluent nature of the sound signal. As previously stated, TE speech has shown to be superior to the other alaryngeal methods

in acceptability, 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. As a result, intelligibility has been essentially overlooked in past years. This is unfortunate and troubling as dissemination of information regarding the specific intelligibility issues to SLPs gives them the ability to tailor therapy sessions with TE speakers around the known errors patterns, potentially creating more intelligible speech for each individual. The ability to refine speech patterns with resultant increases 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 and environments.

Significance of 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 quality of life (Eadie & Doyle, 2004; Karnell et al., 2000; Meyer et al., 2004; Terrell et al., 2004). Quality of life (QoL) is defined by the World Health Organization as:

> *An individuals' perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad ranging concept, incorporating in a complex way a person's physical health, psychological state, level of independence, social relationships, personal beliefs and relationship to salient features of the environment.* (WHO, 1998, p. 17).

This concept of an individual's 'quality of life' plays an important and prominent role in laryngeal cancer due to the unique challenges these individuals face. However, as previously stated, studies have shown speech and communication to be one of the biggest predictors of perceived quality of life 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 of recreation, perception of chewing and swallowing, willingness to eat in public, and reported normalcy of diet. This decreased ability to participate in normal daily activities increases the potential for disability among these individuals. Lower speech intelligibility was also associated with a greater number of altered QoL parameters when compared to their more intelligible counterparts. As well, Karnell, Funk, and Hoffman (2000) evaluated survivors of upper aerodigestive tract cancer and found that speech and eating domains best predicted self-reported QoL scores, further reinforcing the importance of speech rehabilitation in laryngectomized individuals. On a similar note, Rogers, Laher, Overend, and Lowe (1998) found that oropharyngeal cancer survivors rated speech to be among the top three most important QoL domains. Finally, previous research has shown TE speech intelligibility and acceptability to be positively correlated with one another, indicating that speech that is highly intelligible tends to be perceived as also 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 quality of life. The evidence presented in the studies above show that a relationship between highly intelligible speech and

increased QoL exists among laryngeal cancer survivors. This issue is unique to individuals with head and neck cancer and, therefore, deserves attention in order for these individuals to achieve the best possible QoL. This, coupled with the fact that the fundamental purpose 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 research in this area is so needed.

Statement of the Problem

Within this review, it has been noted that limited research has been conducted on TE intelligibility since the late 1980's due at least in part to its spontaneous acquisition and superior acoustic properties, allowing intelligibility to be overlooked. This is troublesome when we consider the significant changes and improvements that have been made to prostheses size, shape, function, and manufacturer since this time. Given that research has been conducted proving the impact speech and communication has on an individual's QoL, it is important that updated research on TE intelligibility be conducted so SLPs and physicians have the most current and up-to-date information outlining the intelligibility patterns of current TE speakers. It is our hypothesis that intelligibility in TE speakers has increased since research was on the subject was first introduced, but that work is still needed in order for each individual to achieve the most intelligible speech. In order to conclude this, this investigation will determine if the intelligibility issues for TE speakers previously documented in the literature are still present in the population, and if new challenges have presented themselves since previous research was conducted. We will also assess overall intelligibility of the TE speakers to allow for comparison to

previous research. Dissemination of this knowledge will allow health care professionals to better structure their treatment and therapy for each individual, giving TE speakers the opportunity to achieve highly intelligible speech, leading to rewarding communication, participation in society, and increased quality of life.

CHAPTER TWO

Methodology

Participants

Speakers

Thirteen male and 2 female (n=15) participants between 39 and 84 (mean = 63.50 years of age served as speakers for this study. The mean age of the male participants was 66.80 years of age (range = 50 to 84) and the mean age of female participants was 51.30 years of age (range = 39 to 60). All participants were native English speakers. As well, all had been previously diagnosed with laryngeal cancer and had undergone total laryngectomy. Additionally, all speakers had undergone TE puncture voice restoration. Each speaker was at least six months postlaryngectomy and TE puncture voice restoration at the time the speech sample was obtained. The participants ranged from 2 to 27 years postlaryngectomy. Individuals with primary and secondary puncture procedures were both included. TE speech served as each speaker's primary mode of alaryngeal communication, although some speakers also used alternate methods (i.e., electrolarynx or esophageal) at times as a secondary communication mode. Informed consent was obtained from all speaker participants prior to the beginning of any speech recordings.

Listeners

Three male and 15 female (n=18) normal hearing adults ranging in age from 21 to 49 years of age (mean=24.10) served as listeners for this study. The mean age of male participants was 23.30 years (range=20.50 to 25.20 years) and the mean age of female participants was 23.70 (range=21.10 to 48.11 years). All participant listeners had no

previous experimental experience with TE speech at the time of participation in the study. As well, all participant listeners were free from any self-identified hearing impairments and were able to read and write English. Participants were recruited voluntarily through undergraduate courses in the Health Studies and Communication Sciences and Disorders programs. Informed consent was obtained from all listeners prior to beginning the research task.

Experimental Stimuli

Speakers

Speaker stimuli were composed of real English words that had been generated from and used in a previous study evaluating speech intelligibility in individuals using electrolaryngeal speech as their primary mode of communication. Published in 1985, the paper, entitled *"Electrolaryngeal Speech Produced By Laryngectomized Subjects: Perceptual Characteristics"* by Weiss and Basilli was comprised of 66 consonant-vowel-consonant (CVC), monosyllabic words with each phoneme in the English language represented equally in both word-initial and word-final positions. Due to phonological rules of English, / ŋ / was omitted word-initially, /h/ and /w/ were omitted word-finally, and /□/ was omitted entirely (see Appendix B).

Listeners

Listener stimuli were the monosyllabic words recorded by the TE speakers described above. Each participant listener was provided with sheets of paper containing numbered, blank lists. This served as an answer sheet for the participants to individually transcribe each word heard during the listening task in English orthographics.

Procedures

Speakers

Speakers were recruited voluntarily through two individual sites. Firstly, at the 2011 International Association of Laryngectomees meeting in Kansas City, Missouri, and secondly, at a surgical clinic at Victoria Hospital in London, Ontario. If interested, potential participants were given a letter of information to consider participation, and if agreed, they were also asked to sign a consent form (see Appendix C). Any questions regarding the study were addressed prior to the start of the experimental procedure. Participants were then asked to read through the word list, saying each word out loud, into a Shure PG81 professional quality, cardioid condenser instrument microphone. All samples were recorded at a sampling rate of 48KHz using SonaSpeech (Kay Pentax, NJ, USA). Participants were also asked to complete a patient demographic sheet in order to obtain general information about treatment history, complications of surgery, associated treatments (e.g., radiotherapy or chemotherapy, etc.) and prosthesis use (see Appendix E). The microphone was placed eight inches away from the participant's mouth, affixed to a stand, to ensure any background noise was not picked up on the recordings. All recordings took place in a quiet room, free from background noise. If words were mispronounced at any point during the task, participants were stopped, instructed on proper pronunciation and asked to continue where they left off. These instances were deleted during analysis of the recordings. The entire recording task required approximately 15 minutes to complete.

After the recording of speech stimuli was completed, each word list was segmented into six separate audio files of 11 words each, using the Amadeus Pro speech editing software (HairerSoft, Kenilworth, UK). These files were saved systematically, according to the speaker and portion of the word list it corresponded to (i.e., 1-15 according to speaker, lists A-F according to the segment of the word list). Periods of silence, each five seconds in length, were also added between each stimulus word to allow listeners time to record their responses between each utterance without needing to pause or stop the track.

A pseudo-randomization listening sequence was then created in order to ensure each speaker, as well as each stimulus word, was represented an equal number of times and at equal positions throughout the listening task. This was done to limit any possible learning or place effects that could occur if unequal representation were to exist. As a result, each listening sequence was comprised of 45 audio files, with 11 words in each file. This meant that each listener was exposed to one half of the entire set of stimulus words (i.e., 33 items) recorded by each speaker. The research team made the decision to expose each listener to only half of the stimuli to reduce the amount of time needed to complete the study. It was our concern that listeners would become fatigued and lose focus if they were required to listen to and transcribe 90 lists of 11 words, a task that would have taken over two hours. To account for the decrease in total stimuli exposure for each listener, we increased the number of listener participants to 18, a number that ensured all stimuli from each speaker was represented an equal number of times across the listeners.

A second randomization sequence of stimuli (n = 99) was also created for each listener to allow agreement to be assessed. Each sequence contained nine audio files the participant was previously exposed to during the listening task. This meant 20% of the stimuli were being repeated to allow for agreement to be analyzed. Similar to the main randomization sequence, the sequence was set-up to ensure each stimulus word was represented an equal number of times, as well as being distributed evenly throughout the listening sequence.

The audio files containing the word lists were then imported into iTunes or Windows Media Player and placed in order, according to the specific randomization sequence for each listener. A separate file was created for each listener for the main listening task, as well as the agreement sequence.

As part of the listening experiment, each listener was exposed to 495 stimuli, as well as 99 agreement samples, totaling 594 stimulus words. Therefore, 10,692 stimuli were presented to the entire group of listeners (594 stimulus words X 18 listeners).

Listeners

Each listener participated in a single listening session that took place in the Voice Production and Perception Laboratory in Elborn College at the University of Western Ontario. Listeners were recruited voluntarily through the Health and Rehabilitation Sciences program, and also through undergraduate courses in Linguistics. Interested individuals were asked to contact the research team through email to set-up a mutually convenient time to complete the task. The entire task took approximately 75 minutes (range = 70 to 85 minutes) and was completed in a single session. Participants listened to

the speech samples through stereo headphones (Sony MDRV-150) at a volume level determined by each listener, attached to a personal computer (Dell Inspiron).

Listening Session

At the beginning of each listening session, each participant was provided a letter of information outlining the nature of the study, and a consent form to sign (See Appendix D). The task was then explained to the participant and any questions that arose were addressed fully. Listeners were then asked to open a folder on the computer corresponding to their participant number that contained the lists of audio files sequentially ordered according to the randomization sequence previously identified. The playlist corresponding to the current listener was then opened on the personal computer and listeners were instructed to press play on the first file when they were ready to begin. The playlist was set-up in such a fashion that once an audio file was played through to completion, the next file would automatically begin, ensuring a continuous flow in the experiment and less distractions for the participant. Listeners were instructed to directly transcribe each word heard using standard English orthographics onto the answer sheet in front of them. Listeners were allowed to stop the files at any point to repeat words or entire segments if necessary but were instructed that once a decision was made, they were not permitted to alter their judgments. Participants were also able to leave a space blank if they were completely unable to determine the word. Each transcription sheet had space for five lists of 11 words and, therefore, had space for 55 responses per page. As well, codes were placed at the top of each list according to the corresponding audio file. This continued until all 45 sets of 11 stimuli were listened to and transcribed. At this point, the same procedure was repeated with the agreement sequence. Following completion of the

general listening and agreement tasks, a debriefing occurred between the researcher and participant that allowed the participant to ask any questions they may have regarding the study, and for the researcher to ensure that the entire task was completed as instructed.

Data Analysis

Once all 18 listeners had completed the listening task, gross, as well as categorical analyses were completed on the collected data. First, gross scoring was conducted to determine the perceptual errors. From this, an overall intelligibility score for each speaker could be determined. Gross scoring allowed the errors to be broken down into 'word-initial', 'word-final', 'vowel', 'whole word', or 'no response' categories. Gross scoring was also conducted for the vowels of each stimulus word to determine if error patterns existed. Next, word-initial and word-final confusion matrices were created for each speaker to record each correct and incorrect phoneme response, allowing for distinctive feature analysis. The individual matrices were then collapsed into one matrix for word-initial and word-final phonemes, incorporating all speakers. The collapsed matrices were also further collapsed to show categories of manner of articulation.

Agreement Assessment

Inter- and intra-rater agreement was also assessed in the present investigation. Both inter and intra-rater agreement was determined through direct sample-by-sample analysis. For intra-rater agreement, word lists transcribed in the agreement portion of the task were compared to transcriptions from the main task to determine if inconsistencies exist. Agreement was then hand calculated by the following formula: number of consistencies/total number of responses x 100. This resulted in a percentage of agreement

among the listeners. Lower agreement could indicate a learning effect occurred among listeners throughout the course of the task.

Due to the nature of the listening task, and the fact that listeners were exposed to only half of the stimuli, inter-rater agreement was determined in two ways, with two groups. Each group consisted of nine listeners, with each listener from their respective group having at least 15% of the stimuli in common. Common stimuli were then hand selected across listeners and evaluated for response consistency. The same formula that was used in the intra-rater agreement portion was again used to determine a percentage value: number of consistencies/total number of responses x 100. This was done independently for both groups, resulting in two separate inter-rater agreement values.

CHAPTER 3

Results

This chapter presents the results obtained on intelligibility measures for the group of 13 male and 2 female TE speakers. Gross perceptual errors were analyzed to determine the predominant location of errors within the stimulus words (i.e., word-initial consonant, word-final consonant, vowel, and whole word). Errors were further analyzed using confusion matrices to identify increased or decreased areas and/or patterns of intelligibility for each individual, as well as the entire group. Intra- and intra-rater analysis was also covered.

Demographic Information of TE Speakers

Relevant information regarding the TE speakers who served as participants in this study can be found in Table 1. A wide range in the number of months since laryngectomy existed between participants, with the shortest being 15 months and the longest being 322 months (26 years, 8 months) (mean=102.25 months) postoperative. The majority of participants (n=10) had their TE puncture procedure completed at the time of their laryngectomy, that is, as a primary surgical procedure. As well, at the time of recording, the majority of participants were using the InHealth TE puncture voice prosthesis; the size (length and French diameter sizing) varied among participants, but the most common prosthesis diameter was 20Fr.

Table 1 *Demographic Data of Speaker Participants*

Variable	Men (n)	Women (n)	Total
Number of Participants	13	2	15
Age (years)	Mean 66.80 (50-84)	Mean 51.30 (39-60)	Mean 63.50 (39-84)
Time postlaryngectomy (months)	Mean 117.10 (15-322)	Mean 28.0 (27-29)	Mean 102.25 (15-322)
Time of TE puncture procedure			
Primary (at time of surgery)	9	1	10
Secondary (after surgery)	3	1	4
Unknown	2	0	2
Type of voice prosthesis			
Blom-Singer (InHealth)	11	2	13
Atos (Provox)	2	0	2
Other	0	0	0
Size of voice prosthesis			
11Fr	1	0	1
16Fr	4	0	4
17Fr	1	0	1
20Fr	5	2	7
Other/Unknown	2	0	2

Individual and Overall Intelligibility Scores

Overall, word intelligibility scores ranged from 54% to 89%, with an overall average of 71%. For the purpose of this study, gross intelligibility was determined through the following formula: # of correct responses/total # of responses x 100. Average intelligibility, as well as minimum and maximum intelligibility values for each speaker are represented in Table 2. Overall intelligibility analyses were based on a total of 594

listener observations for each speaker (33 observations per speaker x 18 listeners). The 71% overall intelligibility of all 15 speakers was based on 8910 listener observations (33 observations x 15 speakers x 18 listeners).

Table 2 *Average Intelligibility and Range of Individual TE Speakers*

Speaker	Average Intelligibility (%)	Range of Listener Intelligibility[2](min-max %)	Intelligibility Rank
1	65	39-85	11
2	83	73-94	2
3	74	61-91	6
4	76	64-85	5
5	72	42-88	7
6	69	58-88	8
7	64	45-73	12
8	89	79-97	1
9	66	48-82	9
10	61	45-85	14
11	66	39-79	10
12	62	36-82	13
13	54	27-82	15
14	81	52-94	3
15	78	70-91	4
	Average – 71%	Range 54% – 89%	

[2] Minimum and maximum intelligibility values taken from individual intelligibility results taken from each listener.

Gross Distribution of Errors

Overall, the greatest percentage of errors involved phonemes occurring in the word-initial position of the stimulus words, comprising 57% of the errors. This is nearly double the number of errors observed for word-final stimuli, which only comprised 30% of the total errors. The remaining 13% of the errors were split between whole word errors (7.50% of the errors), and no responses (5.50% of the errors). Whole word errors pertained to those listener responses that constituted errors in both the word-initial and word-final phonemes, as well as for the vowel within the CVC stimuli. Errors were placed in the "no response" category when a listener failed to record any sound/word in the space the target word was meant to be recorded. Thus, these responses were represented as omissions in the transcribed responses obtained.

Vowels

Overall, vowels were found to be highly intelligible across all TE speakers. Overall, intelligibility of vowels averaged 99.24% across TE speakers. As well, the confusions that were observed involved all front, unrounded vowels (/i/, /□/, /ɛ/, /e/, /æ/). Of these, the most common confusions involved the low front, unrounded vowel /æ/. This vowel was confused most often with middle, front, unrounded vowels /ɛ/ and /e/. Less often, /æ/ was replaced with the high, front, unrounded vowels /i/ and /□/. Less common, but still frequent, was the confusion of the high, front, unrounded vowels /i/ and /□/ with the middle, front, unrounded vowel /e/. Finally, the diphthong vowel /ai/ was

most often replaced by /æ/. However, as previously stated, these vowel confusions were observed in less than 1% of all listener responses (n=66).

Voiced-Voiceless Distinctions

Being able to correctly identify voiced and voiceless consonants is a difficulty that has commonly been reported in the literature related to the intelligibility of TE speakers. In the word-initial position, 50.96% of all errors involved voiceless phonemes being perceived as voiced phonemes (both cognates and open responses). Conversely, voiced phonemes confused for voiceless phonemes made up 24.56% of the errors in the word-initial position. Together, the voiced-voiceless distinction accounted for 75.52% of all errors identified in the word-initial position.

A similar pattern of error, that is, voiceless phonemes being confused for voiced phonemes, was found in the word-final position, although to a smaller extent than what was observed in the word-initial position. Voiceless phonemes confused for voiced phonemes represented 26.15% of the errors, while voiced phonemes confused for voiceless phonemes comprised 10.25% of the errors. Together, 36.40% of all errors in the word-final position involved the voiced-voiceless distinction.

Manner Errors

Listener perceptual data were further analyzed to determine intelligibility by manner classification. Each phoneme was entered into a confusion matrix according to

Figure 1 Word-initial phonemes collapsed into manner of articulation categories.

	-V stops	+V stops	-V fricatives	+V fricatives	-V affricates	+V affricates	nasals	liquids	glides	#	other
-V stops	829	525	48	10	4		3			19	3
+V stops	41	1004	29							8	5
-V fricatives	23	29	1394	63	19	5			12	201	16
+V fricatives		41	272	794			11			22	1
-V affricates	9	8			311	56				8	1
+V affricates						291					
nasals	19	23	12				614			11	2
liquids		7	22	4			4	764	15	4	3
glides		5	8					5	387		
#		9	36	3	2		1				
other	1	2	4	1							

Each cell represents the number of responses elicited from listeners. Stimuli are shown down the left column, while responses from the listeners are across the top of the matrix. Shaded cells indicate correct responses. The symbol # indicates the absence of a consonant, the term "other" indicates responses such as vowels, or consonant clusters.

Figure 2 Word-final phonemes collapsed into manner of articulation categories.

	-V stops	+V stops	-V- fricatives	+V fricatives	-V affricates	+V affricates	nasals	liquids	glides	#	other
-V stops	1167	71	11			2				25	20
+V stops	51	780	3	8			27			17	3
-V fricatives	52	27	1327	132	14	9	2			52	18
+V fricatives	8	8	39	1219						12	7
-V affricates		5			403	32				4	
+V affricates	2	3	6	4	9	404				14	
nasals				4			1145			18	3
liquids		3						866		2	
glides											
#		3	3								
other	4	11		6			3	1			

Each cell represents the number of responses elicited from listeners. Stimuli are shown down the left column, while responses from the listeners are across the top of the matrix. Shaded cells indicate correct responses. The symbol # indicates the absence of a consonant, the term “other” indicates responses such as vowels, or consonant clusters.

TE speaker number and position within the stimulus word (i.e., word-initial or word-final). Individual speaker matrices were then collapsed into master matrices for word-initial and word-final phonemes. Finally, these were further collapsed to reveal manner of articulation errors. Word-initial and word-final matrices showing these error patterns can be found in Figure 1 and Figure 2, respectively. Collectively, 36,720 listener observations were obtained on manner classification (68 stimulus words x 2 phonemes/word x 15 TE speakers x 18 listeners).

Plosives. Plosives were found to be the least intelligible class of phonemes with an overall intelligibility of 80.99%. More errors were found for word-initial plosives, with an average intelligibility of 72.88% across all TE speakers. Conversely, intelligibility was higher in the word-final position, with an average intelligibility of 89.10%; this observation follows the general trend of higher intelligibility for consonants produced in the word-final position. The majority of the errors in the word-initial and word-final positions were confusions between voiced and voiceless phonemes. As well, a great number of errors involved confusing plosives for fricatives. Also of note was the difference in intelligibility found between voiceless and voiced plosives in the word-initial position. Voiced phonemes were found to be more intelligible than voiceless phonemes, with intelligibility for voiced phonemes at 92.36% as compared to an overall intelligibility for voiceless phonemes of 57.50%.

Fricatives: Next to plosives, fricatives were found to be the second least intelligible manner class at 81.19% intelligibility overall. The same trend was observed with greater intelligibility in the word-final position than word-initial position. Word-final intelligibility was 87.01%, word-initial intelligibility averaged

out to 75.37%. Word-initial intelligibility indicated that 28% of the errors involved omissions, namely the voiceless fricative /h/. The remaining word-initial errors mostly involved the confusion of fricatives for plosives, and the voiced-voiceless distinction, with the majority of these errors being voiceless phonemes being substituted for voiced phonemes. For word-final consonants, 35% of the errors involved the confusion of voiceless phonemes for voiced phonemes. The remaining errors were dispersed between voiced for voiceless confusions, substitutions of fricatives for plosives, and omissions. In addition, voiced phonemes were found to be approximately 10% more intelligible than their voiceless counterparts, in both the word-initial and word-final positions.

Affricates: Data collected indicated affricates to be highly intelligible, with an average intelligibility of 89.55%, surpassing that of both plosives and fricatives. As well, intelligibility was found to be fairly constant across phoneme position, with 88% intelligibility in the word-initial position and 91% intelligibility in the word-final position. A distinction can be found between the voiced and voiceless phonemes in the word-initial position, with 100% intelligibility in the voiced affricates, but only 79% observed intelligibility in the voiceless counterpart. Of the errors observed, the majority involved the voiceless affricate (t□) being confused for its voiced counterpart (d□). This was observed in both word-initial and word-final positions.

Nasals: Nasal sounds also were found to be highly intelligible with an average intelligibility of 94.01%. The trend of higher intelligibility among word-final phonemes was also observed, averaging 97.86% across speakers in the word-final position, in comparison to 90.16% intelligibility in the word-initial position. Of the

errors observed, in the word-initial position, most errors were found to involve nasals being substituted for stops and fricatives. In the word-final position, most errors involved omissions, namely involving the phoneme /ŋ/. Another finding indicated nasals being confused for other nasal phonemes (i.e., /m/ being substituted for /n/ and /ŋ/, and vice versa).

Liquids: Liquids (/r/ and /l/) were found to be the most intelligible class in the word-final position, at 99.43%. In the word-initial position, liquids were observed as the second most intelligible class, behind glides, at 92.83% intelligibility across speakers. This resulted in an average intelligibility of 96.13%, the highest overall intelligibility for the various manner of articulation categories. In the word-initial position, the largest amount of errors involved liquids being confused for both voiceless fricatives and glides. Together, these two confusions (substitutions for voiceless fricatives and glides) accounted for 62.70% of the errors observed involving liquids. The remaining errors were evenly distributed across voiced plosives, voiced fricatives and omissions.

Glides Glides were found to be the most intelligible manner of articulation class in the word-initial position, averaging 95.56% intelligibility across speakers. The only errors recorded involved the substitution of glides for liquids, voiceless fricatives, or voiced plosives, with an even distribution across each category (see Figure 1). Phonological rules of English prohibits the use of /w/ in the word-final position, therefore, glides were only considered in the word-initial position.

Omissions

Perceptual omissions as identified by listeners are represented by the symbol "#" in the confusion matrices presented in Figures 1 and 2, and included stimulus words that received no response, or, if the target phoneme was absent in the response by the listener. The most common omission involved the voiceless fricative /h/ in the word-initial position. The stimulus words "hun" and "hung" most often received responses of "un" and "ung", omitting the /h/ phoneme. This phoneme alone comprised 65.57% of all omissions in the word-initial position. Phonological rules of English prohibit the use of /h/ in word-final context, therefore, patterns involving this phoneme were only considered in the word-initial position. Omissions were fairly evenly distributed throughout the remaining manner categories, with the exception of voiced affricates and glides, which revealed no omissions in the word-initial position.

The omissions recorded in the word final position followed the same general pattern as that observed for their word-initial counterparts, with voiceless fricatives receiving the most omissions. The remaining omissions were evenly distributed among the remaining categories, although to a smaller degree due to the smaller number of errors observed in the word-final position. Unlike the word-initial omissions, all manner categories had observed omissions.

Agreement Analysis

Inter- and intra-rater agreement analyses were also performed. Intra-rater agreement was determined by comparing the responses received in the agreement samples to responses received by the same stimuli from the general listening task.

This analysis revealed an average of 87.21% agreement across listeners, with absolute values ranging from 77.78% to 96.97%.

As stated in the methods section of the investigation, inter-rater agreement was determined in two separate groups, due to the nature of the study. Both group A and B were each composed of nine different listeners (half of the total listener participants in each group). Every listener from each of the two groups had 15% of the stimuli (n=77) in common, which were then used to determine agreement among the listeners. Hand analyses determined inter-rater agreement to be 79.08% for Group A, and 78.21% for Group B. Thus, good levels of consistency for perceptual judgments of the phonemes represented was observed across the listeners who participated in this study. Consequently, both intra- and inter-rater agreement measures indicate good-to-excellent consistency in the data acquired as part of the present investigation.

CHAPTER 4

Discussion

The present investigation was designed to explore the intelligibility of males and females who use tracheoesophageal (TE) speech as their primary mode of postlaryngectomy verbal communication. More specifically, the objectives of this project were to: (1) determine overall and individual speech intelligibility for speakers, (2) assess whether patterns exist between speaker age, type of prosthesis/size, length of time postlaryngectomy, and speech intelligibility. (3) analyze errors to determine their rate of occurrence and distribution (i.e., occurring word-initially, word-finally, etc.), and (4) determine if patterns of increased or decreased intelligibility exist according to manner of production . The discussion to follow will address each of the areas outlined above in detail. Inter and intra-rater agreement analyses will also be discussed. As well, clinical implications, and the limitations of the present study will be explored. Finally, directions for future research and overall conclusions will be presented.

Overall and Individual Speech Intelligibility

Analysis of the data indicated an overall intelligibility of 71% across all 15 speakers, with the most intelligible speaker at 89%, and the least intelligible at 54%. This number was determined by whole word scoring. These values are consistent to those found in a study by Doyle, Danhauer, and Reed (1988) that reported a mean overall intelligibility score of 65% for the TE speakers used in their study. These results are also comparable to those found by Doyle and Haaf (1989) and Searl et al.

(2001). Doyle and Haaf (1989) found an average intelligibility of 75.17%, with scores ranging from 69.8% to 77.8% across four speakers. Searl et al. (2001) reported slightly lower intelligibility, at 62.30% on average, with a range of 43.90% to 73%.

Conversely, when compared to multiple other studies assessing the overall intelligibility of TE speakers, the scores found in the present study are approximately 20% lower. Studies by Pindzola and Cain (1988), Tardy-Mitzell et. al., (1985), and Blom, Singer, and Hamaker (1986) reported overall intelligibility scores of 93.20%, 93% and 91.51%, respectively, scores that are remarkably higher than those observed in this study. This finding may be due to factors that are related to study design, including stimuli used and the method of measurement.

In the above three cited studies (Blom, Singer, & Hamaker, 1986; Pindzola & Cain, 1988; Tardy-Mitzell et al., 1985), listeners were given multiple choice responses sheets with six choices for each stimulus word presented, known as a closed-set format response task. This could potentially inflate scores as it forces the listener to choose an answer, and creates the possibility of correctly identifying the word through 'guessing', even if the stimulus word was not produced correctly. This forced-choice method also removes the possibility of whole-word errors and omissions responses from the listener.

In contrast to a closed-set format, the present investigation used an open-set response paradigm. In doing so, this response format allows for greater detection of production errors as the listener must identify the correct word without any cues beyond the sound signal. Furthermore, the studies referenced above that reported

similar intelligibility scores to those found in the present study (Doyle et al., 1988; Doyle & Haaf, 1989; Searl et al., 2001) also used an open-choice response format. This phenomenon has been well documented in the literature with multiple studies comparing intelligibility scores using both open and closed formats (Vigouroux & Miller, n.d.; Yorkston & Beukelman 1978,1980). Each of these studies found that closed format scores were significantly higher than those of open format response scores. Therefore, caution should be taken when attempting to make comparisons between the two methods. In order to fully understand the intelligibility issues troubling the present population of TE speakers, specific investigation into the data needed to be conducted.

Also of note is the range in intelligibility among a subset of the TE speakers (as seen in Table 2). Many of these speakers fell amongst the least intelligible speakers of the group. This could be attributed at least in part to error among the TE speaker participants of the study. Individuals with lower intelligibility may also have voices further away from "normal" in terms of quality. This may have caused some listeners to focus on the quality of the voice rather than the word that was being spoken, leading to varying intelligibility scores. As well, certain listeners may have found the procedure tiresome and lost focus at points during the task, increasing the chance of incorrect responses to be transcribed.

Gross Error Distribution

Gross analysis of the full intelligibility data collected from each listener allowed for the distribution and subsequent frequency of errors to be determined. This

resulted in 57% of the total errors being identified in the word-initial position, 30% occurring for word-final phonemes, 7.5% of the errors involving the entire target word (i.e, word-initial phoneme, vowel, and word-final phoneme), and finally 5.5% involving no recorded response (a frank omission). Vowel errors were also analyzed, but accounted for less than 1% of all observed perceptual errors.

This observation of nearly double the number of errors involving the word-initial phoneme is consistent with findings by Doyle and Haaf (1989) who reported approximately 20% greater intelligibility for word-final phonemes across the four TE speakers used in their study. As originally posited by Doyle and Haaf (1989) the decrease in errors word-finally may be due, at least in part, to context dependent variables. All word-final phonemes followed a vowel, making acoustic cues available to the listener that may assist them in identifying the correct phoneme. As well, the lack of carrier phrase that would have proceeded the word-initial phoneme, in this study, eliminated any potential for additional acoustic cues to assist the listener in identifying the correct phoneme in a word-initial position, also potentially contributing to the increase in errors seen here.

The existence of whole-word errors, and no responses, may also be attributed to the open format response paradigm that was selected to capture the responses made by the listeners in this study. Had a closed, forced-choice format been used, the participant would have been forced to select an answer that differed from the target word by one phoneme, eliminating these categories.

Vowels

As stated in the previous results chapter, vowels were found to be highly intelligible, with errors found in less than 1% of responses. Errors were found to be systematic though, with only front, unrounded vowels involved. The high intelligibility of vowels revealed in this investigation may be attributed to their fundamental acoustic properties. That is, all vowels are voiced phonemes, and high in intensity with minimal to no constriction of the vocal tract during production. Therefore, less control over the PE segment is needed to produce these phonemes, decreasing the potential for error and confusion. As well, vowels are typically held constant for a relatively long duration within CVC syllables, approximately 100 milliseconds (Blood, 1981). Another potential reason why such high vowel intelligibility was found in the present investigation again may be due to context. All vowels were preceded and followed by a consonant, providing acoustic cues that could have helped the listener predict the correct sound. If the word-initial and word-final phonemes of the word were correctly identified by the listener, due to rules of the English language, only certain vowels could occupy the space in order to create a true English word. This decreases the number of the vowels the listener has to choose from and increases the potential of transcribing the correct vowel. However, vowel intelligibility has not been comprehensively assessed in TE speaker and as a result, likely provides a rich area for future study.

Voiced-Voiceless Distinction

The present study found that 50% of all errors involved voiced-for-voiceless distinction confusions. As well, for word-initial consonants, twice as many errors involved substituting voiceless phonemes for voiced phonemes in comparison to the

opposite. The same pattern was observed in word-final position, although to a greater extent, as three times as many errors were voiced-for-voiceless, versus voiceless-for-voiced confusions. This finding is consistent with studies conducted by Doyle et al. (1988, 1989), Miralles and Cervera (1995) and Searl et al. (2001). More specifically, Searl et al. (2001) also found that 50% of the errors observed across TE speakers that participated in their study involved the voicing parameter, with the majority being voiced-for-voiceless confusions as well. Difficulties producing voiceless consonants, and the subsequent substitution of these phonemes for voiced consonants, has been well documented in the alaryngeal speech literature (Doyle et al., 1988, 1989; Miralles & Cervera, 1995; Searl et al., 2001; Searl & Carpenter, 2002).

Two specific issues have been explored as possible reasons as to why voiceless consonants are produced so poorly by TE speakers. The first issue involves the PE segment, or sound source for TE speakers. Unlike the vocal folds, the PE segment does not have the ability to quickly abduct and subsequently, rapidly devoice, despite its vibratory capabilities (Searl & Carpenter, 2002). This lag in the cessation of the voicing feature has the potential to increase the likelihood of a voiced phoneme being produced when a voiceless phoneme is intended. This lag in cessation means that voicing continues on to adjacent phonemes, at times when not intended. Therefore, phonemes that were intended to be voiceless then have an added voicing feature, leading to the perception of a voiced phoneme.

One of the main issues in regard to the voiced/voiceless distinction involving stops/plosives is voice onset time (VOT). VOT is measured beginning at the release burst of the plosive and ending when source vibration begins for the adjacent vowel

(Ferrand, 2007). VOT is known to be remarkably longer (i.e., delayed) in voiceless stops (25 – 100 milliseconds) than for voiced stops (0 – 20 milliseconds). This timing feature is used as a primary factor to distinguish between such voiced and voiceless phonemes in laryngeal speech (Ferrand, 2007). Studies have been conducted on VOT in esophageal and TE speakers and have unanimously found decreased VOT duration in voiceless stops (Christensen, Weinberg, & Alfonso, 1978; Connor, Hamlet, & Joyce, 1985; Robbins, Christensen, & Kempster, 1986; Searl & Carpenter, 2002). This decrease in temporal length brings VOT values to a range closer to that of voiced phonemes, increasing the potential for voiceless phonemes to be confused for voiced ones.

A further element of the VOT issue in alaryngeal speakers pertains to the influence of the PE segment. Reduced motor control and elasticity of the segment diminishes the ability to quickly turn voicing "on and off". This diminished control over the start of voicing can account for the shorter VOTs found in TE speakers. While the PE segment is under some volitional control relative to its tonicity, it is not an adductor-abductor mechanism like the larynx. It is, therefore, not surprising to discover that this phenomenon occurs in both TE and esophageal speakers given that both methods use the PE segment as the postlaryngectomy voice source.

In addition to VOT, other acoustic parameters play a role in signaling the voicing feature. These can include: stop closure duration, fricative noise duration, aspiration noise following burst release, length of adjacent vowel, oral air pressure, and more (Cole & Cooper, 1975; Halle, Hughes, & Radley, 1975; Lisker, 1978; Raphael, Dorman, & Liberman, 1980; Slis & Cohen, 1969). Alterations in any of

these parameters can also impact the voicing parameter in TE speakers, making for additional confusions in voiceless stops and fricatives. Collectively, the impact of these factors likely underlie perceptual confusions that exist for voiced-voiceless cognates. Although VOT and other acoustic parameters were not specifically measured in this investigation, earlier onset of voicing has the potential to explain, at least in part, many of the voiced-for-voiceless stop and fricative confusions found in the present investigation.

Analysis by Manner Classification

Further analysis by manner of articulation classification resulted in the production of a hierarchy of intelligibility according to manner class. Results indicated liquids to be the most intelligible (96.13%) followed by glides (95.56%), nasals (94.01%), and affricates (89.55%), and finally fricatives (81.19%) with plosives (80.99%) as the least intelligible class of sounds. Each of the manner categories investigated in the present study had higher intelligibility scores in the word-final position; in some cases this increase was as large as a 15% difference (as was observed with plosives and fricatives). This hierarchy of intelligibility, as well as the discrepancy in intelligibility between word-initial and word-final phonemes is consistent with previous work published by Doyle and Haaf (1989). This study by Doyle and Haaf (1989) also observed plosives and fricatives to be the least intelligible manner classes (at 83% and 80.50% intelligibility, respectfully), while liquids and glides were found to be the most intelligible (99.50% intelligibility). The difference in intelligibility between word-initial and word-final phonemes in this study was as great as 21%. This discrepancy between word-initial and word-final

phoneme intelligibility is due, at least in part, to context dependent variables. Word-final phonemes were preceded by a vowel, lending acoustic cues to the final phoneme, making it potentially easier to interpret. These acoustic cues were absent in the word-initial phonemes, thereby reducing the intelligibility.

Stops and Fricatives

As noted, stops and fricatives were found to be the least intelligible manner of articulation classes in the present study, at 80.99% and 81.19% intelligibility overall. These findings are consistent with previous studies conducted by Doyle and Haaf (1989), Doyle et al. (1988) and Searl et al. (2001). As mentioned previously, many of the confusions for both manner classes involved the voiced-voiceless distinction.

A phoneme error that deserves particular attention is the voiceless fricative /h/, which comprised 65.57% of all word-initial omissions. We can speculate this is a result of the altered anatomy and physiology of the PE segment of TE speakers. As mentioned earlier, the lack of fine motor control of the PE segment makes precise control of vibration more challenging, increasing the potential that the PE segment may be set into vibration earlier than intended (i.e. during the production of the /h/ phoneme). As well, all /h/ phonemes occurred in the word-initial loci due to language rules and were, therefore, void of any acoustic cues that may have assisted in identifying the phoneme (e.g., the ability to generate laminar flow through a tight, highly controlled sound generating aperture).

Liquids and Glides

Liquids and glides were found to be the most intelligible manner classes (overall intelligibility of 99.43% and 95.56%, respectively). This is consistent with the previous literature exploring manner of articulation intelligibility in TE speakers (Doyle et al., 1988; Doyle & Haaf, 1989; Searl et al., 2001). High intelligibility of these manner classes can be attributed to their production. Both involve minimal constriction of the vocal tract, thus requiring less control by and influence of the PE segment on sound production. As well, glides, the most intelligible class in the present investigation, are also known as "semi-vowels" as they are phonetically similar to vowels. It is, therefore, logical that this category would be perceived as highly intelligible, as vowels accounted for less than 1% of the errors identified in the present investigation.

Patient Demographics and Overall Intelligibility

Demographic information on all TE speakers was collected at the time of voice data recording in order to gather general information regarding age, date of laryngectomy, prosthesis type and size, and time of TE puncture procedure (primary versus secondary surgery). This information allowed us to explore if particular patterns of increased or decreased intelligibility could be attributed to any of these variables.

In evaluating the data, the TE speakers with the highest intelligibility tended to be the individuals who's laryngectomy had been completed in excess of 10 years prior. It could be argued that an increased number of years postlaryngectomy means greater time spent in therapy, greater time using TE speech, and therefore refining it,

and a larger period of time for the surgical site and related structures involved in the speech process to heal completely. Any or all of these factors could contribute to more intelligible speech. With this being said, one of the most intelligible speakers was also one of the individuals with the shortest amount of time postlaryngectomy (two years).

The opposite pattern was also found when further assessing the demographic information gathered. Many of the least intelligible speakers were found to be those with the shortest period of time postlaryngectomy (one to five years). As with the most intelligible individuals, one of the least intelligible speakers was found to be one the longest postlaryngectomy, at approximately 15 years. This time postlaryngectomy alone does not appear to be systematically associated with intelligibility as measured in the present investigation.

These anomalies involving length of time postlaryngectomy and intelligibility reinforces the notion that differences between individuals, namely in the postsurgical anatomy and physiology of the alaryngeal voice tract in general and of the PE segment in specific, plays an important role in the intelligibility of TE speech. Additionally, these variations also may not be able to be completely influenced by therapy and experience using TE speech alone.

Prosthesis type (manufacturer), size (length and French sizing of diameter), and the time elapsed since the TE puncture were also reviewed for corresponding intelligibility patterns. No obvious consistencies could be found among these categories as each variable was fairly evenly distributed among both highly

intelligible and less intelligible speakers. This suggests that it was unlikely any of these variables contributed substantially to the intelligibility of each speaker in the present sample. In turn, this may further reinforce the fact that differences among individuals has the potential to significantly impact speech intelligibility. As well, although not specifically looked at in this study, the overall health and prior medical history of the individuals may have contributed to overall intelligibility in some cases. Unrelated medical complications such as prior strokes, illnesses, medications, etc. all have the potential to impact all or a subset of the speech system, thereby potentially impacting the intelligibility of the speaker. However, combinations of factors may result in patterns that cannot be identified at present. Thus, while no patterns seem to have emerged within the present speaker sample, this suggestion requires empirical confirmation with a greater number of speakers in order to support the external validity of the present findings.

Additionally, there is a potential presence of confounding variables within the study, particularly involving the TE speaker participants. Of note is the prior treatment and/or surgical procedure of the individuals. Extensive surgeries involving surrounding associated structures have the potential to further alter the physiology and subsequent function of the PE segment and oral cavity, thereby potentially impacting the intelligibility of an individuals speech. As well, prior treatment history, such as exposure to chemotherapy and radiation also has the potential to change the function of the speech production system, and therefore also potentially impact an individual's speech intelligibility. Lastly the age of the TE speakers introduces the potential for further confounding variables. As discussed above, comorbidities or

previous medical complications also the potential to impact intelligibility and therefore the findings of the present study.

Agreement Analysis

Intrarater agreement analysis indicated 87.21% agreement across listeners, ranging from 77.78% to 96.97%. Further assessment of the data indicated that in a few cases, stimulus words that were transcribed incorrectly the first time they were heard by the listener were transcribed correctly the second time during the agreement sequence at the end of the task. This may indicate that some learning effects took place during the investigation. Due to the nature of the task, listeners were exposed to each stimulus word seven to eight times, making it possible that listeners learned parts of the word list and were able to anticipate which word would be next in the sequence, or what word they should be hearing versus what they actually heard. However, prior to the beginning of the task, listeners were instructed to transcribe only what they heard, without influence from prior stimuli. As well, debriefing and discussion after completion of the task was also employed in order to try and deter any learning effects.

Interrater agreement analysis resulted in 79.08% and 78.21% agreement for the two separate groups analyzed. This shows that good levels of consistency were observed between the listeners for the experimental task.

Summary

Thus far, overall and individual intelligibility results have been discussed. As well, specific patterns of intelligibility; including error distribution, hierarchy of

intelligibility among manner classes, vowel errors, and omissions have been discussed in detail. Based upon the above, the next section will address potential clinical implications of the present findings, followed by limitations of the current study, and directions for future research.

Clinical Implications

From the time TE speech was first introduced by Blom and Singer in 1979, many studies have been published evaluating the acceptability, intelligibility and acoustic properties of TE speech, as well as in comparison to its alaryngeal speech alternatives (esophageal and electrolaryngeal speech) (Blom et al., 1986; Bridges, 1991; Christensen & Dwyer, 1990; Clark & Stemple, 1982; Cullinan et al., 1986; Doyle et al., 1988; Miralles & Cervera, 1995; Pindzola & Cain, 1988; Robbins, 1984; Robbins et al., 1984; Tardy-Mitzell et al., 1985). Despite this body of literature, relatively limited research has been conducted in the past 20 years regarding the intelligibility of TE speech, with the vast majority of the literature on this subject having been published when the TE puncture procedure was first introduced (Blom et al., 1986; Clark & Stemple, 1982; Cullinan et al., 1986; Doyle et al., 1988, Doyle et al., 1989; Robbins et al, 1984, Smith & Calhoun, 1994; Tardy-Mitzell et al., 1985; Williams & Watson, 1985). Furthermore, limited research directed toward identifying patterns of errors that result in reduced intelligibility in TE speakers has been conducted (Doyle et al., 1988, 1989, Doyle & Haaf, 1989, Miralles & Cervera, 1995). The current investigation determined that despite advances and changes in prostheses type and management, average intelligibility rates have not increased for TE speakers and the error patterns leading to lower intelligibility follow those reported by studies

conducted in the late 1980's. As a result, we may now postulate that much of the intelligibility issues of TE speakers are intrinsic to the individual and not a result of type or size of the prosthesis they are using. This is valuable for to the Speech-Language Pathologist (SLP), who has the potential to influence these issues with direct therapy they may provide to TE speakers.

Jongmans, Rossum, van As-Brooks, Hilgers, and Pols (2008) state that SLPs rarely focus their therapy on speech quality. This may be a result of the spontaneous acquisition associated with TE speech. Much attention may be placed on acquiring functional speech, and once this has been accomplished, teaching prosthesis care. As well, TE speech has been reported by many to be more intelligible than esophageal and 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). This, coupled with the lack of current information regarding intelligibility has the potential to lead to neglect of the intelligibility aspect in therapy. Clearly, reductions in intelligibility do exist and based on the present investigation, a range of intelligibility also exists.

As previously mentioned, a reduction in speech intelligibility has social implications with a potential resultant influence on QoL. Multiple studies have shown that decreased intelligibility leads to altered QoL parameters, less enjoyment of recreation, willingness to eat in public, and more (Karnell et al., 2000. Meyer et al., 2004). As well, more intelligible speech leads to improved communicative ability and independence (Ackerstaff et al., 1994). In addition, environmental factors have the potential to influence communication, for example, in situations of increased

background noise in crowds/public venues, speaking to those with hearing loss associated with age, irregularity of the TE voice to naïve listeners, etc. Any one of these factors has the potential to influence communicative competence in an alaryngeal speaker. Less intelligible speech, coupled with any of these environmental factors has the ability to further impact an individual's communicative ability and effectiveness.

The information presented by Jongmans et al., (2008) shows a lack of formal clinical attention to intelligibility during therapy, combined with the results from the present study that indicate that issues still remain in the area of intelligibility of TE speakers. As such, speech intelligibility concerns continue to warrant clinical attention. Simple intelligibility tests, such as those employed in this study, could be used in clinical practice, giving SLPs the ability to gauge how intelligible each patient is, as well as identify specific areas in need of greater attention. Therapy could then be tailored to each individual in an attempt to achieve as highly intelligible speech as possible.

Limitations to the Present Study

Firstly, the stimulus words spoken by the TE speakers participating in this study were recorded without use of a carrier phrase, a practice that is commonly seen in intelligibility literature. For example, in the study conducted by Weiss and Basili (1985), from which the present stimulus word list was obtained, each word was produced in the carrier phrase "You will write _____". It could be argued that a lack of carrier phrase could decrease the generalizability of the results as conversation

involving the general public is usually conducted in sentence format. Due to the nature of the present study though, and the desire to be able to pinpoint specific patterns of intelligibility, we believed removal of the carrier phrase would remove any potential acoustic cues that could confound the results obtained. Having the ability to explore word-initial and word-final TE productions in a detailed manner does offer the potential to understand both the capacity and limitations of the TE sound source for speech production. Thus, the decision to not employ a carrier phrase was done by design in the present project.

Secondly, despite our best efforts, it is possible that learning effects may have influenced our results to some extent. Due to the nature of the study, and the sheer volume of stimuli presented to each listener, it is possible that some listeners may have started to recognize certain stimuli as they progressed through the perceptual task. If so, this could have potentially influenced judgment of what they heard. However, debriefing was done with each listener after the experimental task to ensure that this was not consciously done; yet it is possible that some listeners may not have been aware of this phenomenon occurring during their participation and transcription of stimuli.

Finally, the potential exists for bias to exist among the speakers and listeners in the present study, with the potential to impact the current findings. A large majority of the TE speakers participating in the current study were recruited from the annual meeting of the International Association of Laryngectomies. Individuals participating in this conference tend to be of the highly motivated and enthusiastic nature, and are therefore more likely to engage in research activities. This may be unrepresentative of

the general TE speaker population as less motivated and less involved individuals were less likely to attend this conference. There was an attempt to counteract this by recruitment at a surgical clinic at Victoria Hospital that sees individuals in all levels of functioning and abilities. Despite this, participation from each site was not equal.

As well, a large gap existed between the age of the listeners (M=24.10) and the TE speakers (M=63.50) in the present study. The fact that most of our listeners were young adults listening to the voices of older adults could potentially introduce the notion of stigma into the data. Stigma towards the voices of these individuals as a result of their age could potentially lead to unrepresentative intelligibility data in some cases.

Directions for Future Research

The present study found that intelligibility of this group of 15 TE speakers is consistent with findings of previous research conducted in the late 1980's, when TE speech was first introduced. As well, patterns of increased and decreased intelligibility have remained fairly constant since this time, indicating that increases in prosthesis technology have not necessarily led to increases (or decreases) in intelligibility. As a result of these findings, multiple recommendations for future research are proposed.

Firstly, valuable information may be gained by using the TE speech samples collected in the current study and acoustically analyzing them for use in future investigations. Examination of the acoustics of the present samples would permit analyses of specific characteristics such as voice onset time, vowel length, and

intrinsic phoneme amplitudes that may combine to aid or reduce speech intelligibility. With this information, we could confirm speculations that abnormalities in these dimensions may have contributed to some of the confusions/errors observed in the present study. Explorations of this type would potentially allow data collected in the present investigation to be more generalizable to previous studies (Blom et al., 1986; Hillman et al., 1998; Robbins, 1984; Robbins et al., 1984; Searl & Carpenter, 2002). However, the ability to discern acoustic and perceptual relationships in a non-normal voicing source such as that of TE speech is a time- and labor-intensive endeavor requiring systematic investigation and replication. As such, efforts that seek to focus on more limited stimulus sets (e.g., plosives) may be more feasible. In such situations, these types of data may be transferred to clinical training tasks that could then be experimentally monitored relative to intelligibility gains.

Additionally, another set of monosyllabic words adapted from a study by Kent, Weismer, Kent, and Rosenbek (1989) was recorded by the TE speakers that participated in the present study. This stimulus word list is of the forced-choice nature and has four responses to choose from for each stimulus word. It would be of interest to determine the intelligibility of the speakers using this task and discern if higher scores were found as a result of the forced-choice paradigm, which has been previously found in the literature (Vigouroux & Miller, n.d.; Yorkston & Beukelman 1978,1980). Performance characteristics associated with this set of experimental stimuli will be the topic of future investigations.

Secondly, it may be valuable to consider the creation of a word intelligibility test specifically designed for the (tracheo)esophageal speaker population. No such

test currently exists and all previous investigations of TE intelligibility used either tests adapted from studies of other clinical populations of alaryngeal speakers or other disordered speakers, or tests that were constructed by the researchers of the study at hand. As a result of this, individuals must be cautious in attempting to compare results across studies that use different methods to measure intelligibility. A universal test/word list would make results more generalizable for alaryngeal speakers. As well, given the unique nature of TE speech with its access to a pulmonary driving source with subsequent influences on the PE segment, a word list specific to TE speech would ensure the specific phonemes or manner of classification are accurately portrayed and sensitive to problems that may exist within the TE speaker population.

An area that was not addressed herein that could have contributed to the intelligibility of individuals was any associated postoperative complications each speaker may have experienced. More intensive surgeries, such as microvascular reconstructions or total laryngectomies combined with neck dissection, flap reconstruction, etc. have the potential to further alter the anatomy and physiology of the PE segment and, therefore, alter the intelligibility of each individual's speech. Chart reviews of the speaker participants from the present study, to determine their surgical history, or replication of the current study with participants with known reconstructive or advanced surgeries would provide an excellent group from which to compare to. This type of information would provide valuable insight to the research team as to additional challenges this subset of individuals may face producing intelligible speech.

Finally, multiple studies that have been conducted indicate TE speech to be judged as more acceptable than esophageal or electrolaryngeal speech (Clark & Stemple, 1982; Pindzola & Cain, 1988; Tardy-Mitzell et al., 1985; Williams & Watson, 1985). It would be valuable to determine if more acceptable speech was also perceived as more intelligible, introducing a quality judgment into the area of TE speech intelligibility testing (Pindzola & Cain, 1988). This could be accomplished by adding an acceptability rating portion to an intelligibility testing experiment, potentially through the use of a visual-analog score. Work of this type is currently underway with both a large population of TE speakers and listeners (Skidmore, Elliott, Sleeth, Bornbaum, & Doyle, 2011).

Summary and Conclusion

This research project was designed to investigate the intelligibility of individuals using TE speech as their primary mode of communication. Fifteen TE speakers recorded 66 monosyllabic words, which were then transcribed into English orthographics by 18 naïve, adult listeners. Gross intelligibility initially was determined by the percentage of correctly transcribed responses. Errors were also entered into confusion matrices in order to evaluate patterns of increased or decreased intelligibility.

Analysis of the data indicated that overall intelligibility of the group of speakers was 71%, and that errors predominantly occurred on the first phoneme of the stimulus word (word-initial). Additionally, a hierarchy of intelligibility was found among manner of production classes with stops and fricatives being the least

intelligible, and liquids and glides being the most intelligible. Many of the errors involving stops/fricatives were as a result of confusions surrounding the voiced-voiceless distinction. Patterns were found surrounding the amount of time postlaryngectomy, indicating that, for the most part, a longer time postlaryngectomy led to more intelligible speech. Exceptions were encountered though, and no other patterns could be determined from the remaining demographic data collected. This observation leads us to believe that individual anatomical and physiological differences among speakers have the greatest potential to influence speech intelligibility.

In addition, intelligibility values obtained from the present investigation, as well as hierarchy of intelligibility, based upon manner classification, follow trends found in literature published when TE speech was first introduced. This indicates that work in the area of TE speech intelligibility may still be fruitful in hopes of identifying clinical treatment protocols. It has been shown that highly intelligible speech not only leads to better communication for individuals using TE speech, but also the potential to increase an individual's QoL. Consequently, there is much to gain from continued research into the area of TE speech intelligibility. The time has come for attention to be brought to this issue, and for intelligibility to become a focus in postlaryngectomy communication. By doing so, laryngectomized individuals may be able to communicate in the most intelligible and effective manner as possible with the goal of achieving the greatest possible level of postlaryngectomy communication rehabilitation.

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.

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 Blom, E.D., Singer, M.I., & Hamaker, R.C. (Eds.), *Tracheoesophageal Voice Restoration Following Total Laryngectomy* (p. 1-8). San Diego, California: 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 esopheageal speech. *Journal of Speech and Hearing Research, 24,* 308-312.

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.

Cole, R.A., & Cooper, W.E. (1975). Perception of voicing in English affricates and fricatives. *Journal of the Acoustical Society of America, 58,* 1280-1287.

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.

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: Singular.

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* (p. 3-15). Texas: PRO-ED, Incorporated.

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* (p. 521-543). Texas: PRO-ED, Incorporated.

Doyle, P.C., & Haaf, R.G. (1989). Perception of pre-vocalic 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.

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.

Farrell, S., Dietrich-Burns, K., & Messing, B.P. (2004). Foundation skis for the artificial larynx part II: intra-oral artificial larynx. In *Voice Points* (part 2). Retrieved from: http://www.webwhispers.org/news/nov2004.htm.

Fauci, A.S., Braunwald, E., Kasper, D.L., Hauser, S.L., Longo, D.L., Jameson, J.L., … Loscalzo, J. (2008). Head and Neck Cancer. In E.E. Vokes (Ed.), Harrison's Principles of Internal Medicine, 17th Edition (Chapter 28). Retrieved from: http://www.accessmedicine.com/content.aspx?aID=2866817.

Ferrand, C.T. (2007). *Speech Science: an integrated approach to theory and clinical practice* (2nd ed.). Boston, MA: Pearson Education, Inc.

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.

Freije, J.E., Beatty, T.W., Campbell, B.H., Woodson, B.T., Schultz, C.J., & Toohill, R.J. (1996). Carcinoma of the larynx in patients with gastroesophageal reflux. *American Journal of Otolaryngology, 17*(6), 386-390.

Gillison, M. L., & Lowry, D.R. (2004). A causal role for human papillomavirus in head and neck cancer. *The Lancet, 363,* 1488-1489.

Gress, C.D., & Singer, M.I. (2005). Tracheoesophageal voice restoration. In P.C. Doyle & R.L. Keith (Eds.), *Contemporary Considerations in the Treatment and Rehabilitation of Head and Neck Cancer: Voice, Speech and Swallowing* (p. 431-452). Texas: PRO-ED, Incorporated.

Halle, M., Hughes, G., & Radley, J. (1957). Acoustical properties of stop consonants. *Journal of the Acoustical Society of America, 29,* 107-116.

Hasibe, M., Brennan, P., Benhamou, S., Castellsague, X., Chen, C., Curado, M.P., … Boffetta, P. (2007). Alcohol drinking in never users of tobacco, cigarette smoking in never drinkers, and the risk of head and neck cancer: Pooled

analysis in the International Head and Neck Cancer Epidemiology Consortium. *Journal of the National Cancer Institute, 99*(10), 777-789.

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* (p. 75-90). Texas: 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.

Jongmans, P., van Rossum, M., van As-Brooks, C., Hilger, F., & Pols, L. (2008). An evidence-based rehabilitation program for tracheoesophageal speakers. Unpublished paper presented at the Invitational Round Table "Evidence-based Voice and Speech Rehabilitation in Head and Neck Cancer", Amsterdam, Netherlands.

Kacker, A., Wolden, S., Pfister, D.G., & Kraus, D.H. (2003). Cancer of the larynx. In Myers, E.N., Suen, J.Y., Myers, J.N., & Hanna E.Y.N. (Eds), *Cancer of the Head and Neck: Fourth Edition* (p. 333-377). USA: Elsevier Science.

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.

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.

Keith, R.L., Shanks, J.C., & Doyle, P.C. (2005). Historical highlights: laryngectomy rehabilitation. In P.C. Doyle & R.L. Keith (Eds.), *Contemporary Considerations in the Treatment and Rehabilitation of Head and Neck Cancer: Voice, Speech and Swallowing* (p. 17-57). Texas: PRO-ED, Incorporated.

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.

Lisker, L., & Abramson, A.S. (1967). Some effects of context on voice onset time in English stops. *Language and Speech, 10,* 1-28.

Marur, S., D'Souza, G., Westra, W.H., Forastiere, A.A. (2010). HPV-associated head and neck cancer: a virus-related cancer epidemic. *The Lancet Oncology, 11,* 781-789.

Matthews, T.W., & Lampe, H.B. (2005). Treatment options in oral cancer. In P.C. Doyle & R.L. Keith (Eds.), *Contemporary Considerations in the Treatment and Rehabilitation of Head and Neck Cancer: Voice, Speech and Swallowing* (p. 153-170). Texas: PRO-ED, Incorporated.

Max, L., De Bruyn, W., & Steurs, W. (1997). Intelligibility of oesophageal and tracheo-oesophageal speech: preliminary observations. *European Journal of Disorders of Communication, 32,* 429-440.

McKaig, R.G., Baric, R.S., & Olshan, A.F. (1998). Human papillomavirus and head and neck cancer: epidemiology and molecular biology. *Head and Neck, 20*(3), 250-265.

Mendenhall, W., Morris, C., Stringer, S., Amdur, R., Hinerman, R., Villaret, D., … Robbins, K.T. (2002). Voice rehabilitation after total laryngectomy and postoperative radiation therapy. *Journal of Clinical Oncology, 20*(10), 2500-2505.

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.

Pai, S., & Westra, W.H. (2009). Molecular pathology of head and neck cancer: Implications for diagnosis, prognosis and treatment. *Annual Review of Pathology: Mechanisms of Disease, 4,* 49-70.

Paz, I.B., Cook, N., Odom-Maryon, T., Xie, Y., & Wilczynski, S.P. (2000). Human papillomavirus (HPV) in head and neck cancer, *Cancer, 79*(3), 595-604.

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

Rajashekar, B., Harrington, K.J., & Kazi, R. (2009). Evolution of voice restoration and the tracheo-oesophageal voice prostheses. In Nayak, U., Kazi, R. (Eds.), *Voice Restoration After Total Laryngectomy: Current science and future perspectives* (p. 1-12) India: Sanat Printers.

Raphael, L., Dorman, M.F., & Liberman, A.M. (1980). On defining the vowel duration that cues voicing in final position. *Language and Speech, 23,* 297-307.

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.

Rogers, S.N., Laher, S.H., Overend, L., & Lowe, D. (2002). Importance-rating using the University of Washington quality of life questionnaire in patients treated by primary surgery for oral and oro-pharyngeal cancer. *Journal of Craniomaxillofacial Surgery, 20,* 125-132.

Schiavetti, N. (1992). Scaling procedures for the measurement of speech intelligibility. In Kent, R.D. (Ed.) *Intelligibility in Speech Disorders,* (p. 11-34). 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.

Skidmore, E., Elliott, H.C., Sleeth, L., Bornbaum, C.C., Doyle, P.C. Relationships between listener comfort and voice severity in tracheoesophageal speech. Unpublished paper presented at the Annual American Speech-Language Hearing Association convention, San Diego, CA, November 2011.

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.

Slis, I.H., & Cohen, A. (1969). On the complex regulating the voiced-voiceless distinction. *Language and Speech, 12,* 137-155.

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

Smoking in Canada. (2011). Retrieved February 27, 2012, from http://www.smoke-free.ca/factsheets/pdf/prevalence.pdf.

Syrjanen, S. (2005). Human papillomavirus (HPV) in head and neck cancer. *Journal of Clinical Virology, 32S,* S59-S66.

Talamini, R., Bosetti, C., La Vecchia, L., Dal Maso, L., Levi, F., Bidoli, E., … Franceschi, S. (2002). Combined effect of tobacco and alcohol on laryngeal cancer risk: a case-control study. *Cancer Causes & Control, 13*(10)*,* 957-964.

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.

Treatment for Laryngeal Cancer. (2011). Retrieved November 22, 2011, from http://http://www.cancer.ca/Canada-wide/About cancer/Types of cancer/Treatmentforlaryngeal cancer.aspx?sc_lang=en.

Trudeau, M.D. (1994). The acoustical variability of tracheoesophageal speech. In R.L. Keith & F.L. Darley (Eds.), *Laryngectomy Rehabilitation* (p. 383-394). Texas: Pro-Ed, Incorporated.

Tyan, Y-S., Liu, S-T., Ong, W-R., Chen, M-L., Shu, C-H., & Chang, Y-S. (1993). Detection of Epstein-Barr virus and human papillomavirus in head and neck tumors. *Journal of Clinical Microbiology, 31*(1), 53-56.

van As, C.J., Higers, 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.

Vigouroux, J., & Miller, N. (n.d.). Intelligibility testing: Issues in closed versus open format scoring. Retrieved February 8, 2012, from:http://www.ncl.ac.uk/linguistics/assets/documents/JENNIFER.VIGOUROUX_NICK.M-FT.pdf.

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

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

World Health Organization. (1947. *Health Promotion Glossary.* Retrieved February 24, 2012 from http://www.who.int/healthpromotion/about/HPG/en/

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.

APPENDIX A

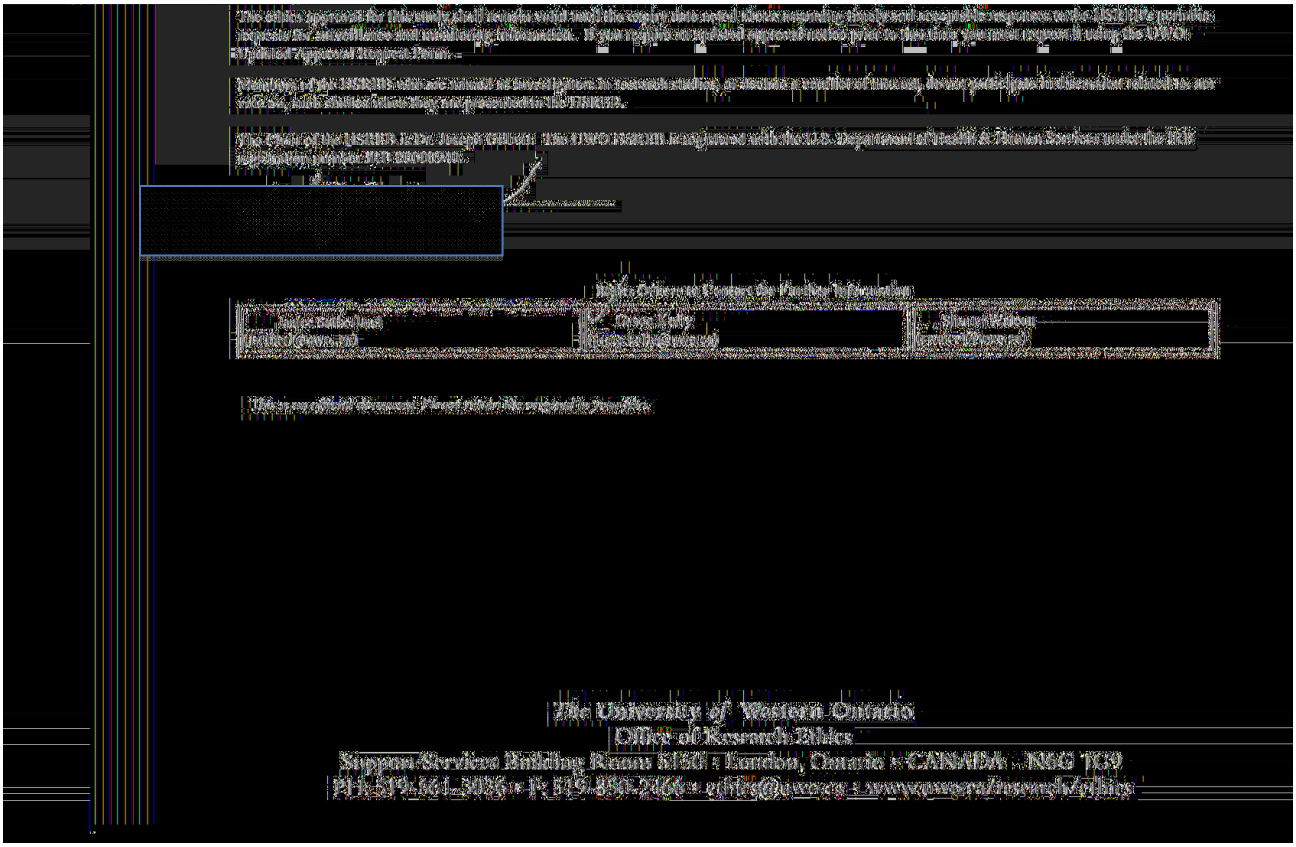
Research Western

Use of Human Participants - Ethics Approval Notice

Principal Investigator: Dr. Philip Doyle
Review Number: 18588E
Review Level: Delegated
Approved Local Adult Participants: 40
Approved Local Minor Participants: 0
Protocol Title: Exploring Speech Intelligibility of Individuals Who Use Tracheoesophageal Speech
Department & Institution: Communication Sciences & Disorders, University of Western Ontario
Sponsor:
Ethics Approval Date: November 24, 2011 **Expiry Date:** December 31, 2012
Documents Reviewed & Approved & Documents Received for Information:

Document Name	Comments	Version Date
UWO Protocol		
Letter of Information & Consent	2011 - 1	
Other	Recruitment Script	

This is to notify you that The University of Western Ontario Research Ethics Board for Health Sciences Research Involving Human Subjects (HSREB) which is organized and operates according to the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans and the Health Canada/ICH Good Clinical Practice Practices: Consolidated Guidelines; and the applicable laws and regulations of Ontario has reviewed and granted approval to the above referenced revision(s) or amendment(s) on the approval date noted above. The membership of this REB also complies with the membership requirements for REB's as defined in Division 5 of the Food and Drug Regulations.



APPENDIX B

1. LEAVE
2. CANE
3. JOG
4. CHEAP
5. CATCH
6. MEAL
7. THY
8. TAB
9. FIVE
10. MASS
11. VEAL
12. RICE
13. PAD
14. WEDGE
15. TEETHE
16. HALF
17. CAME
18. DOPE
19. SACK
20. ICE
21. PAT
22. MASH
23. FEEL
24. WITCH
25. NEAR
26. DAB
27. SAG
28. HUN
29. BAD
30. ZACK
31. EASE
32. RICH
33. TEETH
34. BAT
35. DEER
36. HUNG
37. LEAF
38. KEEP
39. SHAVE
40. ZAG
41. SEEK
42. VEER
43. THING
44. RISE
45. BADGE
46. SHEATH
47. GAB
48. GAIN
49. THIGH
50. PATH
51. GAME
52. EDGE
53. CHAD
54. VET
55. SHEATHE
56. CHIEF
57. THESE
58. FISH
59. ZING
60. JAW
61. THEME
62. GNASH
63. THOU
64. KNOW
65. LOATHE
66. WAY

APPENDIX C

Letter of Information and Consent Form

Study Title: *Voice Production and Perception Laboratory Voice Sample and Voice-Related Quality of Life Database*

Principal Investigator: Dr. Philip C. Doyle, Ph.D.

Co-Investigators: Adam M.B. Day, M.Sc.
, M.Sc.
, M.Sc.
Marie-Ève Caty, M.P.O., S-LP Reg. CASLPO
Lindsay Sleeth, B.H.Sc.
Dr. Kevin Fung, M.D., FRCS(C)

In the sections to follow, the pronouns "you" and "your" should be read as referring to the participant rather than the parent/guardian/next of kin who is signing the consent form for the participant.

Introduction

This letter contains information to help you or your child decide whether or not to participate in this research study. It is important for you to understand why the study is being conducted and what it involves. Please read this letter carefully and feel free to ask questions if anything is unclear or there is something you do not understand.

You or your child are being invited to take part in this study because you have a voice disorder or use a method of alaryngeal speech.

Purpose of Study

The purpose of this study is to collect voice samples and voice-related quality of life data from individuals with voice disorders and individuals who use an alaryngeal methods of voice production. We are interested in building a database to store this information that will allow us to test how listeners perceive your voice and to test how your voice compares to other voices. Additionally, your data will be used to explore how one's voice-related quality of life is impacted as a result of a voice-disorder or use of an alaryngeal method of speech.

Inclusion Criteria
If you are over the age of 5 years old and can read, write, and speak English, you can choose to participate in this study.
Exclusion Criteria
If you are unable to read, write, and speak English, you should not participate in this study.

Description of the Research
This study will require you speak into a microphone so your voice can be recorded. This will involve the recording of several sustained vowels such as "ah", "ee", and "ooh", repeating some short sentences, and the reading aloud of a short paragraph that is age appropriate. The recording will require 10 minutes and will be done in a formal recording suite or quite room within a private setting. As well, you will be asked to complete two written questionnaires, the Voice-Related Quality of Life Questionnaire and the Voice Handicap Index.

Participation in this study will require keeping your voice samples and questionnaire data in a secure database indefinitely for the purposes of future research. If you do not wish for your voice samples and survey data to be used for future research, please do not participate in this study.

Risks & Harms
There are no known or anticipated physical, psychological, or emotional risks or discomforts associated with completing this study. However, if you do experience any problems or discomfort, you can discontinue the task.

Benefits
You may not directly benefit from participating in this study but information gathered may provide benefits to society as a whole.

Voluntary Participation
Participation in this study is voluntary. You may refuse to participate, refuse to answer any questions, refuse to complete a voice task, or withdraw from the study at any time, even in the future, with no effect on your current or future health care. You will not be compensated for your participation in this research.

Refusal to Participate & Discontinuing Participation
The decision to participate is yours to make. If at any time you wish to discontinue your participation you may do so without penalty and all of your information will be destroyed. If at any time you wish to discontinue or withdraw your participation, please contact Dr. Philip Doyle.

In the case that your voice samples and data are being used in an active research project, withdrawal of data will not be permitted until the completion of that research project.

Confidentiality

Your identity and personal information will be coded and known and accessible only by the investigators of this study. Your contact information is being collected so that we can contact you to invite you to participate in future research and to contact you if we experience any threats to your privacy. In addition, representatives of The University of Western Ontario Health Sciences Research Ethics Board may contact you or require access to your study-related records to monitor the conduct of the research.

All of your personal data will be stored electronically in a password protected and encrypted file and as a hard copy in a locked filing cabinet at a locked laboratory at the University of Western Ontario. This locked file is only accessible to the study investigators. Also, a unique identifier will be used instead of your name on all study materials and instruments to protect your confidentiality. If the results of the study are published, your name will not be used and information that discloses your identity will not be released or published.

Each participant's full name will be collected and retained to allow our lab to track each individual over multiple collection points and to allow us to contact them to invite them to participate in further research. Further, because opportunities to collect additional voice and VRQOL data often occur over time (e.g., follow up appointments with head and neck surgeon, attendance at national meetings/conferences, etc.), it is important that we are able to reference individuals by name in the database so that additional data can be attributed to the same individual, and not entered as new participant. When appropriate phone numbers will be collected and retained to allow participant contact for scheduling multiple visits and to allow contact for invitations to participate in future research. Participants' date of birth will be collected to allow comparisons across age, particularly when multiple data collection events occur over years.

For recordings and survey information that may be transferred digitally across an international border, Border Security can ask to see digital information contained on the laptop recording system (encrypted or otherwise). While your information will be coded and known only to the investigators, this potential privacy risk must be brought to your attention.

In the future, your data might be shared with other researchers according to a data sharing agreement. However, if such data sharing is undertaken, those who will have access to this information must complete a separate ethics

submission and data sharing agreement. In this case, your data will not contain any identifiable information.

Waiver of Rights
You do not waive any legal rights by signing the consent form.

This letter and the consent statement are yours to keep.

Page 6 of this document is the investigators' copy of your consent statement.

Consent Statement – **Participant's Copy**

I have read the attached Letter of Information, have had the nature of the study explained to me and agree to participate. All questions have been answered to my satisfaction.

Participant's Name (please print): _____

Participant's Signature: _____

Date: _____

Person Obtaining Informed Consent (please print): _____

Signature: _____

Date: _____

Consent Statement – **Investigators' Copy**

Project Title: *Voice Production and Perception Laboratory Voice Sample and Voice-Related Quality of Life Database*

Study Investigators:

Dr. Philip C. Doyle, Ph.D.
Adam M.B. Day, M.Sc.
Agnieszka Dzioba, M.Sc.
Catherine Bornbaum, M.Sc.
Marie-Ève Caty, M.P.O., S-LP Reg. CASLPO
Lindsay Sleeth, B.H.Sc.
Dr. Kevin Fung, MD, FRCS(C)

I have read the Letter of Information, have had the nature of the study explained to me and agree to participate. All questions have been answered to my satisfaction.

Participant's Name (please print): _____

Participant's Signature: _____

Date: _____

Person Obtaining Informed Consent (please print): _____

Signature: _____

Date: _____

APPENDIX D

Voice Production and Perception Laboratory
Health and Rehabilitation Sciences
University of Western Ontario

Project Title: "Exploring speech intelligibility of individuals who use tracheoesophageal speech"

Principle Investigators:
Philip C. Doyle, Ph.D. *Rehabilitation Sciences and Department of Otolaryngology, UWO*
Lindsay Sleeth, BHSc. *Health and Rehabilitation Science, UWO*
Kevin Fung, MD. *Otolaryngology*

Letter of Information

1. Purpose of the Study

The purpose of this letter is to provide you with information required for you to make an informed decision regarding participation in this research. You are being invited to participate in research that will assess the intelligibility of individuals who use a method of verbal communication termed "tracheoesophageal" speech.

2. Inclusion Criteria

Individuals aged 18 years or older that are of normal hearing and can read and write English can participate in the study.

3. Exclusion Criteria

Individuals with hearing impairments or that are unable to read or write English will not be able to participate in the study. As well, individuals should not previously be familiar with tracheoesophageal speech.

4. Activities of Participants

If you agree to participate, you will be asked to listen to a list of monosyllabic words spoken by tracheoesophageal speakers. After listening to each word, you will be asked to write down the word you heard. It is anticipated that the entire task will be completed in less than one hour, during a single session. The task will be conducted in the Voice Production and Perception Laboratory at Elborn College, Room 2200.

5. Possible Risks and Harms

There are no known or anticipated physical, psychological, or emotional risks or discomforts associated with completing this study. However, if you do experience any problems or discomfort, you may discontinue the task at any time without penalty.

6. **Possible Benefits**

You may not directly benefit from participating in this study but information gathered may provide benefits to society as a whole. You will not be compensated for your participation in this research.

7. **Voluntary Participation**

Participation in this study is voluntary. You may refuse to participate, refuse to answer any questions or withdraw from the study at any time with no effect on your future (care/academic status/employment etc).

8. **Confidentiality**

All data collected will remain confidential and accessible only by the investigators of this study. If the results are published, your name will not be used. If you choose to withdraw from this study, your data will be removed and destroyed from our database.

This letter is yours to keep for future reference.

REB Approval # 18588E

Letter of Consent

Project Title: "Exploring speech intelligibility of individuals who use tracheoesophageal speech"

I have read the "Letter of Information" and have had the nature of the study explained to me and I agree to participate. All questions have been answered to my satisfaction.

Participant's Name (please print): _____

Participant's Signature: _____

Date: _____

Person Obtaining Informed Consent (please print): _____

Signature: _____

Date: _____

APPENDIX E

Voice Production and Perception Laboratory
Demographic Information Questionnaire

CODE

|____|____|____|____|____|____|

Questions about your treatment:

Neck dissection: Y | N If yes, which side: Left | Right | Both

Radiation: Y | N If yes, pre or post surgery: Pre | Post | Both

Chemotherapy: Y | N If yes, pre or post surgery: Pre | Post | Both

Questions about your voice:

Primary Speech Mode: Tracheoesophageal (TE) | Esophageal (ES) | Electrolarynx (EL)

If TE, primary (at time of surgery) or secondary (after surgery): Primary | Secondary

If TE, which prosthesis: Blom-Singer - InHealth | Atos - Provox | Other

If "other", please specify: _____

If TE, size_____ indwelling device: Y | N

For communication purposes, overall, I would rate my voice as:

Very poor | Poor | Fair | Good | Excellent

Specific to my expectations, the method of postlaryngectomy communication that I use:

___ Falls extremely short of my expectations
___ Falls somewhat short of my expectations
___ Meets my expectations
___ Somewhat exceeds my expectations
___ Substantially exceeds my expectations

Other treatment or health related notes: _____

_____

_____

CURRICULUM VITAE

LINDSAY E. SLEETH

Education

M.Sc. Health & Rehabilitation Sciences, Rehabilitation Sciences
The University of Western Ontario, London, Ontario, Canada.
Degree to be conferred October 2012.

B.H.Sc. Health Sciences, Specialization in Health Sciences
The University of Western Ontario, London, Ontario, Canada.
Degree conferred June 2010.

Honours & Awards

2011-2012	Queen Elizabeth II Graduate Scholarship in Science and Technology -Obtained award in the value of $15,000
2011-2012	Western Graduate Research Scholarship -Obtained award in the value of $10,000
2010-2011	Western Graduate Research Scholarship -Obtained award in the value of $10,000
2008-2010	Dean's Honour List
2009	Bruno Da Silva Community Service Award -Awarded to an individual in recognition of their dedication to education and involvement in the community.

Publications & Presentations

Published Abstracts and Conference Presentations:

Sleeth, L., Skidmore, E., Elliot, H.C., Bornbaum, C.C., & Doyle, P.C. Auditory-perceptual assessments of voice severity and listener comfort in postlaryngectomy tracheoesophageal speakers. Poster presented at the 4th annual Aging, Rehabilitation & Geriatric Care Research conference, London, ON, February 2011.

Sleeth, L., Skidmore, E., Elliot, H.C., Bornbaum, C.C., & Doyle, P.C. Using Voice Severity and Listener Comfort as a Psychophysical Evaluation of Postlaryngectomy Tracheoesophageal Speech. Poster presented at the 23rd annual Western Research Forum, London, Ontario, February 2011.

Sleeth, L., Skidmore, E., Elliot, H.C., Bornbaum, C.C., & Doyle, P.C. Auditory-perceptual assessments of voice severity and listener comfort in postlaryngectomy tracheoesophageal speakers. Poster presented at the Faculty of Health Sciences Research Day, London, ON, March 2011.

Bornbaum, C.C., Doyle, P.C., Day, A.M.B., Dzioba, A. & **Sleeth, L.** Construct validation of the V-RQOL in alaryngeal speakers. Paper presented at the Annual American Speech-Language Hearing Association convention, San Diego, CA, November 2011.

Sleeth, **L**., Skidmore, E., Elliott, H.C., Bornbaum, C.C., & Doyle, P.C. Relationships Between Listener Comfort and Voice Severity in Tracheoesophageal Speech. Poster presented at the Annual American Speech-Language Hearing Association convention, San Diego, CA, November 2011.

Sleeth, L., Day, A.M.B., & Doyle, P.C. Exploring speech intelligibility in individuals using tracheoesophageal speech. Poster presented at the 5th annual Health and Rehabilitation Sciences Graduate Research Forum, London, ON, February 2012.

Professional Activities

Employment:

09/2010 – 01/2011	Teaching Assistant. HS2300 – Functional Anatomy. The University of Western Ontario, London, Ontario, Canada.

Memberships:

2010 – Present	Rehabilitation Sciences Journal Club Graduate Student Member
2009 – 2010	Faculty of Health Sciences Faculty Council Executive Member
2009 – 2010	Faculty of Health Sciences "Dream Team" Charitable Initiative Director

Volunteer Activities:

2010	The Faculty of Health Sciences Alumnus Ambassador
2007 – present	The Make-A-Wish Foundation of Southwestern Ontario Volunteer
2009 – 2010	Parkwood Hospital, London, ON Speech Pathology Clinic Volunteer