February 2019

An Experimental Evaluation of Stop-Plosive and Fricative Consonant Intelligibility by Tracheoesophageal Speakers in Quiet and Noise

Sebastiano Failla
The University of Western Ontario

Supervisor
Doyle, Philip C.
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

Abstract

Despite functional levels of postlaryngectomy communication, individuals who undergo total laryngectomy and tracheoesophageal (TE) puncture voice restoration continue to experience significant communication difficulties in noisy environments. In an effort to identify and further characterize TE speakers' intelligibility in noise, the current auditory-perceptual study investigated stop-plosive and fricative intelligibility of TE speech in quiet and in the presence of multi-talker noise. Eighteen listeners evaluated monosyllabic consonant-vowel-consonant words produced by 14 TE speakers using an open-response paradigm. Our findings indicate that overall intelligibility was significantly lower in noise. Further examination showed a differential effect of noise on intelligibility according to manner and phoneme position. While overall error patterns remained consistent across conditions, voicing distinction was affected differentially according to manner and position. The present investigation provides valuable insight into the difficulties faced by TE speakers in noisy speaking environments, as well as a basis for optimization of counseling and postlaryngectomy voice rehabilitation.

Keywords: tracheoesophageal, alaryngeal, intelligibility, stop-plosive, fricative, noise

Acknowledgments

First and foremost, I wish to express my sincerest gratitude to Dr. Philip Doyle, without whom this project would not have been possible. You have been an exceptional mentor and guide throughout my graduate experience, as well as other aspects of my life. I am truly grateful that I had the opportunity to meet and work with you. Thank you for everything.

To the volunteers at the Voice Production and Perception Laboratory, thank you for assisting with the copious amounts of data this project yielded.

To my significant other, thank you for being my rock in times of need. Going through the same experience, you knew exactly what I was going through and supported me every step of the way.

Lastly, to my parents, thank you for your unwavering support and encouragement through both the good and bad times. I can say without a doubt that there is nobody else in this world that could have put up with as much as you did for me. I would not have been able to achieve what I have without you. I truly cannot express how grateful I am for the support you have provided and continue to provide.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

Review of Literature

Laryngeal cancer is a disease which has the potential to impact all areas of quality of life including physical, psychological, and social well-being (Eadie & Doyle, 2004, 2005; Meyer et al., 2004). In cases of significant disease progression, treatment may involve the removal of the entire larynx (i.e., laryngectomy) and subsequent loss of typical verbal communication. Fortunately, for this situation, several postlaryngectomy “alaryngeal” voice rehabilitation methods have been developed in an effort to restore verbal communication. Despite functional levels of postlaryngectomy communication, alaryngeal speakers continue to experience lower intelligibility and difficulty communicating in the presence of competing noise (Clark, 1985; Clark & Stemple, 1982; Dudley, 1984; McColl, Fucci, Petrosino, Martin, & McCaffrey, 1998). Given the significant role of speech communication in influencing alaryngeal speakers’ quality of life (Eadie & Doyle, 2004, 2005; Meyer et al., 2004), it is imperative that investigations that seek to identify specific areas of communication difficulty are pursued (Terrell et al., 2004).

Accordingly, the chapter to follow will provide a brief introduction to laryngeal cancer including its epidemiology and etiological factors, as well as a review of treatment and rehabilitative options. Furthermore, literature pertaining to postlaryngectomy communication outcomes will be reviewed. As part of this review, there will be a particular focus on speech intelligibility associated with one specific method of alaryngeal communication, namely, tracheoesophageal (TE) speech.

Epidemiology

It is estimated that 1,150 new cases of laryngeal cancer were diagnosed in Canada in 2017, with approximately 440 deaths attributable to the disease (Canadian Cancer Sociey, 2017). Approximately 50% of newly diagnosed individuals present with advanced (Stage III or IV) disease (Smith et al., 2018). Men are more commonly diagnosed with laryngeal cancer, accounting for 80% to 85% of the patient population (Canadian Cancer Sociey, 2017; MacNeil et al., 2015). Fortunately, incidence rates have decreased significantly over the last 30 years (Canadian Cancer Sociey, 2017). However, despite advances in technology and clinical management, five-year survival rates have remained relatively unchanged, hovering around 57% since the mid 1990's (MacNeil et al., 2015). The relative stability of survival rates suggests that the potential for loss of one's larynx due to malignancy remains a clinical and postlaryngectomy rehabilitation concern.

Etiological Factors

A number of risk factors have been implicated in the etiology of laryngeal cancer; with alcohol consumption and tobacco smoking comprising the most significant factors associated with the development of laryngeal cancer.

Both ethanol and its metabolite – acetaldehyde – have been classified as group 1 carcinogens by the International Agency on Research for Cancer (IARC) (IARC, 2012). A meta-analysis conducted by Islami et al. (2010) reported that overall, alcohol consumption was associated with a 2-fold increased risk of developing laryngeal cancer compared to its occasional use or non-consumption. Their results indicate that light consumption (≤1 drink/day) was not associated with an increased risk. However, moderate (>1 to <4 drinks/day) and heavy (≥4 drinks/day) alcohol consumption were associated with a 1.5 and 2.5-fold increased risk, respectively (Islami et al., 2010). Beyond its direct carcinogenic effects, alcohol also has the potential to prolong the exposure of other carcinogens within the body due to its nature as a chemical solvent. That is, the presence of alcohol may delay the rate of metabolism and excretion of other carcinogens. In particular, a number of studies have demonstrated that alcohol and tobacco have a synergistic effect on the risk of developing laryngeal cancer, with studies reporting odds ratios as high as 177 (Hashibe et al., 2009; Talamini et al., 2002).

Beyond tobacco's synergistic role, it has also been classified as a group 1 carcinogenic substance itself (IARC, 2012). A meta-analysis of 30 studies including 14,293 cases of laryngeal cancer concluded that duration and frequency of tobacco smoking were significantly correlated with an increased risk of developing laryngeal cancer (Zuo et al., 2017). More specifically, individuals who smoked more than 30 cigarettes per day were 7 times more likely to develop cancer of the larynx. Moreover, relative risk continually rose within the first 40 years of smoking, with those who smoked more than 40 years achieving a 5-fold increased relative risk (Muscat & Wynder, 1992; Wynder, Mushinski, & Spivak, 1977; Wynder & Stellman, 1977; Zuo et al., 2017).

The association and clinical significance of human papillomavirus (HPV) with head and neck cancers, most notably oropharyngeal cancers, has gained considerable attention over the last several decades. A recent study investigating the prevalence of HPV among 3680 cases of head and neck cancers worldwide found that only 3.5% of laryngeal squamous cell carcinomas were positive for the presence of biologically active HPV (Castellsagué et al., 2016). In other words, only 3.5% of cases could have been directly the result of HPV. In contrast, however, it has been suggested that 22.4% of oropharyngeal cancers had biologically active HPV. While it is clear that HPV has a causative role in head and neck cancers (Castellsagué et al., 2016; Gillison et al., 2000; Kreimer, Clifford, Boyle, & Franceschi, 2005; Wei et al., 2012; zur Hausen, 2000), it is important to note its contribution is highly site-specific (Castellsagué et al., 2016), with the oropharynx being the most common site of malignancy.

A number of other risk factors for laryngeal cancer have also been proposed in the literature; these include, but are not limited to, gastroesophageal reflux (Zhang, Zhou, Chen, Zhou, & Tao, 2014), Epstein-Barr virus (de Oliveira et al., 2006; Rota, Fidan, Muderris, Yesilyurt, & Lale, 2010), and several genetic polymorphisms (Li & Liu, 2014; Qi & Zhou, 2014; Starska et al., 2014). Exposure to asbestos has also been proposed as a risk factor, however, a recent systematic review conducted by Ferster, Schubart, Kim, and Goldenberg (2017) illustrates a lack of evidence to support this notion. Based on their assessments, Ferster et al. (2017) state that very few studies found an association between asbestos exposure and laryngeal cancer, and those that did, failed to account for tobacco and alcohol as confounding factors. While the aforementioned factors (including HPV) are not insignificant, their etiological role is less well documented in comparison to that of alcohol and tobacco use, which have historically been considered the largest etiological factors of laryngeal cancer when consumed in excess or in combination (Wynder et al., 1977).

Treatment

Currently, several treatment modalities are available for the management of malignant laryngeal disease. Treatment is dependent upon a variety of factors including those specific to the tumor, as well as the patient (Angel, Doyle, & Fung, 2011). Early stage tumors (Stage I and II) classified as T1 or T2 are primarily treated with either primary endoscopic excision or radiotherapy, both of which have been shown to have comparable oncological outcomes (Cohen, Garrett, Dupont, Ossoff, & Courey, 2006; Higgins, Shah, Ogaick, & Enepekides, 2009; Hristov & Bajaj, 2008; Kazi et al., 2008; Rosier et al., 1998). More advanced disease (Stage III and IV) with tumors that have progressed to T3 or higher require more aggressive approaches. These include tumours limited to the larynx with vocal cord fixation and/or invasion of the preepiglottic and paraglottic space, cricoid and thyroid cartilage, and postcricoid area (Stevenson, 2018). Laryngeal tumors with extracapsular spread and those with an increased likelihood of lymphatic invasion or distant metastasis also constitute advanced disease irrespective of T staging (Stevenson, 2018).

The current primary choice of treatment for advanced disease is combined chemoradiotherapy (CRT) (Angel et al., 2011). CRT has been shown to have similar survival outcomes compared to the historical standard practice of radical surgery (i.e., total laryngectomy with bilateral neck dissection) combined with postoperative radiotherapy (Forastiere et al., 2003; Jacobi, van der Molen, Huiskens, van Rossum, & Hilgers, 2010; Trivedi et al., 2008). Moreover, CRT has the added benefit of preserving anatomical and physiological function of the larynx – namely, those functions related to breathing, swallowing, and speech. However, CRT can have a variety of adverse effects including, but not limited to, radiodermatitis, xerostomia, and dysgeusia (Stenson et al., 2012; Xiao et al., 2013). In circumstances where CRT proves to be unsuccessful, total laryngectomy – the complete removal of the larynx – may be used as a salvage therapy (Rassekh & Haughey, 2010).

While surgery is not always the primary choice for treatment of advanced laryngeal cancer, laryngectomy continues to play a significant role in treatment. When clinically feasible, surgical organ preservation methods (i.e., partial laryngectomy) are generally preferred over complete laryngeal resection. Both transoral microsurgery and transoral robotic partial laryngectomy have been shown to provide excellent oncological and functional outcomes with relatively low associated morbidity (Cabanillas, Rodrigo, Llorente, & Suarez, 2008; Grant et al., 2007; Kayhan, Kaya, Altintas, & Sayin, 2014; Mendelsohn, Remacle, Van Der Vorst, Bachy, & Lawson, 2013; Ozer et al., 2013). Similar to CRT, total laryngectomy may be used as a salvage therapy when partial laryngectomy proves unsuccessful. In cases where laryngeal tumors have cartilaginous invasion and destruction, extensive involvement of the cricoid cartilage, and/or subglottic extension, total laryngectomy may be indicated as a primary treatment (Rassekh & Haughey, 2010).

The complete removal of the larynx causes significant disruptions in anatomical structure and function - namely, the separation of the alimentary canal from the upper respiratory tract. As a result, the trachea is brought forward surgically to form a tracheostoma at the base of the anterior neck from which patients will be required to breathe from on a permanent basis. In addition to changes in breathing and swallowing function, removal of the larynx results in the subsequent loss of the active abductor/adductor mechanism required for fine tuning of voice and speech production. Fortunately, several methods of postlaryngectomy voice rehabilitation exist. Information regarding these methods will be discussed below.

Postlaryngectomy Voice Rehabilitation

Currently, there are two types of voice rehabilitation methods available postlaryngectomy – extrinsic and intrinsic. Extrinsic methods rely on external vibratory sources for the production of sound, whereas intrinsic methods rely on internal vibratory sources originating from biological tissue (Doyle, 1994).

Electrolaryngeal (EL) speech is the sole external method of postlaryngectomy voice. For this type of alaryngeal voice, tone is generated externally through the use of a small hand-held electronic sound source known as an artificial electrolarynx. By placing the device on the neck or cheek, vibrations are transmitted through the skin and into the vocal tract where they can be articulated into speech (Hilgers & van den Brekel, 2010). Vibrations may also be introduced intraorally through the use of a plastic tube attachment when extensive scar tissue in the neck prevents adequate vibrational transmission (Hilgers & van den Brekel, 2010). When taught with excellent instruction, the result can provide an effective form of verbal communication that can often be acquired shortly after laryngectomy. While EL speech is typically easily learned and can be utilized by almost all individuals postlaryngectomy, its unnatural sound has been associated with reduced voice-related quality of life (QoL) outcomes compared to other methods of postlaryngectomy voice (Moukarbel et al., 2011). Nevertheless, EL speech continues to be a viable option as a primary method of verbal communication or as a secondary method should other postlaryngectomy voice methods fail.

Esophageal (ES) speech on the other hand, is considered an intrinsic method of alaryngeal speech due to its vibratory source originating from the pharyngoesophageal (PE) mucosa. Insufflation of the esophagus followed by expulsion of air vibrates the PE segment to create a sound source. As the sound source travels into the upper vocal tract and oral cavity, it may be articulated into speech (Hilgers & van den Brekel, 2010). While ES speech is an effective method of postlaryngectomy communication, it requires expert instruction, which can at times be difficult to obtain (Doyle & Finchem, in press). Moreover, ES is characterized by relatively low pitch because of the tissue mass that is used for voicing, as well as short phonation times due to the limited quantity of air available for PE mucosa vibration (Robbins, 1984; Robbins, Fisher, Blom, & Singer, 1984).

Singer and Blom's (1980) introduction of the tracheoesophageal (TE) method of alaryngeal voice – the second intrinsic method – was an elegant solution to the aerodynamic limitations experienced by ES speakers. Like ES speech, the PE segment acts as the internal vibratory source. However, in contrast to the “swallowed air” or insufflation technique of ES speech, users of TE speech have access to normal lung volumes for outward airflow generation used to vibrate the PE mucosa. In order for individuals to acquire the characteristic pulmonary aerodynamic drive unique to TE speakers, the posterior wall of the trachea and anterior wall of the esophagus must be connected via a surgically-formed midline fistula. A one-way valved prosthesis is then inserted into the puncture such that material (including air and fluids) can only flow from the trachea to the esophagus, preventing aspiration (Hilgers & van den Brekel, 2010). Upon digital tracheostomal occlusion and subsequent exhalation, air is diverted into the esophagus via the TE prosthesis, where it vibrates the PE mucosa to form the sound source. As vibrations travel up the vocal tract and into the oral cavity, they may be articulated to form speech (Hilgers & van den Brekel, 2010). In general, and although voice quality is not normal (Eadie & Doyle, 2004), TE voice is regarded as being the most natural sounding compared to other methods of alaryngeal voice. As a result, TE speech has become an extremely viable option for postlaryngectomy voice rehabilitation (Evans, Carding, & Drinnan, 2009; Hilgers & van den Brekel, 2010).

Assessing Tracheoesophageal Communication Outcomes

Several studies have sought to investigate postlaryngectomy communication outcomes using a metric termed speech intelligibility – the extent to which a speaker’s message is recovered by a listener (Hillman, Walsh, & Heaton, 2005; Kent, Weismer, Kent, & Rosenbek, 1989). Early studies primarily focused on listeners' judgements of speaker intelligibility through the use of scaling procedures including visual analogue and Likert scales because of their ease of application and scoring (Schiavetti, 1992). However, as intelligibility testing gained popularity in other disordered speech populations, scaling procedures received considerable criticism for their limitations in estimating intelligibility and identifying specific error patterns influencing overall intelligibility (Schiavetti, 1992). As a result, there has been a shift towards more objective word identification procedures where intelligibility can be expressed as the percentage of message recovery by listeners. While word identification procedures allow for more accurate descriptions of intelligibility and identification of specific areas of communication difficulty, they are not without limitations. More specifically, intelligibility may vary as a function of a number of factors including listener experience (i.e., experienced vs. naïve) and age (i.e., younger vs. older), stimulus type (i.e., word vs. sentence, nonsense stimuli vs. real words), type of noise present (i.e., quiet vs. multi-talker noise vs. white noise vs. amplitude-modulated white noise), and response paradigm (i.e., open vs. closed-choice) (Bridges, 1991; Clark, 1985; Danhauer, Doyle, & Lucks, 1985).

Several studies have shown that listeners with greater familiarity and/or formal training with alaryngeal speech (i.e., speech-language pathologists) report speaker intelligibility as being higher compared to naïve listeners (Beukelman & Yorkston, 1980; Bridges, 1991; Doyle, Swift, & Haaf, 1989; Williams & Watson, 1985). Likely explanations for this discrepancy include familiarity with the speaker population, as well as the stimuli used for intelligibility assessment. Moreover, naïve listeners' lack of common exposure to alaryngeal speech may lead to a shift in focus away from identification of speech sounds and toward the unnatural quality of voice, resulting in lower intelligibility scores. Given that the majority of individuals that TE speakers communicate and interact with are unfamiliar with alaryngeal voice, it is likely that intelligibility estimates derived from assessments by experienced listeners overestimate functional alaryngeal intelligibility. For this reason, most intelligibility studies, including those described below, have chosen naïve listeners to better approximate everyday TE intelligibility.

Intelligibility of Tracheoesophageal Speech

Since the introduction of the Blom-Singer method (Singer & Blom, 1980), TE speech has been studied extensively. Many studies have illustrated TE's superiority over other alaryngeal methods with respect to intelligibility (Blom, Singer, & Hamaker, 1986; Doyle, Danhauer, & Reed, 1988; Pindzola & Cain, 1988; Robbins, 1984; Tardy-Mitzell, Andrews, & Bowman, 1985; Williams & Watson, 1985). Intelligibility values have ranged anywhere from 65% to 100% (Doyle et al., 1988; Pindzola & Cain, 1988; Sleeth, 2012; Smith & Calhoun, 1994; Tardy-Mitzell et al., 1985). Between-study differences in intelligibility are likely the result of inconsistent methodological approaches. More specifically, the aforementioned studies utilized markedly different stimuli and response paradigms.

Pindzola and Cain (1988) and Tardy-Mitzell et al. (1985) chose a closed- or forced-choice response paradigm. That is, listeners' responses were confined to a limited number of options. In contrast, Sleeth (2012) and Doyle et al. (1988) used an open-choice response paradigm where listeners were required to identify stimuli through transcription rather than choosing from a limited number of response options. Smith and Calhoun

(1994) further illustrate the significance of different response paradigms by investigating TE intelligibility using both an open and closed-choice response paradigm. Interestingly, there was an 11% discrepancy in intelligibility between open and closed-response paradigms (82% vs 71%, respectively). To complicate matters further, all of the aforementioned studies (Doyle et al., 1988; Pindzola & Cain, 1988; Sleeth, 2012; Smith & Calhoun, 1994; Tardy-Mitzell et al., 1985) used different sets of stimuli, further adding to the variability in intelligibility.

Despite receiving extensive attention, few studies have investigated the specific error patterns contributing to overall TE intelligibility. Doyle et al. (1988) were the first to investigate TE intelligibility at the phonemic level. In their study, naïve listeners identified nonsense consonant-vowel-consonant-vowel-consonant (CVCVC) stimuli through the use of an open-response paradigm. Doyle et al. (1988) found stop-plosives, fricatives, and affricates to be the least intelligible manner classes (intelligibility = 63%, 60%, and 57%, respectively). Furthermore, they illustrated TE speakers' particular difficulty with voiced-voiceless distinctions. That is, voiceless phonemes were produced when voiced phonemes were intended.

Shortly thereafter, a similar study conducted by Doyle and Haaf (1989) showed similar findings regarding manner feature intelligibility hierarchy and voicing errors. However, rather than using CVCVC stimuli as Doyle et al. (1988) did, Doyle and Haaf (1989) used CVC stimuli to identify linguistic positional issues related to intelligibility. Their findings indicate that prevocalic consonants were less intelligible than postvocalic consonants. That is, the context of where a given sound appears within a target word (e.g., word-initial or word-final) directly influences intelligibility measures gathered.

Another study by Searl, Carpenter, and Banta (2001) corroborated the aforementioned studies' (Doyle et al., 1988; Doyle & Haaf, 1989) findings regarding voiced-voiceless distinction errors in TE speech through the use of nonsense CVCV stimuli. That is, the majority of errors came from the misperception of voiceless phonemes as voiced phonemes. Interestingly, voicing error was more prevalent for fricatives than stop-plosives. Once again, the issue of phonetic position was confirmed to be an important factor for consideration in assessments of intelligibility in postlaryngectomy speakers.

A more recent descriptive analysis of TE speech by Sleeth (2012) also found that voicing accounted for the majority of errors for prevocalic consonants of real-word CVC stimuli. More specifically, two-thirds of all prevocalic voicing errors were perceptions of voiceless for voiced phonemes. Postvocalic voicing errors followed a similar trend (i.e., mostly perceptions of voiceless for voiced phonemes). In contrast however, voicing did not account for the majority of errors for postvocalic consonants. Moreover, similar to Doyle et al. (1988), Sleeth (2012) found stop-plosives and fricatives to be among the least intelligible manner of articulation classifications (intelligibility= 80.99% and 81.19%, respectively). Therefore, patterns of perceptual errors based on the context of where a sound appears within a target word or stimulus must be considered. In this regard, and consistent with the very first findings for TE speakers provided by Doyle et al. (1988), word scores are unlikely to represent the true intelligibility of any given speaker.

Intelligibility of Tracheoesophageal Speech in Noise

While TE speech intelligibility has been studied extensively since the introduction of the Blom-Singer method (Singer & Blom, 1980), few studies have investigated the impact of the presence of background noise. Clark and Stemple (1982) were among the first to investigate alaryngeal intelligibility in the presence of background noise. Through the use of a closed-response synthetic sentence identification task, they assessed the relative intelligibility of four speakers (one TE, ES, EL, and a normal laryngeal speaker) who were judged by three experienced speech-language pathologists to have above-average voice quality. Speaker recordings of 10 synthetic sentences developed by Speaks and Jerger (1965) served as stimuli for their study. Additionally, 20 normal-hearing individuals between 19 and 30 years of age were exposed to 10 stimuli from each speech mode at three signal-to-noise ratios (SNRs) (0, -5, and -10 dB). SNR was used to indicate the relative volumes of speech signal and noise. For example, a -5 dB SNR indicates a speech signal which is 5 dB quieter than background noise. Conversely, a positive SNR (e.g., +5 dB SNR) would indicate that the speech signal is louder than the present noise by the specified magnitude. Listeners were asked to identify each synthetic sentence by choosing one of 10 options provided. Speaker intelligibility was determined as the percentage of correct responses in each SNR.

Interestingly, Clark and Stemple (1982) found that there were no significant differences in intelligibility between the four speech modes at the 0 dB SNR (TE intelligibility=99%), with intelligibility ranging from 97.5% to 100%. However, of the various speech modes examined, TE speech was found to have the lowest intelligibility in both the -5 and -10 SNR conditions (TE intelligibility=75.5% and 12.0%, respectively). This finding suggests that TE speech is particularly susceptible to signal degradation in the presence of noise compared to other alaryngeal speech modes. Clark and Stemple (1982) hypothesized this finding to be the result of the lack of listener familiarity with alaryngeal voice and/or the similarity between the fundamental frequency of TE and the masking noise, allowing TE to be more easily masked. However, spectral analysis is required to validate their speculations.

The effect of noise on alaryngeal speech intelligibility was also assessed in a similar study conducted by Clark (1985) shortly after the initial report by Clark and Stemple (1982). Clark (1985) employed the same methodological protocol as Clark and Stemple (1982), with the exception that there were two distinct listener groups to investigate the potential effect of listener age; the first consisting of 11 normal-hearing individuals between 21 and 30 years of age (younger listener group), and the second consisting of 11 normal-hearing individuals between 50 and 66 years of age (older listener group). Similar to Clark and Stemple's (1982) findings, there were no significant differences identified between speech modes at the 0 dB SNR for the younger listener group (TE intelligibility=100%). Moreover, TE had significantly lower intelligibility compared to EL and laryngeal speech in the -5 dB SNR (TE intelligibility=76.36%). However, there was no significant difference noted between TE and ES intelligibility. For the -10 dB SNR, TE was found to be the least intelligible speech mode for the younger group (TE intelligibility=0%). The older group displayed similar results overall, with the exception that there was no significant difference between TE and normal laryngeal speech at the -10 dB SNR.

A dissertation by Dudley (1984) investigated the comparative intelligibility of TE, ES, and laryngeal speech in the presence of a multi-talker noise competitor. Dudley’s study employed a closed-response paradigm to investigate the intelligibility of 24 average alaryngeal speakers (TE=12, ES=12) and 12 normal laryngeal speakers. Stimuli consisted of 30 monosyllabic words chosen from a modified version of the Northwestern University Intelligibility Test (1975). All words were read by the speakers in a standard carrier phrase; “*Say the word* ___ *again*”. For ES speakers, the carrier phrase was shortened to, “*Word* ___ *again*”, to account for shorter phonation durations. Fifty-four young naïve adult listeners (range: 21-33 years of age) were exposed to all words in three different listening conditions (quiet, 0 dB, and +6 dB SNR) and were asked to identify the target words by choosing one of four selection options.

Dudley (1984) found that there was no significant interaction effect between speech mode and noise condition on overall intelligibility. Furthermore, there was a significant main effect of noise (i.e., as SNR decreased, intelligibility decreased). TE intelligibility scores were reported to be 78.7%, 75.5%, and 78.7% in the quiet, +6 dB, and 0 dB SNR conditions, respectively. Lower intelligibility findings relative to Clark (1985) and Clark and Stemple (1982) are likely due to differences in speaker proficiencies (i.e., average vs. excellent speakers). Target analysis revealed similar TE error patterns between quiet and 0 dB SNR conditions. More specifically, fricatives were found to be the least intelligible manner of articulation, followed by stop-plosives, affricates, nasals, glides, and laterals, respectively. However, it is important to note that stop-plosives and fricatives had the highest frequency of occurrence (approximately six times as frequent as nasals, glides, affricates, or laterals). Examination of raw scores revealed differences in intelligibility between quiet and noise to be the result of consistent increases in the number of errors across all manners of articulation.

Fourteen years later, McColl et al. (1998) investigated the intelligibility of one superior TE and one normal laryngeal speaker in nine different noise conditions (quiet, +20 dB, +15 dB, +10 dB, +5 dB, 0 dB, -5 dB, -10 dB, and -15 dB SNR) through the use of a scaling procedure. Audio recordings of speakers producing the following sentence pair from Fairbanks' (1960) Sentences of Phonetic Inventory served as stimuli for their study*: "Part way up the slope above the pool was a popular camping spot. Many people stopped there for picnic supper".* Fifty listeners were each presented with nine samples from each speaker - eight of which were accompanied by multi-talker noise presented at various SNRs - and asked to assign each sample a number according to how well it was understood. Higher numbers indicated lower intelligibility. McColl et al. (1998) found a significant interaction effect between speech mode and noise condition. That is, the effect of noise was dependent upon speech mode. Moreover, and as expected, TE intelligibility rankings declined as the SNR decreased (i.e., noise level increased).

Summary

After careful examination and review of the literature, the following conclusions can be drawn regarding TE speech intelligibility in quiet. First, TE speech is a viable method of postlaryngectomy voice production with speech intelligibility ranging anywhere from 65% to 100%. Second, the most common class of errors specific to manner of production occur with stop-plosives and fricatives. Third, TE speakers have been shown to exhibit a particular difficulty with voiced-voiceless distinctions. Finally, research has shown that postvocalic consonants tend to be more intelligible overall when compared to prevocalic consonants. In addition to the above information, the following conclusions can also be drawn for TE intelligibility in noise: 1) as the intensity of noise increases relative to the speech signal, there is a decrease in overall intelligibility, and 2) while there is one study which investigated specific error patterns in noise, limitations regarding the choice of stimuli make it difficult to make clear conclusions. Thus, questions related to speech intelligibility in the context of competing noise persist in relation to TE speech production.

Statement of Problem

Despite potentially functional levels of postlaryngectomy communication following total laryngectomy, alaryngeal speakers have significantly lower intelligibility compared to normal laryngeal speakers (Dudley, 1984; Searl et al., 2001; Williams & Watson, 1987). While TE speech has received a considerable amount of attention over the last several decades, only a handful of studies have investigated intelligibility in the presence of background noise (Clark, 1985; Clark & Stemple, 1982; Dudley, 1984; McColl et al., 1998). Of these studies, only one (Dudley, 1984) has directly investigated specific error patterns influencing overall intelligibility. However, limitations in phonemic representation make it difficult to make clear conclusions. Another significant concern regarding the validity of the data regarding TE intelligibility relates to the remote time period in which these studies were conducted. Since the early 1980s, there have been considerable advances in TE prosthesis technology and management which may lend to changes in intelligibility characteristics.

Given that it is seldom the case that communication occurs in ideal quiet conditions, it is imperative that additional research that is designed to better approximate TE speakers' less-than-ideal daily speaking environment be conducted. Such information is clearly necessary if we are to better understand the communication difficulties faced by today's TE speakers. As such, the present investigation sought to: 1) identify and compare overall stop-plosive and fricative TE intelligibility in quiet and noise, and 2) analyze and compare specific error patterns in quiet and noise.

Chapter 2

Method

The current study was conducted in accordance with Tri-Council Policy ethical guidelines. Ethics approval was obtained from Western's Research Ethics Board (110335).

Participants

Speakers. Twelve male (M = 66.8, range = 50 – 84 years) and 2 female (M = 51.3, range = 39 – 60 years) native English speaking individuals who underwent total laryngectomy as treatment for advanced laryngeal cancer served as participant-speakers. All used TE speech as their primary mode of verbal communication and all were judged by an experienced clinician to have excellent intelligibility. Consent was obtained from all speakers prior to obtaining voice recordings.

Listeners. Eight male (M = 24.1, range = 21 – 28 years) and 10 female (M = 22.3, range = 20 – 26 years) native English speaking adults between the ages of 18 and 30 with no prior hearing difficulties served as voluntary participant-listeners. None had formal training in voice disorders or prior exposure to alaryngeal speech and, thus, were considered to be naïve listeners. All were recruited from undergraduate courses in the faculty of Health Sciences at Western University using the script approved by Western's Research Ethics Board (see Appendix A). Listeners were provided with a letter of information and written informed consent was obtained prior to study participation (see Appendix B).

Experimental Stimuli and Speaker Recordings

All speaker recordings were obtained from an archival voice database located in the Voice Production and Perception Laboratory at Western University in London, Ontario. Participant-speakers were recorded while reciting a list of real English monosyllabic consonant-vowel-consonant (CVC) words originally developed and reported by Weiss and Basili (1985) using a cardioid condenser microphone affixed to a microphone stand. A 15-cm microphone-to-mouth distance was maintained throughout all recordings. No carrier phrase was used in the acquisition of voice samples. This was done in an effort to avoid the influence of onset phenomena associated with the use of a carrier phrase. All samples were recorded at 44.1 kHz using a preamplifier and Kay-Pentax Sona Speech II software (Pine Brook, NJ) in a professional quality recording environment free of ambient noise.

Stimulus Preparation. Fifteen CVC words were chosen from Weiss and Basili's (1985) 66-item word list for use in this study. These 15 stimuli were extracted from pre-recorded speaker samples so that six stop-plosive (/p/, /t/, /k/, /b/, /d/, /g/) and seven fricative (/f/, /θ/, /s/, /ʃ/, /v/, /ð/, /z/) consonants were each represented in both word-initial and word-final positions (see Appendix C). The extracted words were normalized relative to average vowel intensity for use in the quiet condition after removing extraneous noise (see Appendix D). One second of silence was concatenated to the beginning and end of each word.

Quiet condition stimuli were duplicated and overlaid with a multi-talker noise complex obtained from the National Center for Audiology for use in the noise condition. Multi-talker noise was overlaid such that a signal-to-noise ratio (SNR) of +3 decibels (dB) was achieved.

Auditory signal distortion and/or clipping were prevented by ensuring that no stimulus exceeded a maximum intensity of -2 dB relative to full scale - approximately 80% of the maximum signal amplitude an electronic device can handle without audio signal distortion - after both normalization and noise overlay procedures. All stimuli were prepared using Audacity version 2.0.6 (Pittsburgh, PA).

Word Lists. A total of 12 randomized word lists (6 quiet and 6 noise) were created. Each list was comprised of 252 samples (210 stimuli + 42 duplicate samples for agreement assessment) presented as separate ".wav" audio files in a pseudorandomized sequence. Word lists were pseudorandomized to ensure that presentation of the original sample preceded presentation of the corresponding duplicate sample. Three predetermined stimuli from each speaker (3 words X 14 speakers = 42 duplicate samples) were repeated in a randomized sequence in each word list using a split-half reliability protocol to assess for the presence of potential listener learning and exposure bias (see Appendix C). Duplicate samples were randomly assigned to be presented as a group midway through, or at the end of each word list. All sequence randomizations were carried out using a batch script developed by the author (see Appendix E).

Listening Procedure

Listeners participated in a single listening session that took place in the Voice Production and Perception Laboratory at Western University. The session lasted approximately 70 minutes (range = 53 – 95 minutes). Participants were presented two word lists (one quiet and one noise) in a counterbalanced fashion. That is, the first listener was presented with the quiet condition first, followed by the noise condition; the second listener was presented with the noise condition first, followed by the quite condition, and so on. Following the acquisition of consent, listeners were instructed to listen to and then transcribe, in Standard English orthographics, each word heard onto an answer sheet. Listeners were requested to listen to each sample in a sequential manner (see Appendix F). As noted, the presentation of word lists was counterbalanced to prevent any potential order bias. All samples were presented binaurally through Sony MDRV-150 stereo headphones at a volume determined to be comfortable by each listener.

Participants were permitted to listen to each stimulus sample as many times as desired prior to transcribing each item. However, once a determination was made and transcription occurred, participant-listeners were asked to proceed to the next stimulus without returning back to any prior transcription. All perceptual identifications were to be made independent of all other samples. Participants were also instructed to leave partial or complete blanks on stimuli they were unable to make judgements on. That is, they were asked to transcribe any portion of the word (i.e., one or more of the CVC phonemes) they heard and indicate those they were unable to identify with the use of an underscore (e.g., “pa_”) on the transcription sheet. Listeners were permitted to take as many breaks as desired while completing the listening task. However, all listeners were required to take a mandatory 10-minute break after completing transcription of the first word list, regardless of whether they were presented with the quiet or noise list first.

A debriefing session followed the completion of the entire listening task to ensure that the task was completed as instructed. To prevent potential transcription errors, participants were asked if they could distinguish the Standard English orthography of “teeth” and “teethe”. Listeners who were unable to make this distinction were asked to return to a series of samples and listen to ensure their transcription corresponded to their initial judgment. Corrections were made to responses which did not correspond to the initial judgment. At this time, participants were also given the opportunity to ask any questions about the study.

Data Analysis

All orthographic listener responses were transcribed using the International Phonetic Alphabet (IPA) and recorded into an Excel database for intelligibility analysis by the author and four volunteers. All volunteers were properly instructed on proper transcription practices and monitored periodically during transcription. Additionally, all database entries were checked for quality control purposes by the author prior to final intelligibility calculations. Whole-word, word-initial, and word-final stop-plosive and fricative intelligibility were calculated for each speaker by collapsing intelligibility scores across listeners. Whole-word intelligibility was calculated as the percentage of correctly identified words. In order for a word to be considered correct, all word-initial, vowel, and word-final phonemes had to be correctly identified. Similarly, word-initial and word-final intelligibility were calculated as the percentage of correctly identified phonemes in their respective position and manner of articulation classification.

A repeated-measures multivariate Hotelling's T^2 test was conducted to compare whole-word, word-initial, and word-final intelligibility in quiet and noise conditions. Hotelling's T^2 was followed up with five univariate paired sample t-tests. A Bonferroni correction was applied to all univariate follow-up tests to control for alpha inflation. Thus, alpha was set to 0.01 to evaluate the significance of each univariate test. Statistical analyses were conducted using the 'ICSNP' statistical package on R computing software version 3.4.4 (Vienna, Austria). Word-initial and word-final confusion matrices for the control and noise conditions were constructed to allow for categorization of errors based on distinctive feature analysis.

Assessment of Listener Agreement

Intra-rater agreement was assessed through direct sample-by-sample analysis. Inconsistencies between original judgments of samples and their corresponding duplicate samples were assessed and used for agreement calculation. Agreement was calculated as the percentage of consistent phonemes relative to the total number of agreement phonemes. Sample responses were considered to be consistent if transcription was identical on the corresponding duplicate sample. Note that samples could be considered consistent irrespective of whether or not target phonemes were correctly identified. Listener learning and exposure bias was assessed by calculating two separate agreements for each condition; one for the agreement samples presented midway through, and one for those presented at the end of each condition.

Chapter 3

Results

Absolute Agreement

Overall intra-rater absolute agreement ranged from 81.4% to 97.2% (*M = 92.2%*). First and second half agreement for each condition remained relatively consistent across all listeners, indicating no observable listener learning or exposure bias (see Table 1). A considerable variation in agreement was noted between speakers. Examination of individual listener task completion times revealed no discernable pattern among listeners with lower agreement scores. Although all speakers reported normal hearing, significant differences in hearing thresholds may have existed and contributed to some listeners having more difficulty with the listening task compared to others. Additionally, it is possible that some attentional differences among listeners may have been present during the task.

Intelligibility

Whole-word intelligibility scores were based on a total of 3,780 observations (15 stimuli X 14 speakers X 18 listeners) for each condition. Word-initial and word-final stop-plosive intelligibility scores were comprised of 1,512 observations (6 stimuli X 14 speakers X 18 listeners) for each condition. Word-initial and word-final fricative intelligibility scores were comprised of 2,142 observations (7 stimuli X 14 speakers X 16 listeners) for each condition. A total of 22,176 observations were used for this investigation. All intelligibility scores were derived from the constructed confusion matrices (see Figures 1-4). Note that the diagonal, indicated by shaded cells in Figures 1 through 4, indicates correct responses by listeners. The maximum possible for each of these shaded cells is 252.

All statistical test assumptions were met (i.e., multivariate normality). Multivariate analysis revealed a significant main effect of noise; $T^2(4, 10) = 11.41$, $p = .001$. Follow-up univariate tests revealed significant findings in all noise-quiet comparisons with the exception of word-initial stop-plosive and word-final fricative intelligibility; $t(13) = 2.34$, $p = .036$; $t(13) = 52.15$, $p = .051$, respectively. Whole-word intelligibility was significantly lower in noise; $t(13) = 5.49$, $p < .001$. Similarly, word-initial fricatives and word-final stop-plosives were significantly less intelligible in noise; $t(13) = 3.80$, $p = .002$; $t(13) = 6.51$, $p < .001$, respectively (see Table 2).

Listener Response

Consonant Target	**p**	**b**	**t**	**d**	**k**	**g**	**s**	**z**	**f**	**v**	**θ**	**ð**	**ʃ**	**ʒ**	**tʃ**	**dʒ**	#	**Int**	**Other**
p	40	180							9		2	1					8		12
b	3	238				1			1		1						2		6
t	1		191	16					2	1	5	14					11	1	11
d			1	234													4		13
k			2	1	136	104											3	1	6
g				9	17	163											15		48
s							212	18		2			1				3		16
z		1					87	150		2			2			3	6		1
f	1	6							226	1	6	7					5		
v		13							5	137	2	71	1				16	1	7
θ				1			1		3		238	5					3		1
ð		2		2							1	240					2		5
ʃ						1	23						218		2	7			1

Figure 1. Word-Initial phonemes in quiet. Target stimuli are shown along the vertical axis and responses are shown along the horizontal axis. Shaded cells indicate correct responses. # indicates the absence of a consonant. *Other* indicates other responses including consonant blends and other consonants or vowels. *Int* indicates intrusions - the introduction of a phoneme.

Listener Response

Consonant Target	p	b	t	d	k	g	s	z	f	v	θ	ð	ʃ	ʒ	tʃ	dʒ	#	Int	Other
p	246	2								1							3	1	
b	5	231		1													10	4	5
t	1		221	9	1				1		2						14	7	3
d			4	227		1					10						4	2	6
k	4		2		233	11											2		
g		1		18	6	220		1								1	5		
s							212	3			11		4		11		4	8	7
z								249				1					2		
f	4			2			5	3	155	46	11	7					18	2	1
v	2									247							1		2
θ			5	9			5	1		1	220	3	2				5		1
ð	1		3	7		1		9	2	10	137	56					22	5	4
ʃ							9	3					225	5		1	6	1	3

Figure 2. Word-Final phonemes in quiet. Target stimuli are shown along the vertical axis and responses are shown along the horizontal axis. Shaded cells indicate correct responses. # indicates the absence of a consonant. *Other* indicates other responses including consonant blends and other consonants or vowels. *Int* indicates intrusions - the introduction of a phoneme.

Listener Response

Consonant Target	p	b	t	d	k	g	s	z	f	v	θ	ð	ʃ	ʒ	tʃ	dʒ	#	Int	Other
p	42	157		1					17	1		1					18		15
b	1	212				2	4		8	1		6					5		13
t			203	16	1	1			3	4	3	7			1		6	1	7
d			2	207		1			1			1					11		29
k	1		2	1	140	94					2	1					9	2	2
g		3		14	32	129	1	1	2			2					19		49
s							213	13	3	1			1				4		17
z		2					119	106	3	7	1	1					13		
f		2							235	4	2	4					5		
v		14				1			24	133	6	43					13	1	18
θ	1	1		2					9	1	218	7					7	1	6
ð				3			2		10		3	205					8	1	21
ʃ							23						211		4	10	4		

Figure 3. Word-Initial phonemes in noise. Target stimuli are shown along the vertical axis and responses are shown along the horizontal axis. Shaded cells indicate correct responses. # indicates the absence of a consonant. *Other* indicates other responses including consonant blends and other consonants or vowels. *Int* indicates intrusions - the introduction of a phoneme.

Listener Response

Consonant Target	p	b	t	d	k	g	s	z	f	v	θ	ð	ʃ	ʒ	tʃ	dʒ	#	Int	Other
p	227		1	1		1			1	5							13	3	3
b	16	188	8	5		1			1	3	3						17	10	10
t	4	2	188	24							1						16	14	17
d	1	2	8	200		5					14						6	5	16
k	7		2	5	226	10											1	4	1
g			1	23	5	191				2			1				13	2	16
s			1	1	1		182		2		20		4		13		9	10	19
z							2	242									6	2	2
f	4		2	2		1	3	5	154	37	18	5					18	2	3
v	1			4	1					221							11		14
θ		1	10	15			5	1			208	1	1				6		4
ð		2	4	5		2		4		5	135	56					23	3	16
ʃ						1	4	4					231	4	1		6	3	1

Figure 4. Word-Final phonemes in noise. Target stimuli are shown along the vertical axis and responses are shown along the horizontal axis. Shaded cells indicate correct responses. # indicates absence of a consonant. *Other* indicates other responses including consonant blends and other consonants or vowels. *Int* indicates intrusions - the introduction of a phoneme.

Distribution of Errors by Manner and Position

Word-Initial Stop-Plosives. A total of 512 and 582 stop-plosive errors were identified in the word-initial position in quiet and noise, respectively (see Figures 1 & 3). The majority of errors in the quiet condition were attributed to voiced-voiceless phoneme confusions (66.4%). Similarly, 57.8% of errors were attributed to voiced-voiceless phoneme confusions in the noise condition. Although there were more overall errors in noise, there were fewer voiced-voiceless distinction errors compared to the quiet condition. “Other” errors were also notable contributors of error in both the quiet (24.8%) and noise (30.2%) conditions (see Table 3). That is, a significant proportion of errors were attributed to confusions for other phonemes of similar voicing status, blends, or other phoneme classes including affricates, nasals, and liquids.

Word-Final Stop-Plosives. A total of 148 and 330 stop-plosive errors were identified in the word-final position in the quiet and noise conditions, respectively (see

Figures 2 & 4). Overall, there were fewer stop-plosive errors in the word-final position compared to the word-initial position in both quiet and noise conditions. In contrast to the word-initial position, “other” errors were the largest contributors of error in both the quiet (32.4%) and noise (36.4%) conditions. Voiced-voiceless distinction errors were also significant contributors of error in both quiet (32.4%) and noise (32.1%) conditions (see Table 3).

Table 1. Absolute agreement of listeners in percent consistency

Listener	Quiet		Noise		Overall
	I	II	I	II	
1	90.48	96.83	88.89	88.89	91.27
2	88.89	90.48	92.06	95.24	91.67
3	80.95	84.13	77.78	82.54	81.35
4	98.41	96.83	88.89	90.48	93.65
5	82.54	88.89	82.54	90.48	86.11
6	96.83	100.00	96.83	95.24	97.22
7	98.41	92.06	96.83	90.48	94.44
8	95.24	92.06	96.83	95.24	94.84
9	95.24	100.00	95.24	96.83	96.83
10	98.41	96.83	79.37	96.83	92.86
11	98.41	98.41	79.37	93.65	92.46
12	93.65	100.00	82.54	82.54	89.68
13	96.83	98.41	88.89	93.65	94.44
14	92.06	98.41	90.48	96.83	94.44
15	95.24	100.00	88.89	84.13	92.06
16	95.24	93.65	76.19	82.54	86.90
17	100.00	95.24	87.30	90.48	93.25
18	100.00	96.83	90.48	95.24	95.63

Note. Listeners 1-9 completed the quiet condition first and listeners 10-18 completed the noise condition first. *I* indicates the reliability portion completed midway through its respective condition. *II* indicates the reliability portion completed at the end of its respective condition.

Word-Initial Fricatives. A total of 344 and 446 fricative errors were identified in the word-initial position in the quiet and noise conditions, respectively (see Figures 1 & 3). "Other" errors accounted for almost half (47.1%) of fricative errors in the quiet condition. In contrast, perceptual confusions between voiced and voiceless phonemes accounted for the largest proportion of error (47.8%) in noise. Voiced-voiceless distinction errors (42.4%) and "other" errors (39.5%) were notable contributors of error in both quiet and noise, respectively (see Table 3).

Word-Final Fricatives. A total of 416 and 490 fricative errors were identified in the word-final position in the quiet and noise conditions, respectively (see Figures 2 & 4). Overall, there were more fricative errors in the word-final position compared to the word-initial position in both quiet and noise. The majority of fricative errors in quiet were attributed to perceptual confusions between voiced and voiceless phonemes (55%). Similarly, 45.1% of errors in noise were attributed to voiced-voiceless distinction errors. Although there were more overall errors in noise, a smaller proportion of errors were due to confusions between voiced and voiceless phonemes (45.1%) (see Table 3).

Table 2. Intelligibility means, standard deviation, minimum, and maximum (%)

		Mean ± SD	Min	Max
*Whole-Word				
	Quiet	65.71 ± 10.17	49.63	80.74
	Noise	58.23 ± 10.96	41.11	74.81
Word-Initial				
Stop-plosives				
	Quiet	66.27 ± 8.79	51.85	82.41
	Noise	61.71 ± 14.29	30.56	76.85
*Fricatives				
	Quiet	80.56 ± 10.85	60.32	95.24
	Noise	74.89 ± 10.80	55.56	89.68
Word-Final				
*Stop-plosives				
	Quiet	91.14 ± 9.29	69.44	100
	Noise	80.69 ± 13.46	50	99.07
Fricatives				
	Quiet	77.32 ± 8.48	60.32	88.10
	Noise	73.36 ± 10.76	54.76	89.68

Note. Paired t-tests were performed to compare quiet and noise for each manner of articulation class at each locus.
*Significant, $p < .01$

Table 3. Distribution of errors by condition, position, and manner of articulation (raw)

	Cognate		Non-Cognate		Omissions	Intrusions	Other	Total
	V+/V-	V-/V+	V+/V-	V-/V+				
Quiet								
Word-Initial								
Stop-plosives	21	300	2	17	43	2	127	512
Fricatives	93	24	5	24	35	1	162	344
Word-Final								
Stop-plosives	15	22	10	1	38	14	48	148
Fricatives	137	57	8	27	58	16	113	416
Noise								
Word-Initial								
Stop-plosives	35	267	16	17	68	3	176	582
Fricatives	146	24	22	21	54	3	176	446
Word-Final								
Stop-plosives	29	34	29	14	66	38	120	330
Fricatives	137	42	6	36	79	20	170	490

Note. Other includes responses such as blends, and other consonant responses without voiced-voiceless distinction error. V+/V- ; (target/response) voiced targets incorrectly identified as voiceless phonemes; V-/V+ ; (target/response) voiceless targets incorrectly identified as voiced phonemes.

Chapter 4

Discussion

The current study was designed to address questions related to the intelligibility of phonemes produced by tracheoesophageal (TE) speakers. At present, there is a paucity of literature pertaining to phonemic TE intelligibility obtained in the presence of less-than-ideal communication environments. More specifically, the present study sought to investigate stop-plosive and fricative intelligibility in the presence of a multitalker noise competitor. In designing this project, an effort was made to equally represent stop-plosive and fricative phonemes across word-initial and word-final positions within stimuli; additionally, the stimuli used were comprised of real-word consonant-vowel-consonant (CVC) stimuli.

In the sections to follow, findings pertaining to whole-word intelligibility will first be addressed, followed by a discussion of phonemic intelligibility stratified by manner of articulation and position. Finally, the distribution of perceptual errors will be presented. Accordingly, aspects of both the strengths and limitations of this study will also be outlined. The final sections of the chapter will present the clinical implications of the current study and directions for future research.

Whole-Word Intelligibility

Examination of individual speaker intelligibility scores in quiet showed considerable between-speaker variation with a spread of approximately 30%. This finding has also been demonstrated and reported in the literature (Dudley, 1984; Searl et al., 2001; Sleeth, 2012). Based on the variability in intelligibility scores generated, it is clear that TE speakers are not “similar” in their communicative abilities despite some incorrect, anecdotal assumptions. These data indicate that TE speakers are unique.

Nevertheless, our intelligibility scores obtained in the quiet condition were found to be comparable to previous investigations conducted by Doyle et al. (1989) and Sleeth (2012). However, scores of the present TE speakers were substantially lower than those reported by Pindzola and Cain (1988) and Tardy-Mitzell et al. (1985). Discrepancies between intelligibility scores of the current study and those in the literature may be largely attributed to variability in methodological approaches (i.e., stimuli, speaker proficiency, response paradigm, etc.).

More specifically, the current study utilized an open-response paradigm to gauge intelligibility in both quiet and noise, whereas Pindzola and Cain (1988) and Tardy-Mitzell et al. (1985) used a forced-choice response paradigm. The use of the latter response method places strict limitations on listeners' responses such that response errors are limited to a pre-selected, a priori set of options. Moreover, there is an opportunity to correctly identify words by 'guessing' based on the perception of partial information. As a result of the limited number of responses associated with closed-set formats, there may be a subsequent inflation of intelligibility scores generated.

Similar to the intelligibility findings in quiet, there was considerable variation between our intelligibility in noise findings and those documented in prior investigations (Clark, 1985; Clark & Stemple, 1982; Dudley, 1984; McColl et al., 1998). Once again, discrepancies could be explained by differences in methodological approaches. However, in this case, between-study variability in intelligibility findings can be largely attributed to two factors pertaining specifically to noise – noise level and noise type. Despite there being only a handful of studies investigating TE intelligibility in noise, there is a lack of consistency in the level of noise applied to TE speech signals (i.e., SNRs). This suggests that the level of competition may be a critical factor. While it is obvious that competing noise will influence the perception of even normal speech, the nature of TE speech may in fact reveal very specific profiles of intelligibility challenge.

The second factor relates to differences in the type of noise used between studies. In the context of normal speech production, past research has clearly documented that the type of noise used influences speaker intelligibility scores (Danhauer et al., 1985). Further, Danhauer et al. (1985) found that the masking efficiency of a noise competitor increased as it became more similar to the speech signal it was masking. These findings are also likely to influence degraded speech signals such as those assessed in the current research project. Although most studies have chosen to use multitalker noise as the masker for TE speech signals, few studies have used the same variant of the masker. That is, studies have failed to use multitalker maskers derived from the same speakers and/or number of speakers. Slight variations in multitalker noise (i.e., number of speakers and quality of speakers) may lead to acoustic changes resulting in altered masking properties, thus, altering measures of intelligibility. Regardless of differences in multitalker noise, the present finding that whole-word intelligibility was significantly lower in the presence of noise was consistent with similar investigations conducted by Clark (1985), Clark and Stemple (1982), Dudley (1984), and McColl et al. (1998). Thus, the external validity of data gathered in the current investigation are supported.

Manner Intelligibility Stratified by Position

Examination of intelligibility by manner and position revealed similar intelligibility hierarchies in both quiet and noise conditions. More specifically, in the word-initial position, fricatives were more intelligible than stop-plosives in both the quiet and noise conditions (80% vs. 66%, and 74% vs. 61%, respectively). In the word-final position, however, stop-plosives were more intelligible than fricatives in the quiet and noise conditions (91% vs. 77%, and 80% vs. 73%, respectively). These findings support those reported by Doyle and Haaf (1989), who also found a similar shift in intelligibility hierarchy between word-initial and word-final positions. Based on these data, even at the simplest level, phonetic position of phoneme is of importance to TE speech intelligibility.

It appears that the hierarchical shift observed between word-initial and word-final positions was largely due to considerable increases in stop-plosive intelligibility. In particular, there was a 25%, and 19% increase in stop-plosive intelligibility in quiet and noise, respectively, when shifting from the word-initial to the word-final position. Doyle and Haaf (1989) hypothesized improved word-final intelligibilities to be the result of phonetic context. They state that consonants preceded by vowels may have augmented acoustic cues, allowing easier identification by the listener.

Interestingly, and in contrast to Doyle and Haaf's (1989) findings, there appeared to be no positional effect on fricatives in either condition (80% vs. 77% in quiet and 74% vs. 73% in noise). This finding may be the result of the continuant nature of fricatives. More specifically, the relatively long duration of fricatives may allow them to be more easily coded, thus, allowing significant leniencies regarding their perception as compared to phonemes with much shorter durations. In fact, during the production of normal speech, fricatives may have durations ranging anywhere from 50-200 milliseconds without a perceptual degradation. That is, despite significant temporal deviations (i.e., relative shortening or lengthening), perceptual identification of fricatives remains intact. It should also be noted that TE speakers may unknowingly also extend the duration of their speech due to their ability to access a relatively substantial power supply from the lungs (Moon & Weinberg, 1987; Robbins et al., 1984; Weinberg, Horii, Blom, & Singer, 1982).

The current finding that intelligibility remained consistent across conditions for fricatives in the word-final position (77% vs. 73%) may have also been the result of the relatively long duration of fricatives. That is, despite being masked by noise, a sufficient amount of acoustic cues may have remained intact in word-final fricatives, thereby, allowing listeners to identify word-final fricatives as easily as in the quiet condition. In contrast, word-initial fricatives were found to be significantly less intelligible in noise. However, closer examination of these data revealed that the difference was only 6% as compared to the nonsignificant difference of 4% obtained in the word-final position. Given these minute differences, it is not unreasonable to hypothesize that the effect of noise on intelligibility may be mitigated by a similar phenomenon in the word-initial position. Although there is a clear issue regarding fricative intelligibility for TE speakers, the latter findings suggest that the continuant nature of fricatives may provide protection, at least to some extent, against perceptual degradation of phonetic entities in the presence of noise. However, it is important to note that manner class is also of importance in understanding TE speech intelligibility. This is certainly true relative to the production and perception of stop-plosives.

The current investigation revealed that stop-plosives were much more susceptible to noise as compared to fricatives. While there was only a 5% decrease in word-initial stop-plosive intelligibility when presented in noise (66% vs. 61%), there was an 11% decrease in intelligibility of word-final stop-plosives (91% vs. 80%). This finding suggests that stop-plosives are particularly susceptible to noise and positional effects. More specifically, word-final stop-plosives appear to be the most susceptible to perceptual degradation in the presence of noise. This finding can likely be explained by the relative ease with which stop-plosives are coded in comparison to fricatives. In contrast to fricatives, stop-plosives are composed of several short segments (i.e., stop gap, release burst, aspiration and voice onset time), each providing necessary acoustic perceptual cues. In reality, stop phonemes are relatively brief in their construct. Given the relatively short durations, durational alterations may leave insufficient acoustic information for perception. Moreover, the relative complex structure of stop-plosives may result in these phonemes being coded in such a way that permits fewer leniencies regarding missing acoustic information. Therefore, stop-plosives' greater susceptibility to noise and positional effects may be the result of a more complex perceptual coding of these specific phonemes.

Error Analysis

Further examination of the error type distribution by manner of production, as well as by word-initial and word-final position showed that overall error patterns tended to remain consistent across conditions. In general, voicing errors and errors classified as other were the most prominent error type followed by omissions, and intrusions, respectively. While word-initial fricatives and word-final stop-plosives displayed slight variations in error type hierarchy across conditions, differences were limited to reversals of the most prominent and second most prominent error types (i.e., voicing confusions and errors classified as other). Moreover, in cases where other errors were more common than voicing errors, differences were limited to less than 20 errors (an approximately 10% difference between voicing errors and errors classified as others), suggesting that the change in hierarchy may have simply been a chance occurrence. Overall, these collective findings suggest that despite slight differential noise effects on manner of articulation and position, the hierarchical distribution of errors of the current stimuli remain relatively intact in the presence of noise.

Voicing Distinctions

The current study identified voiced-voiceless confusions to be one of the most common types of error overall in both the noise and quiet conditions. These types of errors include those in which voiceless targets are perceived as voiced phonemes (i.e., /p/ perceived as /b/), as well as those in which voiced targets are perceived as voiceless phonemes (i.e., /b/ perceived as /p/). This finding is consistent with previous investigations of TE speech (Doyle et al., 1988; Doyle & Haaf, 1989).

As expected, the majority of errors for stop-plosives in the word-initial position in both the quiet and noise conditions consisted of voiceless targets being misperceived as voiced phonemes. Given the fact that the alaryngeal voice source for TE speech lacks active adductor-abductor functionality, this perceptual pattern of errors is not unexpected. That is, the ability to fine-tune vibratory “on-off” control that serves to signal voicing distinctions is lost with the PE segment used in TE speech. In contrast however, voiced-for-voiceless and voiceless-for-voiced confusions were equally prevalent for stop-plosives in the word-final position in both the quiet and noise conditions. Word-initial and final fricatives, on the other hand, exhibited a different pattern of voicing errors when the present listener data were evaluated. In fact, for these phonemes (i.e., word-initial and final fricatives), the majority of errors came from voiced targets being perceived as voiceless phonemes in both the quiet and noise conditions. Doyle et al (1988) contend that voicing distinction errors (both voiced-for-voiceless and voiceless-for-voiced) may be the result of changes in temporal aspects pertaining to phonatory onset/offset of the PE mucosa. In particular, because the voicing source (i.e., the PE mucosa) of TE speakers lacks the fine motor adductor/abductor control required for temporal tuning of voicing cues (i.e., voice onset time), there may be a resultant altered voicing cue, and ultimately voicing perception (Doyle et al., 1988). The ultimate issue is that TE speech has access to a pulmonary air source; high volumes of air as a driving source for the alaryngeal vibratory source has considerable capacity to modify how the source vibrates, as well as how the upper airway influences patterns of vibration. As such, airway changes and the creation of turbulent noise sources secondary to increased air volumes may create voiceless for voiced misperceptions.

Despite similar voicing error patterns when comparing the quiet condition of each manner in each position to its respective noise condition, careful examination of how errors changed revealed a differential effect of noise according to manner of articulation and position. More specifically, word-initial stop-plosives had an equal increase and decrease in the number of voiced targets perceived as voiceless phonemes (+28 errors) and voiceless targets perceived as voiced phonemes (-32 errors) when comparing the quiet and noise conditions. This resulted in no net change in the number of voicing errors for word-initial stop-plosives despite the introduction of noise. In contrast, word-final stop-plosives had a roughly equal increase in both voiced-for-voiceless (+25 errors) and voiceless-for-voiced (+33 errors) confusions when comparing the quiet and noise conditions.

Word-initial fricatives showed a different pattern of voicing error change when shifting from quiet to noise; with a general increase in the number of voiced targets perceived as voiceless phonemes (+70 errors) and no change in the number of voiceless targets perceived as voiced phonemes (-3 errors). Word-final fricatives, yet again showed a different voicing error change pattern; these phonemes showed virtually no change in both voiced-for-voiceless (-6 errors) and voiceless-for-voiced (-2 errors) confusions despite the introduction of noise. These findings suggest that the effect of noise on voicing feature may differ according to both the manner of production and the phonetic position where the phoneme appears. In particular, it appears that word-final fricative voicing cues are resistant to the effect of the noise used for this investigation. However, these phenomena must be investigated further to gain greater insight as to why the observed patterns exist.

Summary

The current study was designed to address questions related to the intelligibility of phonemes produced by tracheoesophageal (TE) speakers. A particular focus of this work centered on describing word-initial and word-final stop-plosive and fricative intelligibility in noisy environments. While TE speech has been regarded as a viable and preferred method of postlaryngectomy verbal communication, it does not come without limitations. The combination of an altered voicing source and the lack of a proper adductor-abductor voicing mechanism makes TE speech particularly susceptible to intelligibility deficits, regardless of whether listening occurs in quiet or in the presence of competing noise. As indicated by the current findings, it is evident that TE speakers have significant difficulty communicating even in ideal speaking conditions. Communication difficulties only become further exacerbated with the introduction of competing noise.

The fact that the presence of noise had differential effects on specific components of intelligibility highlights the importance of two key factors when considering intelligibility of TE speakers. First, phonemic position of phonemes is in fact important. This was clearly evident for word-final stop-plosives, which were much more susceptible to signal degradation in the presence of noise compared to their word-initial counterparts. Second, manner of articulation matters. The present data indicate that stop-plosives were much more susceptible to the effects of noise compared to fricatives. A potential explanation for this is a difference in how these different manner classes may be coded by the listener, with fricatives providing more acoustic leniencies regarding perceptual cues. Nevertheless, the current investigation provides valuable insight regarding important factors influencing TE intelligibility in realistic speaking environments commonly encountered by TE speakers. As such, the present findings offer additional clinical insights that are relevant to voice and speech rehabilitation outcomes following TE puncture voice restoration.

Clinical Implications

It has long been suggested that TE speech is the preferred method of postlaryngectomy alaryngeal speech. Regardless of the fact that TE voice restoration may be achieved rather rapidly in many cases, the voice and speech signal produced is not normal. As stated by Doyle, listeners will need to simultaneously deal with an abnormal, and often noisy vocal signal in addition to reductions in speech intelligibility. The present findings add new data to support the notion that although TE speech is an important and viable postlaryngectomy verbal communication method, it remains characterized by suboptimal levels of speech intelligibility even in ideal speaking environments (i.e., quiet) with ideal listeners (i.e., young normal-hearing individuals).

In light of the findings of the current study, it is clearly evident that TE speakers' present communication difficulties become further exacerbated when in the presence noisy environments. Given that it is seldom the case that verbal communication occurs in ideal speaking environments, strategies to optimize communication in noisy environments should be an integral part of counselling and the rehabilitative process. Further, the development of clinical measures which allow for the consideration of noisy environments in the assessment of functional communication outcomes might be a worthwhile endeavor. Such efforts would help to more accurately characterize and monitor TE speakers' communication outcomes. By acknowledging the realistic environment in which day-to-day communication occurs, we may better understand the communication difficulties faced by TE speakers and be better equipped to facilitate rehabilitative efforts. This suggestion certainly raises questions about how TE speech intelligibility is measured and how representative such measures may be relative to actual communication situations. While a standard method for such assessment does not currently exist, efforts within our center continue to focus on this concern. However, clinicians must be mindful of the concerns raised in the hope of interpreting intelligibility data in a fair manner.

Limitations of the Present Study

To fully contextualize the findings of the current investigation, several limitations need to be acknowledged. First, the voiced alveolar-palatal fricative /ʒ/ (as in *beige*) was not represented in the stimuli of the current study. As a result of the lack of representation within Weiss and Basili's (1985) original word list, /ʒ/ was not able to be represented within the current investigation. However, given that /ʒ/ has a relatively rare occurrence in the English language and may often be closed by surrounding sounds, we do not anticipate our intelligibility measures to deviate significantly from what might be expected of TE speakers' everyday conversational speech.

Second, the use of single word stimuli which lack the prosodic and contextual cues present in everyday conversational speech may have led to an underestimation of functional intelligibility. That is, intonation, tone, stress, and rhythm, as well as semantic context provided by surrounding words and sentence structure provide meaningful perceptual cues to assist the listener in the acquisition of a speaker's intended message.

Without these cues, the accurate portrayal of a message becomes significantly more difficult, thus, reducing intelligibility. However, the aforementioned cues were omitted by design in an effort to limit the investigation of the effect of noise to phonetic productions rather than on prosodic and contextual cues.

Lastly, given that the current findings are based on auditory-perceptual data, conclusions and rationales for the observed phenomena are speculative at best. Most importantly, the hearing capacity of those who make such judgments must be considered. Given that the majority of alaryngeal speakers may be older, it is possible that their primary conversational partners would be likely to exhibit hearing loss. Thus, multiple factors related to both the speaker and the listener must be considered in future research endeavors.

Directions for Future Research

The current study evaluated intelligibility of TE speakers through the use of normal-hearing, young adult listeners. As noted earlier, given that a significant proportion of TE communication occurs through interactions with peers (i.e., older adults) who may not have the same hearing capabilities as young adults, future studies which use older adults as listeners may provide valuable information. Additionally, future studies which investigate how noise affects other factors relating to intelligibility (e.g., vowel context, prosodic cues, sentence-level contextual cues, etc.) may provide a more comprehensive understanding of the communication difficulties face by TE speakers. Lastly, as noted previously, the current investigation was based on auditory-perceptual data, limiting explanations of observed phenomena to speculations. Future studies which employ spectral and acoustic methods of analysis may provide greater insight into the observed patterns and phenomena.

Conclusions

The current study provides additional data to support the notion that although TE speech is a viable postlaryngectomy communication option, these speakers continue to experience communication difficulties even in ideal speaking environments. Moreover, our findings illustrate how intelligibility deficits are exacerbated when TE speakers communicate in more realistic, less than ideal speaking environments. Overall, TE intelligibility was significantly lower in noise. More detailed examination of intelligibility showed noise to have differential effects according to manner of articulation and position. Nevertheless, overall error patterns remained consistent even in the presence of noise. However, differential effects on voicing feature distinction were noted according to manner and position.

In light of the current findings, it is imperative to take into consideration these data if we are to depict more accurately the day-to-day communication difficulties faced by TE speakers. Additionally, these data provide evidence that investigations conducted under ideal speaking conditions are not sufficient enough to accurately describe TE speakers' realistic functional communication outcomes. Finally, it is critical to point out that the nature of the TE speech signal is not normal; consequently, the quality of the speech and voice signal must also be considered in the context of the speech produced. This suggests that a listener may be distracted to some extent by the perceptual characteristics of the voice signal (i.e., aperiodic and noisy) which may influence one's attention to the speech signal proper. This then becomes an additional factor relative to speaker variability and the environment within which communication takes place. By acknowledging the less-than-ideal speaking environment TE speakers face every day, we may be better equipped to guide voice rehabilitation in a manner which leads to meaningful functional improvements in verbal communication.

References

Angel, D., Doyle, P. C., & Fung, K. (2011). Measuring voice outcomes following treatment for laryngeal cancer. *Expert Review of Pharmacoeconomics & Outcomes Research*, *11*(4), 415–420. https://doi.org/10.1586/erp.11.37

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

Blom, E. D., Singer, M. I., & Hamaker, R. C. (1986). A prospective study of tracheoesophageal speech. *Archives of Otolaryngology--Head & Neck Surgery*, *112*(4), 440–447.

Bridges, A. (1991). Acceptability ratings and intelligibility scores of alaryngeal speakers by three listener groups. *The British Journal of Disorders of Communication*, *26*(3), 325–335.

Cabanillas, R., Rodrigo, J. P., Llorente, J. L., & Suarez, C. (2008). Oncologic outcomes of transoral laser surgery of supraglottic carcinoma compared with a transcervical approach. *Head & Neck*, *30*(6), 750–755. https://doi.org/10.1002/hed.20778

Canadian Cancer Sociey. (2017). Canadian Cancer Statistics Special topic : Pancreatic Cancers. *Canadian Cancer Society*, 1–132. https://doi.org/0835-2976

Castellsagué, X., Alemany, L., Quer, M., Halec, G., Quirós, B., Tous, S., … Xavier Bosch, F. (2016). HPV Involvement in Head and Neck Cancers: Comprehensive Assessment of Biomarkers in 3680 Patients. *Journal of the National Cancer Institute*, *108*(6), 1–12. https://doi.org/10.1093/jnci/djv403

Clark, J. G. (1985). Alaryngeal speech intelligibility and the older listener. *The Journal of Speech and Hearing Disorders*, *50*(1), 60–5.

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

Cohen, S. M., Garrett, C. G., Dupont, W. D., Ossoff, R. H., & Courey, M. S. (2006). Voice-related quality of life in T1 glottic cancer: irradiation versus endoscopic excision. *The Annals of Otology, Rhinology, and Laryngology*, *115*(8), 581–586. https://doi.org/10.1177/000348940611500803

Danhauer, J. L., Doyle, P. C., & Lucks, L. (1985). Effects of noise on NST and NU 6 stimuli. *Ear and Hearing*, *6*(5), 266–9.

de Oliveira, D. E., Bacchi, M. M., Macarenco, R. S. S., Tagliarini, J. V., Cordeiro, R. C., & Bacchi, C. E. (2006). Human papillomavirus and Epstein-Barr virus infection, p53 expression, and cellular proliferation in laryngeal carcinoma. *American Journal of Clinical Pathology*, *126*(2), 284–293. https://doi.org/10.1309/UU2J-ADUE-HDWA-TVM9

Doyle, P. C. (1994). *Foundations of voice and speech rehabilitation following laryngeal cancer* (1st ed.). San Diego: Singular Publishing Group.

Doyle, P. C., Danhauer, J. L., & Reed, C. G. (1988). Listeners' perceptions of consonants produced by esophageal and tracheoesophageal talkers. *The Journal of Speech and Hearing Disorders*, *53*(4), 400–7.

Doyle, P. C., & Finchem, E. (n.d.). Teaching esophageal speech: A process of collaborative instruction. In P. C. Doyle (Ed.), *Clinical Care and Rehabilitation in Head and Neck Cancer*. New York: Springer Nature.

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), 350–3.

Doyle, P. C., Swift, E., & Haaf, R. G. (1989). Effects of listener sophistication on judgements of tracheoesophageal talker intelligibility. *Journal of Communication Disorders*, *22*(2), 105–113.

Dudley, B. L. (1984). *The comparative intelligibility of tracheoesophageal esophageal and laryngeal speech*. (Doctoral Dissertation). Proquest Dissertations & Theses.

(303304663).

Eadie, T. L., & Doyle, P. C. (2004). Auditory-perceptual scaling and quality of life in tracheoesophageal speakers. *The Laryngoscope*, *114*(4), 753–9. https://doi.org/10.1097/00005537-200404000-00030

Eadie, T. L., & Doyle, P. C. (2005). Quality of life in male tracheoesophageal (TE) speakers. *Journal of Rehabilitation Research and Development*, *42*(1), 115–24.

Evans, E., Carding, P., & Drinnan, M. (2009). The Voice Handicap Index with post-laryngectomy male voices. *International Journal of Language & Communication Disorders*, *44*(5), 575–586. https://doi.org/10.1080/13682820902928729

Fairbanks, G. (1960). *Voice and articulation drillbook* (2nd ed.). New Tork: Joanna Cotler Books.

Ferster, A. P. O., Schubart, J., Kim, Y., & Goldenberg, D. (2017). Association Between Laryngeal Cancer and Asbestos Exposure: A Systematic Review. *JAMA Otolaryngology-- Head & Neck Surgery*, *143*(4), 409–416. https://doi.org/10.1001/jamaoto.2016.3421

Forastiere, A. A., Goepfert, H., Maor, M., Pajak, T. F., Weber, R., Morrison, W., … Cooper, J. (2003). Concurrent chemotherapy and radiotherapy for organ preservation in advanced laryngeal cancer. *The New England Journal of Medicine*, *349*(22), 2091–2098. https://doi.org/10.1056/NEJMoa031317

Gillison, M. L., Koch, W. M., Capone, R. B., Spafford, M., Westra, W. H., Wu, L., … Sidransky, D. (2000). Evidence for a causal association between human papillomavirus and a subset of head and neck cancers. *Journal of the National Cancer Institute*, *92*(9), 709–720.

Grant, D. G., Salassa, J. R., Hinni, M. L., Pearson, B. W., Hayden, R. E., & Perry, W. C. (2007). Transoral laser microsurgery for carcinoma of the supraglottic larynx. *Otolaryngology--Head and Neck Surgery : Official Journal of American Academy of Otolaryngology-Head and Neck Surgery*, *136*(6), 900–906.

https://doi.org/10.1016/j.otohns.2006.12.015

Hashibe, M., Brennan, P., Chuang, S.-C., Boccia, S., Castellsague, X., Chen, C., … Boffetta, P. (2009). Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the International Head and Neck Cancer Epidemiology Consortium. *Cancer Epidemiology, Biomarkers & Prevention : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *18*(2), 541–550. https://doi.org/10.1158/1055-9965.EPI-08-0347

Higgins, K. M., Shah, M. D., Ogaick, M. J., & Enepekides, D. (2009). Treatment of early-stage glottic cancer: meta-analysis comparison of laser excision versus radiotherapy. *Journal of Otolaryngology - Head & Neck Surgery = Le Journal d'oto-Rhino-Laryngologie et de Chirurgie Cervico-Faciale*, *38*(6), 603–612.

Hilgers, F., & van den Brekel, M. (2010). Vocal and speech rehabilitation following laryngectomy. In P. Flint, B. Haughey, & V. Lund (Eds.), *Cummings Otolaryngology Head and Surgery* (5th ed., pp. 1594–1610). PA, USA: Mosby Elsevier.

Hillman, R. E., Walsh, M. J., & Heaton, J. T. (2005). Laryngectomy speech rehabilitation - A review of outcomes. In P. C. Doyle & R. . Keith (Eds.), *Contemporary Considerations in the Treatment and Rehabilitation of Head and Neck Cancer: Voice, Speech and Swallowing* (1st ed., pp. 75–90). Pro Ed.

Hristov, B., & Bajaj, G. K. (2008). Radiotherapeutic management of laryngeal carcinoma. *Otolaryngologic Clinics of North America*, *41*(4), 715–40, vi. https://doi.org/10.1016/j.otc.2008.01.017

IARC. (2012). Personal habits and indoor combustions. Volume 100 E. A review of human carcinogens. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans*, *100*(Pt E), 1–538.

Islami, F., Tramacere, I., Rota, M., Bagnardi, V., Fedirko, V., Scotti, L., … La Vecchia, C. (2010). Alcohol drinking and laryngeal cancer: Overall and dose-risk relation - A systematic review and meta-analysis. *Oral Oncology*, *46*(11), 802–810. https://doi.org/10.1016/j.oraloncology.2010.07.015

Jacobi, I., van der Molen, L., Huiskens, H., van Rossum, M. A., & Hilgers, F. J. M. (2010). Voice and speech outcomes of chemoradiation for advanced head and neck cancer: a systematic review. *European Archives of Oto-Rhino-Laryngology : Official Journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : Affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery*, *267*(10), 1495–1505. https://doi.org/10.1007/s00405-010-1316-x

Kayhan, F. T., Kaya, K. H., Altintas, A., & Sayin, I. (2014). Transoral robotic supraglottic partial laryngectomy. *The Journal of Craniofacial Surgery*, *25*(4), 1422–1426. https://doi.org/10.1097/SCS.0000000000000572

Kazi, R., Venkitaraman, R., Johnson, C., Prasad, V., Clarke, P., Newbold, K., … Harrington, K. (2008). Prospective, longitudinal electroglottographic study of voice recovery following accelerated hypofractionated radiotherapy for T1/T2 larynx cancer. *Radiotherapy and Oncology : Journal of the European Society for Therapeutic Radiology and Oncology*, *87*(2), 230–236. https://doi.org/10.1016/j.radonc.2008.01.006

Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility testing in dysarthria. *The Journal of Speech and Hearing Disorders*, *54*(4), 482–499.

Kreimer, A. R., Clifford, G. M., Boyle, P., & Franceschi, S. (2005). Human papillomavirus types in head and neck squamous cell carcinomas worldwide: a systematic review. *Cancer Epidemiology, Biomarkers & Prevention : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *14*(2), 467–475. https://doi.org/10.1158/1055-9965.EPI-04-0551

Li, Q., & Liu, M. (2014). Glutathione S-transferase T1 null genotype and laryngeal cancer risk: a meta-analysis. *Tumour Biology : The Journal of the International*

Society for Oncodevelopmental Biology and Medicine, *35*(9), 8781–8785. https://doi.org/10.1007/s13277-014-2121-8

MacNeil, S. D., Liu, K., Shariff, S. Z., Thind, A., Winquist, E., Yoo, J., … Garg, A. X. (2015). Secular trends in the survival of patients with laryngeal carcinoma, 1995-2007. *Current Oncology (Toronto, Ont.)*, *22*(2), e85-99. https://doi.org/10.3747/co.22.2361

McColl, D., Fucci, D., Petrosino, L., Martin, D. E., & McCaffrey, P. (1998). Listener ratings of the intelligibility of tracheoesophageal speech in noise. *Journal of Communication Disorders*, *31*(4), 279–89.

Mendelsohn, A. H., Remacle, M., Van Der Vorst, S., Bachy, V., & Lawson, G. (2013). Outcomes following transoral robotic surgery: supraglottic laryngectomy. *The Laryngoscope*, *123*(1), 208–214. https://doi.org/10.1002/lary.23621

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*(11), 1977–81. https://doi.org/10.1097/01.mlg.0000147932.36885.9e

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

Moukarbel, R. V., Doyle, P. C., Yoo, J. H., Franklin, J. H., Day, A. M. B., & Fung, K. (2011). Voice-related quality of life (V-RQOL) outcomes in laryngectomees. *Head & Neck*, *33*(1), 31–36. https://doi.org/10.1002/hed.21409

Muscat, J. E., & Wynder, E. L. (1992). Tobacco, alcohol, asbestos, and occupational risk factors for laryngeal cancer. *Cancer*, *69*(9), 2244–2251.

Ozer, E., Alvarez, B., Kakarala, K., Durmus, K., Teknos, T. N., & Carrau, R. L. (2013). Clinical outcomes of transoral robotic supraglottic laryngectomy. *Head & Neck*, *35*(8), 1158–1161. https://doi.org/10.1002/hed.23101

Pindzola, R. H., & Cain, B. H. (1988). Acceptability ratings of tracheoesophageal speech. *The Laryngoscope*, *98*(4), 394–397. https://doi.org/10.1288/00005537-198804000-00007

Qi, Y., & Zhou, X. (2014). Haplotype analysis of RECQL5 gene and laryngeal cancer. *Tumour Biology : The Journal of the International Society for Oncodevelopmental Biology and Medicine*, *35*(3), 2669–2673. https://doi.org/10.1007/s13277-013-1351-5

Rassekh, C., & Haughey, B. (2010). Total laryngectomy and laryngopharyngectomy. In P. Flint, B. Haughey, & L. V (Eds.), *Cummings Otolaryngology Head and Neck Surgery* (5th ed., pp. 1563–1576). PA, USA: Mosby Elsevier.

Robbins, J. (1984). Acoustic differentiation of laryngeal, esophageal, and tracheoesophageal speech. *Journal of Speech and Hearing Research*, *27*(4), 577–585.

Robbins, J., Fisher, H. B., Blom, E. C., & Singer, M. I. (1984). A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production. *The Journal of Speech and Hearing Disorders*, *49*(2), 202–10. https://doi.org/10.1044/jshd.4902.202

Robbins, J., Fisher, H. B., Blom, E. D., & Singer, M. I. (1984). Selected acoustic features of tracheoesophageal, esophageal, and laryngeal speech. *Archives of Otolaryngology (Chicago, Ill. : 1960)*, *110*(10), 670–2.

Rosier, J. F., Gregoire, V., Counoy, H., Octave-Prignot, M., Rombaut, P., Scalliet, P., … Hamoir, M. (1998). Comparison of external radiotherapy, laser microsurgery and partial laryngectomy for the treatment of T1N0M0 glottic carcinomas: a retrospective evaluation. *Radiotherapy and Oncology : Journal of the European Society for Therapeutic Radiology and Oncology*, *48*(2), 175–183.

Rota, S., Fidan, I., Muderris, T., Yesilyurt, E., & Lale, Z. (2010). Cytokine levels in patients with Epstein-Barr virus associated laryngeal carcinoma. *The Journal of Laryngology and Otology*, *124*(9), 990–994.

https://doi.org/10.1017/S0022215110001416

Schiavetti, N. (1992). Scaling procedures for the measurement of speech intelligibility. In R. . Kent (Ed.), *Intelligibility in Speech Disorders* (pp. 11–34). Philadelphia: Jon Benjamins Publishing Company.

Searl, J. P., Carpenter, M. A., & Banta, C. L. (2001). Intelligibility of stops and fricatives in tracheoesophageal speech. *Journal of Communication Disorders, 34*(4), 305–21. https://doi.org/10.1016/S0021-9924(01)00052-1

Singer, M. I., & Blom, E. D. (1980). An endoscopic technique for restoration of voice after laryngectomy. *The Annals of Otology, Rhinology, and Laryngology*, *89*(6 Pt 1), 529–533. https://doi.org/10.1177/000348948008900608

Sleeth, L. E. (2012). *Exploring Intelligibility in Tracheoesophageal Speech : A Descriptive Analysis*. (Masters Thesis). Electronic Thesis and Dissertations Repository. (623).

Smith, L., Bryan, S., De, P., Rahal, R., Shaw, A., Turner, D., … Dixon, M. (2018). Members of the Canadian Cancer Statistics Advisory Committee Project management.

Smith, L. F., & Calhoun, K. H. (1994). Intelligibility of tracheoesophageal speech among naive listeners. *Southern Medical Journal*, *87*(3), 333–5.

Speaks, C., & Jerger, J. (1965). Method for Measurement of Speech Identification. *Journal of Speech Language and Hearing Research*, *8*(2), 185. https://doi.org/10.1044/jshr.0802.185

Starska, K., Krzeslak, A., Forma, E., Olszewski, J., Lewy-Trenda, I., Osuch-Wojcikiewicz, E., & Brys, M. (2014). Genetic polymorphism of metallothionein 2A and risk of laryngeal cancer in a Polish population. *Medical Oncology (Northwood, London, England), 31*(7), 75. https://doi.org/10.1007/s12032-014-0075-8

Stenson, K. M., MacCracken, E., Kunnavakkam, R., W. Cohen, E. E., Portugal, L. D., Villaflor, V., … Vokes, E. E. (2012). Chemoradiation for patients with large-volume laryngeal cancers. *Head & Neck*, *34*(8), 1162–1167. https://doi.org/10.1002/hed.21888

Stevenson, M. M. (2018). Oral Cavity and Laryngeal Cancer Staging. Retrieved from https://emedicine.medscape.com/article/2048034-overview

Talamini, R., Bosetti, C., La Vecchia, C., 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 : CCC*, *13*(10), 957–964.

Tardy-Mitzell, S., Andrews, M. L., & Bowman, S. A. (1985). Acceptability and intelligibility of tracheoesophageal speech. *Archives of Otolaryngology (Chicago, Ill. : 1960)*, *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 & Neck Surgery*, *130*(4), 401–8. https://doi.org/10.1001/archotol.130.4.401

Trivedi, N. P., Swaminathan, D. K., Thankappan, K., Chatni, S., Kuriakose, M. A., & Iyer, S. (2008). Comparison of quality of life in advanced laryngeal cancer patients after concurrent chemoradiotherapy vs total laryngectomy. *Otolaryngology--Head and Neck Surgery : Official Journal of American Academy of Otolaryngology-Head and Neck Surgery*, *139*(5), 702–707. https://doi.org/10.1016/j.otohns.2008.06.002

Wei, W., Shi, Q., Guo, F., Zhang, B.-Y., Chen, C., Zhang, N.-S., & Dong, X.-P. (2012). The distribution of human papillomavirus in tissues from patients with head and neck squamous cell carcinoma. *Oncology Reports*, *28*(5), 1750–1756. https://doi.org/10.3892/or.2012.1990

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

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

Williams, S. E., & Watson, J. B. (1985). Differences in speaking proficiencies in three laryngectomee groups. *Archives of Otolaryngology (Chicago, Ill. : 1960)*, *111*(4), 216–219.

Williams, S. E., & Watson, J. B. (1987). Speaking proficiency variations according to method of alaryngeal voicing. *The Laryngoscope*, *97*(6), 737–9.

Wynder, E. L., Mushinski, M. H., & Spivak, J. C. (1977). Tobacco and alcohol consumption in relation to the development of multiple primary cancers. *Cancer*, *40*(4 Suppl), 1872–1878.

Wynder, E. L., & Stellman, S. D. (1977). Comparative epidemiology of tobacco-related cancers. *Cancer Research*, *37*(12), 4608–4622.

Xiao, C., Hanlon, A., Zhang, Q., Ang, K., Rosenthal, D. I., Nguyen-Tan, P. F., … Bruner, D. W. (2013). Symptom clusters in patients with head and neck cancer receiving concurrent chemoradiotherapy. *Oral Oncology*, *49*(4), 360–366. https://doi.org/10.1016/j.oraloncology.2012.10.004

Zhang, D., Zhou, J., Chen, B., Zhou, L., & Tao, L. (2014). Gastroesophageal reflux and carcinoma of larynx or pharynx: a meta-analysis. *Acta Oto-Laryngologica*, *134*(10), 982–989. https://doi.org/10.3109/00016489.2014.927592

Zuo, J. J., Tao, Z. Z., Chen, C., Hu, Z. W., Xu, Y. X., Zheng, A. Y., & Guo, Y. (2017). Characteristics of cigarette smoking without alcohol consumption and laryngeal cancer: overall and time-risk relation. A meta-analysis of observational studies. *European Archives of Oto-Rhino-Laryngology*, *274*(3), 1617–1631. https://doi.org/10.1007/s00405-016-4390-x

zur Hausen, H. (2000). Papillomaviruses causing cancer: evasion from host-cell control in early events in carcinogenesis. *Journal of the National Cancer Institute*, *92*(9), 690–698.

Appendices

Appendix A: Verbal Recruitment Script

Hello, my name is Sebastiano Failla and I am a graduate student in Rehabilitation Sciences. I would like to invite you to participate in a research study that will investigate intelligibility of speakers who have had their voice box removed and are required to use an alternative method of voice and speech. This project is being conducted under the supervision of Dr. Kevin Fung and Dr. Philip Doyle.

I am currently recruiting participants between 18 and 30 years of age and whose native language is English. We also ask that you provide a self-report that your hearing is normal and that you do not have any past history of hearing loss. Briefly, this study will ask you to listen to a series of voice samples. You will be asked to identify the word spoken in each sample. The listening procedure will require that you participate in one listening session that will last approximately 60 minutes. Participation in this study will be conducted within a quiet listening laboratory located in Elborn College.

Participation in this study is completely voluntary. Your identity will be maintained as confidential. Your decision of whether you wish to participate or not will not affect your course evaluation or grades in any way.

If you are interested in participating or have any questions, please email me at [redacted]

Thank you for your time and consideration.

Appendix B: Letter of Information and Consent

Letter of Information

Project Title: "An analysis of tracheoesophageal speech intelligibility in the presence of noise"

Investigators: Sebastiano Failla, BHSc, Philip C. Doyle, PhD, Kevin Fung, MD, FRCS, FACS

Introduction and Purpose of Study

The purpose of this letter is to provide you with the information you require to make an informed decision regarding participation in this research. This study will examine speaker intelligibility, the percentage of speech items correctly identified by a listener. Word Identifications will be made on samples provided by both men and women who use tracheoesophageal speech as their primary mode of verbal communication.

Activities You Will Take Part In

The listening procedure will require you to spell out, in plain English, words you heard after listening to a series of speaker samples; further instructions will be provided. It will require that you participate in one listening session that will last approximately 60 minutes. The session will be held in Elborn College (Room 2200), University of Western Ontario.

Inclusion Criteria

If you are between the ages of 18 and 30 years, self-report hearing to be within normal limits, have no formal training in voice or voice disorders, and English is your native language, you are welcome to participate.

Exclusion Criteria

If you fall outside of the age range, have any past history of hearing impairment or hearing loss, have any level of formal training, or English is not your native language, you will be excluded from participation.

Any Possible Risks or Discomforts

Fatigue associated with the repetitive nature of the task.

Any Possible Benefits

Due to the nature of this study, you will not directly benefit from the data obtained. Potential benefits to society include a greater understanding of communication difficulties of tracheoesophageal speakers which may help guide future alaryngeal voice rehabilitation.

<u>Voluntary Participation</u>

Your participation is voluntary and you may refuse to answer any question without penalty or academic consequence. You may withdraw from the study at any point without penalty or academic consequence. You may withdraw your data from study inclusion at any point, including after participation is completed (you have left the lab). Data withdrawal requests can be made by contacting any one of the investigators (contacts listed below). Withdrawal of your data will result in its exclusion from the study.

<u>Compensation</u>

You will not receive compensation for your participation.

<u>Confidentiality</u>

All data obtained will remain confidential. The investigators will keep any personal information about you in a secure and confidential location for a minimum of 5 years. Your data will be identified by a code known only to the investigators. A master list linking your code with your name will be kept separate from your study file in a locked cabinet in the Voice Production & Perception Laboratory at The University of Western Ontario. Only the investigators will have access to the locked cabinet. As such, confidentiality will be maintained as much as possible. Representatives of The University of Western Ontario Non-Medical Research Ethics Board may require access to your study-related records to monitor the conduct of the research.

<u>Waiver of Rights</u>

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

<u>Contacts for Further Questions</u>

If you have questions or require more information about the study itself, please contact Dr. Philip Doyle by e-mail at [redacted] or by telephone at [redacted]. If you have concerns or questions about your rights as a participant or about the way the study is conducted, you may contact:

University of Western Ontario Health Sciences Research Ethics Board
c/o Office of Research Services: Room 5150 Support Services Building,
1393 Western Road,
London, Ontario, Telephone: [redacted]
Canada, N6G 1G9 E-mail: [redacted]

Letter of Consent

Project Title: "An analysis of tracheoesophageal speech intelligibility in the presence of noise"
Investigators: Sebastiano Failla, BHSc, Philip C. Doyle, PhD, Kevin Fung, MD, FRCS, FACS

Consent: I have read the Letter of Information, 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 Printed Name ____

Participant's Signature ________________________________ Date:_______________

Person Obtaining Informed Consent

Printed Name ____

Signature ____ Date:________________

Appendix C: Word List

1. CANE
2. MASS
3. TEETHE
4. DOPE
5. SACK
6. BAD*****
7. LEAF
8. SHAVE
9. ZAG
10. GAB
11. PATH*****
12. VET
13. THESE*****
14. FISH
15. THEME

***** *Predetermined reliability stimulus.*

Appendix D: Stimulus Normalization Protocol

Step 1. Normalize all words to the minimum root mean square (RMS) vowel amplitude across all words.

$I_v + A_1 = I_{min}$

$A_1 = I_{min} - I_v$

Where

A_1 = amplification factor for individual word (dB)

I_v= RMS vowel amplitude of individual word (dBFS)

I_{min} = minimum RMS vowel amplitude across all stimuli (dBFS)

Step 2. Determine the maximum word amplitude (I_{max}) across all words after applying first amplification factor.

Step 3. Amplify all stimuli by the same factor (A_2) determined by the following formula.

$I_{max} + A = -2$

$A_2 = -(I_{max} + 2)$

Where

I_{max} = global maximum word amplitude after normalization (dBFS)

A_2 = second amplification factor applied to all words (dB)

Appendix E: Randomization Script

```
REM This script allows you to randomize a set of files within a folder of your choice
REM It provides the translation in a txt files
REM Files can also be de-randomized by running the script again and choosing option 2
REM Sebastiano Failla
REM The University of Western Ontario, February, 2018
REM sfailla@uwo.ca

@ECHO OFF

SETLOCAL EnableExtensions EnableDelayedExpansion

:START
cls
COLOR 0A
TITLE File Randomizer

SET TF1=Translation1.txt
SET TF2=Translation2.txt
ECHO.
ECHO.
ECHO                               FILE RANDOMIZER
ECHO.
ECHO ____________________________________________________________________________________________
ECHO.
ECHO Choose one of the following [1/2/3]:
ECHO.
ECHO 1:RANDOMIZE FILE NAMES
ECHO 2:UNDO RANDOMIZATION
ECHO 3:EXIT
ECHO.
ECHO.

SET /P Choice=

IF %Choice%==1 (GOTO PROCEED)
```

```
IF %Choice%==2 (GOTO UNDO)

IF %Choice%==3 (GOTO :EOF)

IF NOT %Choice==1,2,3 (
                    ECHO INVALID ENTRY
                    pause
                    GOTO START)

:LOOP1
:PROCEED
cls

ECHO.
ECHO INDICATE FILE EXTENSION TYPE[EX; "wav"]
ECHO.
ECHO.
SET /P Ext=

ECHO.
ECHO INSERT LIST NAME[LETTER]
ECHO.
ECHO.
SET /P List=

ECHO.
ECHO NUMBER TO START COUNT AT
ECHO.
ECHO.
SET /P Number1=

ECHO.
ECHO.
ECHO CONFIRM RANDOMIZATION:[y/n]
ECHO.
ECHO.
SET /P Confirm=

IF %Confirm%==y (GOTO RANDOMIZE)
IF %Confirm%==n (GOTO START)
IF NOT %Confirm%==y,n (
```

```
                    ECHO INVALID ENTRY
                      pause
                      GOTO LOOP1


:randomize

ECHO Randomized/Original>%TF1%
ECHO ----------------------------------------------------------------------------------------------------
----------------------------->>%TF1%

FOR /F "tokens=1 delims=~" %%A IN ('DIR /A:-D /B') DO (
        IF NOT %%A==%~nx0 (
                IF NOT %%A==%TF1% (
                        SET Use=%%A
                        SET Modified=!RANDOM!------!Use!

                        RENAME "%%A" "!Modified!"

                        ECHO !Modified!/%%A>>%TF1%
                )
        )
)

ECHO Randomized/Original>%TF2%
ECHO ----------------------------------------------------------------------------------------------------
----------------------------->>%TF2%

REM Creates array with files in order
for %%A in (*.!Ext!) do (
  for /F "delims=-" %%n in ("%%A") do (
    set "number=00000%%n"
    set "file[!number:~-6!]=%%A"
   )
)

REM Process the filenames in right order

SET /A Count =%Number1%-1

for /F "tokens=2 delims==" %%A in ('set file[') DO (
```

```
        IF NOT %%A==%~nx0 (
                IF NOT %%A==%TF1% (
                    IF NOT %%A==%TF2% (
                        SET Use=%%A
                        SET /A Count +=1
                        SET Modified=!List!!Count!.!Ext!

                        RENAME "%%A" "!Modified!"

                        ECHO !Modified!    /%%A>>%TF2%
                    )
                )
        )
)

FOR /F "tokens=1 delims=~" %%A IN ('DIR /A:-D /B')
GOTO :EOF

:UNDO
:LOOP2
cls
ECHO.
ECHO CONFIRM UNDO:[y/n]
ECHO.
ECHO.
SET /P confirmchoice=

IF %confirmchoice%==y (GOTO FinishUndo)
IF %confirmchoice%==n (GOTO START)
IF NOT %confirmchoice%==y,n (
                    ECHO INVALID ENTRY
                    GOTO LOOP2
                )

:FinishUndo

IF NOT EXIST %TF2% (
            cls
            ECHO.
            ECHO.
            ECHO Translation reference unavailable
```

```
            ECHO Press any key to return to main menu
            ECHO.
            ECHO.
            pause>nul
            GOTO START
            )

FOR /F "skip=2 tokens=1,2 delims=/" %%A IN (%TF2%) DO (RENAME "%%A"
"%%B")

DEL /F /Q %TF2%

IF NOT EXIST %TF1% (
            cls
            ECHO.
            ECHO.
            ECHO Translation reference unavailable
            ECHO Press any key to return to main menu
            ECHO.
            ECHO.
            pause>nul
            GOTO START
            )
FOR /F "skip=2 tokens=1,2 delims=/" %%A IN (%TF1%) DO (RENAME "%%A"
"%%B")
DEL /F /Q %TF1%
```

Curriculum Vitae

Name:	Sebastiano Failla
Post-secondary Education and Degrees:	The University of Western Ontario London, Ontario, Canada 2019 M.Sc.
	The University of Western Ontario London, Ontario, Canada 2015 B.H.Sc.
Honours and Awards:	Western Graduate Research Scholarship 2016-2017, 2017-2018
	Western University Founder's Award 2014
Related Work Experience	Clinical Research Assistant Lawson Research Institute – Parkwood Hospital 2018-2019
	Graduate Teaching Assistant The University of Western Ontario 2017
	Research Assistant Voice Production and Perception Laboratory The University of Western Ontario 2014-2017

Published Abstracts

You, P., **Failla, S.,** Rajakumar, C., Dworschak-Stokan, Doyle, P.C., A., & Husein, M. (2019). Characteristics of velopharyngeal dysfunction in 22q11.2 deletion syndrome, annual American Society of Pediatric Otolaryngology (ASPO) Meeting, Austin, TX.

Caughlin. S., Longval, M., **Failla, S.**, Mirkowski, M., Mehta, S., McIntyre, A., Sequeira, K., Loh, E., & Teasell, R. (2019). Psychosocial factors and long-term outcomes in mild traumatic brain injury (mTBI): The role of anxiety sensitivity, 8th annual GTA Rehab Network Best Practices Day, Toronto, ON.

Caughlin. S., Longval, M., **Failla, S.**, Mirkowski, M., Mehta, S., McIntyre, A., Sequeira, K., Loh, E., & Teasell, R. (2019). Psychosocial factors and long-term outcomes in mild traumatic brain injury (mTBI): The role of experiential avoidance, 8th annual GTA Rehab Network Best Practices Day, Toronto, ON.

Failla, S., & Doyle, P.C. (2018, November). The influence of multi-talker noise on stop-plosive and fricative intelligibility of tracheoesophageal speakers, annual American Speech-Language-Hearing Association (ASHA) Convention, Boston, MA.

Failla, S., Al Zanoon, N., Smith, N., & Doyle, P.C. (2017, November). Effects of contextual priming on listener judgments of alaryngeal speech, annual American Speech-Language-Hearing Association (ASHA) Convention, Los Angeles, CA.

Cox, S.R., **Failla, S.**, & Doyle, P.C. (2015, November). The impact of clear speech on listener judgements of electrolaryngeal speech, annual American Speech-Language-Hearing Association (ASHA) Convention, Denver, CO.

Reitzel, K., Cox, S.R., **Failla, S.**, & Doyle, P.C. (2015, November). Temporal modifications of electrolaryngeal speech during a clear speech task, annual American Speech-Language-Hearing Association (ASHA) Convention, Denver, CO.

Smith, N., Cox, S.R., Day, A.M.B., **Failla, S.**, & Doyle, P.C. (2015, November). Situationally-bound judgements of listener comfort for postlaryngectomy voice & speech, annual American Speech-Language-Hearing Association (ASHA) Convention, Denver, CO.

Jackman, K., Cox, S.R., **Failla, S.**, Leblanc, C., & Doyle, P.C. (2014, November). Auditory-perceptual assessment of acceptability, listener comfort, & voice-related quality of life in female tracheoesophageal speakers, annual American Speech-Language-Hearing Association (ASHA) Convention, Orlando, FL.

Wilkinson, J.L., Cox, S.R., **Failla, S.**, Leblanc, C., & Doyle, P.C. (2014, November). Determining the presence of a gender bias in auditory-perceptual evaluations of tracheoesophageal speakers, annual American Speech-Language-Hearing Association (ASHA) Convention, Orlando, FL.