

Electronic Thesis and Dissertation Repository

3-10-2016 12:00 AM

The Application of Clear Speech in Electrolaryngeal Speakers

Steven R. Cox

The University of Western Ontario

Supervisor

Dr. Philip C. Doyle

The University of Western Ontario

Graduate Program in Health and Rehabilitation Sciences

A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of
Philosophy

© Steven R. Cox 2016

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Speech Pathology and Audiology Commons](#)

Recommended Citation

Cox, Steven R., "The Application of Clear Speech in Electrolaryngeal Speakers" (2016). *Electronic Thesis and Dissertation Repository*. 3533.

<https://ir.lib.uwo.ca/etd/3533>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlsadmin@uwo.ca.

Abstract

The present work was comprised of a series of experiments that investigated the application of clear speech (CS) in a group of electrolaryngeal (EL) speakers. Three experiments were conducted to assess the impact of CS on three important aspects of EL speech. More specifically, Experiment 1 sought to identify the impact of CS on EL speakers' word and consonant intelligibility; Experiment 2 examined the influence of CS on the acoustic characteristics of words and vowels in EL speech; and finally, Experiment 3 sought to identify the influence of CS produced by EL speakers on auditory-perceptual ratings by naïve listeners. Results revealed that overall word and consonant intelligibility were minimally different when EL speakers used CS compared to their everyday, 'habitual' speech (HS) (Experiment 1). Secondly, EL speakers' use of CS significantly increased word durations, but did not have a substantial impact on fundamental and formant frequency characteristics of vowels (Experiment 2). Finally, due to the productive changes associated with CS involving a slower rate of speech, over-articulation, and increased mouth-opening, listeners judged EL speech to be significantly less acceptable to listen to when compared to HS. However, no significant effect of speaking condition was noted on listeners' comfort levels (Experiment 3). Overall, findings suggest that the acoustic deficits in EL speech might be too complex to derive further benefit from CS in the areas of speech intelligibility, the acoustic structure of EL speech and/or auditory-perceptual ratings of EL speakers. Clinical implications and future directions for research are discussed.

Keywords

electrolarynx, clear speech, intelligibility, speech acoustics, speech acceptability, listener comfort

Co-Authorship Statement

The primary author of the work contained within this thesis was Steven Randall Cox.

The other author on Chapters 2, 3, and 4 was Philip C. Doyle, Ph.D., the candidate's supervisor.

Acknowledgements

To my mentor, Dr. Philip Doyle. There are many things that I have to be grateful for during my doctoral journey. The majority of these items are the result of your consistent and tireless support, borne out of a general desire to see all of your students succeed. Thank you for your patience, motivation, and knowledge. There is no other person in this world that could have provided the level of doctoral training quite like you, especially because of the knowledge and experience that you shared during my time in your lab. This was accentuated by your unique ability to balance all of the important things in life that contribute toward improving one's quality of life. You and Betsy also treated me like family, welcoming me into your home for wonderful celebrations to honor graduating colleagues and holiday parties filled with bountiful food and drink. Above all else, you supported me when I lost the most important person in my life. For this reason, and many more, I will always consider you more than a mentor. Thank you.

To my advisory committee, Dr. Kevin Fung and Dr. J. B. Orange. Thank you for your insight and feedback as I developed my dissertation projects. It was always a pleasure to speak with both of you, whether in the Head and Neck clinic at the London Regional Cancer Program, or in the halls of Elborn College. Your support throughout my doctoral training will always be remembered and appreciated.

To my thesis examination committee, Dr. J. B. Orange, Dr. Jeffrey Searl, Dr. Sandi Spaulding and Dr. Jeff Tennant. Thank you for your questions and feedback during my defense. The value and importance of the doctoral defense is something that I will cherish as I move forward in my academic career. I will continue to strengthen what I know and continue to further my knowledge in what may be currently unknown.

To the 'Labrats', Dr. Catherine Bornbaum, Dr. Marie-Eve Caty, Mr. Eric Davis, Dr. Adam Day, Dr. Agnes Dzioba, Mrs. Melissa Nash, Ms. Grace Scott, and Ms. Lindsey Sleeth. I was blessed to have the opportunity to work alongside such an intelligent group of people (not to mention a group that understands something about work-life balance). Thank you.

To the individuals who participated in the dissertation projects, the undergraduate and graduate students at Western University, and the laryngectomees. This

work would not have been possible without any of you. I am forever thankful for your time and willingness to participate.

To Mr. Sebastiano Failla and the lab volunteers. You are vital to the daily operations and pursuits of Dr. Doyle's lab. Never forget how thankful I am to have you as part of the team. This project would not been completed without you.

To the faculty and student members of the Rehabilitation Sciences Journal Club. Thank you for the opportunity to be part of such a supportive learning environment. I appreciate several invitations to share my own research, the feedback that came from these experiences, in addition to the opportunity to learn about other member's research interests, too.

To members of the Department of Otolaryngology – Head and Neck Surgery and the London Regional Cancer Program. Words cannot express how thankful I am to have worked alongside Canada's brightest minds. Further, I am forever grateful for the invitation to numerous department gatherings and weekly grand rounds.

To my colleagues in the Department of Communication and Sciences Disorders at Adelphi University, Dr. Robert Goldfarb and Dr. Lawrence Raphael, for their review and valuable feedback during the revisions of this doctoral thesis. In addition, Dr. Carolyn Springer in the Department of Psychology at Adelphi University for her assistance with the statistical analyses used in this doctoral dissertation.

To my immediate and extended families, especially my brother Craig and sister Melissa. Although mom left us during my doctoral training, it was because I had both of you that I was able to carry on. Always remember that I am thankful for every text, every phone call, every visit to London, and every helping hand during a move.

To my late Uncle John Pagliacci. You inspired me to begin investigating individuals' communication and quality of life after head and neck cancer treatment.

To my wife, Dr. Katia Ciampa. I cannot thank you enough for the love and support you have provided me, especially while I was writing my dissertation. We lost a very special woman during my journey. However, it was you that helped me endure, looking toward a brighter future with fond memories. You are my destiny.

Last, but not least, to my late mother, Aida Lucia (née Pilato) Macdonald. All of my life, you encouraged me to pursue my dreams. You always reminded me that anything I wanted to accomplish was possible. The completion of my doctorate is a testament to this. I move forward with the intent of making you proud as you look down from above. Grazie per “l’amor che move il sole e l’altre stelle.”

Table of Contents

	Page
Abstract	ii
Co-Authorship Statement	iii
Acknowledgements	iv
Table of Contents	vii
List of Tables	xii
List of Figures	xiv
List of Appendices	xvi
Chapter One	1
Introduction and Review of Literature	1
Medical Management of Laryngeal Cancer	2
Postlaryngectomy Voice and Speech	4
History of Electrolaryngeal Voice and Speech	7
Alaryngeal Speech and Voice-Related Quality of Life	11
Speech Intelligibility	13
Acoustic Properties of Electrolaryngeal Speech	18
Intensity and Signal-to-Noise Ratio	18
Frequency	20
Speaking Rate	24
Suprasegmental Features of Electrolaryngeal Speech	26
Intonation	26
Stress	27
Rhythm	28
Juncture	29
Perceptual Features of Electrolaryngeal Speech	31
Speech Acceptability	31
Listener Comfort	33
Social Consequences of Electrolaryngeal Speech	35
Experimental Attempts to Improve Acoustic Characteristics of Electrolaryngeal Speech	37

Clear Speech	39
Rationale for the Present Studies	45
References	48
Chapter Two	65
The Impact of Clear Speech on Word Intelligibility of Electrolaryngeal Speakers	65
Method	70
Participant Speakers	70
Speech Stimuli	71
Acquisition of Speech Stimuli	71
Editing Speech Stimuli	72
Evaluation of Intelligibility	73
Participant Listeners	73
Listening Procedure	74
Data Analyses	76
Statistical Analyses	76
Word-Level Analyses	76
Word-Position Analyses	77
Relationships Between Speech Intelligibility Scores	77
Agreement and Reliability	77
Results	78
Word Intelligibility	78
Comparison of Word Intelligibility by Device Group	80
Relationship Between Speaking Conditions	80
Intelligibility by Consonant Position: Word-Initial and Word-Final	82
Word-Initial Position	82
Word-Final Position	82
Comparison of Word Position by Device Group	86
Voicing Feature	87
Voiced Consonants	87
Word-Initial Voiced Consonants	87
Word-Final Voiced Consonants	89

Voiceless Consonants	91
Word-Initial Voiceless Consonants	91
Word-Final Voiceless Consonants	92
Comparison of Voiced and Voiceless Consonants	92
Comparison of Voiced and Voiceless Consonants by Speaking Condition	93
Comparison of Voiced and Voiceless Consonants by Word Position	94
Manner Feature	94
Nasals	95
Fricatives	95
Plosives	95
Affricates	95
Comparison of Individual Speaker Scores by Word Position	98
Omissions	100
Discussion	103
Word Intelligibility	103
Comparison of Word Intelligibility by Device Group	105
Intelligibility by Consonant Position: Word-Initial and Word-Final	106
Comparison of Word Position by Device Group	107
Voicing Feature	108
Manner Feature	111
Plosives	111
Affricates	111
Omissions	113
Plosives	113
Fricatives	113
Conclusions	114
References	117
Chapter Three	123
The Influence of Clear Speech on Acoustic Characteristics of Electrolaryngeal Speech	123
Method	130

Speakers	130
Speech Stimuli	131
Data Acquisition	132
Recording of Speech Stimuli	132
Editing Word Stimuli	133
Acoustic Analysis	133
Duration Measurement	134
Overall Word Duration	134
Overall Vowel Duration	134
Fundamental Frequency	134
Data Analyses	135
Results	136
Whole-Word Stimuli	136
Overall Word Duration	136
Vowel Stimuli	144
Overall Vowel Duration	144
Non-Speech Fundamental Frequency	146
Vowel Fundamental and Formant Frequencies	146
Discussion	164
Conclusions	170
References	171
Chapter Four	179
The Impact of Clear Speech on Auditory-Perceptual Judgments of Electrolaryngeal Speech	179
Method	182
Participant Speakers	182
Preliminary Intelligibility Assessment	183
Data Collection – Experimental Speech Samples	184
Recording of Speech Stimuli	184
Listener Stimuli	186
Participant Listeners	187

Auditory-Perceptual Rating Procedure	188
Data Analysis	190
Agreement and Reliability	191
Results	194
Listener Ratings	194
Speech Acceptability	194
Listener Comfort	194
Correlational Analysis	194
Discussion	197
Clinical and Research Implications	204
References	206
Chapter Five	213
General Discussion and Integration of Findings	213
General Overview	213
Integration of Findings	217
Auditory-Perceptual Assessment Following the Application of Clear Speech	225
Limitations of the Present Work	232
Clinical Implications	236
Directions for Future Research	241
Conclusions	245
References	247
Appendices	255
Curriculum Vitae	273

List of Tables

Table	Description	Page
2.1	Individual Speaker Raw and Percentage Scores for Overall Words	79
2.2a	Overall Perceptual Confusion Matrix for Word-Initial Consonants Spoken with Habitual Speech/Clear Speech	83
2.2b	Overall Perceptual Confusion Matrix for Word-Final Consonants Spoken with Habitual Speech/Clear Speech	84
2.3	Overall Individual Speaker SI Raw and Percentage Scores for Word-Initial and Word-Final Positions Across Habitual Speech and Clear Speech Conditions	85
2.4a	Overall Individual Speaker SI Raw and Percentage Scores for Word-Initial Phonemic Voicing Features Across Habitual Speech and Clear Speech Conditions	88
2.4b	Overall Individual Speaker SI Raw and Percentage Scores for Word-Initial Phonemic Voicing Features Across Habitual Speech and Clear Speech Conditions	90
2.5a	Individual Speaker Raw and Percentage SI Scores for Word-Initial Consonants By Manner Class Across Habitual Speech and Clear Speech Conditions	96
2.5b	Individual Speaker Raw and Percentage SI Scores for Word-Final Consonants By Manner Class Across Habitual Speech and Clear Speech Conditions	97
2.6a	Total Number of Omissions By Manner Feature for Habitual Speech and Clear Speech Conditions	101
2.6b	Individual Speaker Omissions By Manner Feature for Habitual Speech and Clear Speech Conditions	102
3.1	Overall Mean Word Durations for EL Speakers During Habitual Speech and Clear Speech	137
3.2a	Overall Mean Word Durations for Servox Speakers During Habitual Speech and Clear Speech	140

3.2b	Overall Mean Word Durations for TruTone Speakers During Habitual Speech and Clear Speech	141
3.3	Mean Mean Vowel Durations for EL Speakers During Habitual Speech and Clear Speech	136
3.4	Fundamental Frequency of Non-Speech Data by Electrolaryngeal Speakers	147
3.5	Average Fundamental and Formant Frequencies of Vowels Produced By Servox Digital and TruTone Speakers in Habitual Speech and Clear Speech	148
4.1	Individual Speaker Intelligibility Scores for Word-Initial and Word-Final Stimuli	185
4.2a	Agreement for Listener Ratings of Speech Acceptability (ACC)	192
4.2b	Agreement for Listener Ratings of Listener Comfort (LC)	193

List of Figures

Figure	Description	Page
2.1	Relationship Between Individual Speaker Intelligibility In Habitual Speech and Clear Speech	81
3.1	Overall Word Durations and Ranges by Electrolaryngeal Speakers	138
3.2a	Mean Word Durations by Servox Speakers	142
3.2b	Mean Word Durations by TruTone Speakers	143
3.3a	F1/F2 Plots of Monophthongs Produced by Servox Speakers	150
3.3b	F1/F2 Plots of Diphthongs Produced by Servox Speakers	151
3.4a	F1/F2 Plots of Monophthongs Produced by TruTone Speakers	152
3.4b	F1/F2 Plots of Diphthongs Produced by TruTone Speakers	153
3.5a	F1/F2 Plots of Monophthongs Produced by Speaker 1	154
3.5b	F1/F2 Plots of Diphthongs Produced by Speaker 1	154
3.6a	F1/F2 Plots of Monophthongs Produced by Speaker 2	155
3.6b	F1/F2 Plots of Diphthongs Produced by Speaker 2	155
3.7a	F1/F2 Plots of Monophthongs Produced by Speaker 3	156
3.7b	F1/F2 Plots of Diphthongs Produced by Speaker 3	156
3.8a	F1/F2 Plots of Monophthongs Produced by Speaker 4	157
3.8b	F1/F2 Plots of Diphthongs Produced by Speaker 4	157
3.9a	F1/F2 Plots of Monophthongs Produced by Speaker 5	158
3.9b	F1/F2 Plots of Diphthongs Produced by Speaker 5	158
3.10a	F1/F2 Plots of Monophthongs Produced by Speaker 6	159
3.10b	F1/F2 Plots of Diphthongs Produced by Speaker 6	159
3.11a	F1/F2 Plots of Monophthongs Produced by Speaker 7	160
3.11b	F1/F2 Plots of Diphthongs Produced by Speaker 7	160
3.12a	F1/F2 Plots of Monophthongs Produced by Speaker 8	161
3.12b	F1/F2 Plots of Diphthongs Produced by Speaker 8	161
3.13a	F1/F2 Plots of Monophthongs Produced by Speaker 9	162
3.13b	F1/F2 Plots of Diphthongs Produced by Speaker 9	162
3.14a	F1/F2 Plots of Monophthongs Produced by Speaker 10	163

3.14b	F1/F2 Plots of Diphthongs Produced by Speaker 10	163
4.1	Mean Listener Ratings of Acceptability for Electrolaryngeal Speakers Between Habitual Speech and Clear Speech Conditions	195
4.2	Mean Listener Ratings of Listener Comfort for Electrolaryngeal Speakers Between Habitual Speech and Clear Speech Conditions	196

List of Appendices

Figure	Description	Page
Appendix A	Word stimuli list (Chapters 2 and Chapter 3)	255
Appendix B	Ethical approval (Chapters 2, 3, and 4, Participant Speakers)	256
Appendix C	Letter of information and consent (Chapters 2, 3, and 4, Participant Speakers)	257
Appendix D	Demographic questionnaire (Chapter 2, 3, and 4, Participant Speakers)	263
Appendix E	Ethical approval (Chapters 2 and 4, Participant Listeners)	264
Appendix F	Letter of information and consent (Chapter 2, Participant Listeners)	265
Appendix G	Letter of information and consent (Chapter 4, Participant Listeners)	269

Chapter 1

Introduction and Review of Literature

The larynx is a critical structure in human functioning and survival. Due to its anatomical position at the top of the airway, the larynx is involved in respiration, protection of the airway, and is also the source of the human voice. It contains three anatomical divisions that are often described in relation to the glottis (the variable area between the true vocal folds). These regions include the supraglottis (area extending above the vocal folds) and the subglottis (the area extending below the vocal folds). A threat or violation to any of these anatomical divisions can lead to a wide-range of consequences; for example, a sudden change in voice quality or complete loss of one's voice. One threat to the larynx that has existed for thousands of years is cancer (Snidecor, 1968).

Cancer involves the uncontrolled proliferation of abnormal cells within the body (American Cancer Society, 2015). Cancer of the larynx most often arises from the squamous epithelium of the true vocal folds, but can also extend into the supra- and/or subglottic regions. Laryngeal cancer is often described by a set of staging guidelines developed by the American Joint Committee on Cancer (AJCC) and the Union for International Cancer Control (UICC) in relation to tumour size (T), involvement of lymph nodes (N), and the presence (or absence) of distant metastasis (M) (AJCC, 2010; UICC, 2009).

Recent estimates indicate that there will be a proportionally similar number of new diagnoses of laryngeal cancer in Canada and the United States with 1,050 and 13,560 cases, respectively (American Cancer Society, 2015; Canadian Cancer Society,

2015; Siegel, Miller, & Jemal, 2015). Due to medical advances, however, improvements in the early detection of laryngeal cancer have been observed (Doyle, 1994). This has resulted in individuals surviving longer after initial diagnosis and without significant differences in patient survival between treatment modalities (Department of Veterans Affairs Laryngeal Cancer Study Group, 1991; Doyle, 1994; Finizia, Hammerlid, Westin, & Lindstrom, 1998; Silver, Beitler, Shaha, Rinaldo, & Ferlito, 2009; Timmermans, de Gooijer, Hamming-Vrieze, Hilgers, & van den Brekel, 2014). The National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program estimates that approximately two-thirds of individuals with laryngeal cancer live at least five years after their diagnosis and these rates have remained stable since 1975 (SEER, 2014). As a result, it is important to consider the potential needs of this population after diagnosis and subsequent treatment. Research has indicated that an individual's needs following laryngeal cancer treatment vary greatly and are based on the treatment(s) selected (The Department of Veterans Affairs Laryngeal Cancer Study Group, 1991; Finizia et al., 1998; Hanna et al., 2004; Rinkel et al., 2014; Robertson, Yeo, Sabey, Young, & MacKenzie, 2013). Therefore, the following section will highlight three standard treatments currently offered for laryngeal cancer. In addition, the consequences of laryngeal cancer treatment on communication will be discussed with an emphasis on total laryngectomy (TL).

Medical Management of Laryngeal Cancer

Three standard treatments for laryngeal cancer include surgery, radiation therapy (RT), and concomitant chemoradiotherapy (CCRT) (National Cancer Institute, 2014; Silver et al., 2009). Surgical intervention generally involves resection of the tumour and a

margin surrounding it, and the option for removal of regional lymph nodes (i.e., neck dissection). RT employs the use of internal or external radiation (e.g., brachytherapy or intensity-modulated radiation therapy, respectively) to ameliorate malignant cells. Adjuvant RT involves treatment after surgery to remove any remaining, though undetected cancer cells. CCRT involves the use of combined RT and chemotherapy (CT), which utilizes drug therapy to shrink and prevent the division of cancer cells. CCRT has been shown to provide similar survival rates when compared to surgical intervention (TL) alone (Department of Veterans Affairs Laryngeal Cancer Study Group, 1991; Forastiere et al., 2003). While CCRT is often used as part of a ‘conservation’ approach to preserve the larynx, advanced laryngeal cancer tumours are often treated with TL in addition to RT ((Forastiere et al., 2003; Timmermans et al., 2014). Timmermans et al. (2014) indicated that the majority of patients with advanced laryngeal cancer continue to rely on surgery with RT even though no significant differences have been found in survival between CCRT and surgery. Further, Timmermans et al. (2014) reported more recurrences of cancer in individuals treated with RT or CRT alone when compared to TL (e.g., 32.4% for RT and 30% for CRT compared to 13.3% following TL). If RT or CCRT are selected as the initial treatment method, TL or modified surgical procedures might be the last option for controlling regional and/or distant disease.

Partial laryngectomy or other conservation surgical procedures can be used in an attempt to spare the function of the larynx. This is especially true for early stage laryngeal cancer or to treat disease recurrence (Bailey, 1971; Biacabe, Creiver-Buchman, Hans, Laccourreye, & Brasnu, 1999; Silver & Ferlito, 1996). For example, a partial laryngectomy may involve the removal of one true vocal fold while maintaining some

level of function of the remaining vocal fold for breathing, swallowing, and/or phonatory function. In contrast, TL remains the most radical surgical treatment for laryngeal cancer and involves the complete removal of the entire larynx and surrounding structures. In addition, the trachea is detached from the upper aerodigestive tract and is sutured to the front of the neck to create a permanent tracheostoma for breathing. It is not surprising, then, that individuals face a host of postlaryngectomy issues related to breathing and stoma care, as well as those related to voice and speech. Thus, loss of the larynx will require acquisition of a new speaking method postlaryngectomy (or, what is termed ‘alaryngeal’ voice/speech). Since several communication options are currently available for laryngectomees¹, it is important to describe how voice and speech can be produced without a larynx. Therefore, the following section will examine the communication options that exist after TL with a special focus on the electrolarynx (EL) and its importance as a postlaryngectomy communication option.

Postlaryngectomy Voice and Speech

Research in the area of postlaryngectomy communication has evolved considerably over the past 140 years. Voice and speech production following TL began with an artificial larynx developed by Lieter in 1873 and led to the introduction of the electronic, neck-type artificial larynx by Bell Laboratories in 1959 (Barney, Haworth, & Dunn, 1959). Presently, three alaryngeal speaking methods are typically offered to individuals postlaryngectomy; this includes esophageal (ES), tracheoesophageal (TE),

¹ Although not “person-first” language, Doyle (in press) has indicated that this term is preferred by those who have undergone TL.

and electrolaryngeal (EL) speech. Unfortunately, the electrolarynx (EL) has been historically viewed as an inferior alaryngeal communication option by some physicians and Speech-language pathologists (Berry, 1978; Duguay, 1978; Gates et al., 1982; Lauder, 1968). This also resulted as a consequence of comparisons between alaryngeal speech methods (Doyle & Eadie, 2005). Therefore, a description of each specific alaryngeal speaking method is necessary to provide a better understanding regarding the differences between each method.

Laryngectomees trained to use ES generate voice by injecting or insufflating air into the esophageal reservoir. This is followed by a controlled release of air that passes across reconstructed pharyngeal and esophageal anatomical tissues that comprise the pharyngoesophageal (PE) segment (Diedrich, 1968; Doyle & Eadie, 2005). The PE segment is set into vibration and the resulting sound energy travels into the oral cavity where it can be articulated into speech (Diedrich, 1968). In comparison, the production of TE speech is similar to ES in that it depends on the PE segment for voicing. However, TE speech differs in two ways: 1) a reliance on pulmonary air as the driving source (Doyle, Danhauer, & Reed, 1988), and 2) the use of a prosthesis that is placed in a surgically created puncture site in the common tissue wall that separates the trachea anteriorly and esophagus posteriorly. TE speech production begins with the introduction of air through the tracheostoma at the front of the neck, followed by occlusion of the tracheostoma with a finger or hands-free valve (Blom, Singer, & Hamaker, 1986; Singer & Blom, 1980). Closing the airway in this manner directs pulmonary air into the esophagus through the prosthesis which serves as a conduit into the PE reservoir. TE “voicing” is created as air pressure increases in the esophagus and eventually moves across the PE segment. The

resulting vibratory sound energy is directed up into the oral cavity where it is articulated into speech (Doyle, 1994; Singer & Blom, 1980). However, unlike ES and TE which are ‘intrinsic’ alaryngeal speaking methods that rely on internal, reconstructed tissues of the pharynx and esophagus, EL speech involves use of an ‘extrinsic’ electronic voicing source that can provide the transmission of vibratory sound energy via neck tissues or intra-orally (Keith & Darley, 1986; Salmon & Goldstein, 1978; Weinberg, 1982). Therefore, the following sections will briefly discuss postlaryngectomy voice and speech produced using an EL.

Two options exist for laryngectomees who use EL speech: neck-type (transcervical) or intra-oral (transoral) methods. However, neck-type EL devices are the most commonly used option (Saikachi, Stevens, & Hillman, 2009). EL speech is produced when the vibratory head of a transcervical EL device is placed against the neck and transmits sound energy through those tissues into the vocal tract. This sound energy moves up into the oral cavity where it is eventually articulated into speech. Conversely, an intra-oral adapter can be added to many neck-type devices in order to provide a sound source that is introduced directly into the mouth where it is then articulated (Doyle, 1994). Regardless of the option used, the EL can act as a primary alaryngeal communication option, as well as serving as a dependable standby in the event that a laryngectomee experiences difficulties or complications with other alaryngeal speaking methods (Hillman, Walsh, Wolf, Fisher, & Hong, 1998). Accordingly, the following paragraph will discuss the history of the EL as a postlaryngectomy communication option and provide insight into how EL voice and speech differ from voice and speech produced using ES and TE communication methods.

History of EL voice and speech. Voice and speech produced without a larynx has a longstanding and rich history². The earliest, commercially available EL was developed by Bell Laboratories in the 1950s (Barney et al., 1959). The vast majority of individuals who undergo TL use an EL in the immediate, postsurgical period (Hillman et al., 1998; Ward, Koh, Frisby, & Hodge, 2003). At one year postlaryngectomy, reports on EL device use have ranged from approximately 30% to 85% (Hillman et al., 1998; Ward et al., 2003). At two years postlaryngectomy, approximately 50% of laryngectomees have been reported to use an EL (Hillman et al., 1998; Mendenhall et al., 2002). These statistics on EL use may reflect the relative ease and prompt voicing provided to many laryngectomees. In addition, these features offer some of the benefits of this alaryngeal communication option when proper speech rehabilitation is provided (Doyle, 1994, 1999; Goldstein, 1978).

When EL devices use is considered, it is also important to highlight the potential difficulties that individuals may experience with other alaryngeal speaking methods. For example, previous reports on the acquisition of ES suggest that less than a third of individuals are capable of acquiring it (Gates et al., 1982) and less than half of those who are successful are unable to produce “acceptable” speech (Damste, 1979). The percentage of laryngectomees who use ES speech at two years postlaryngectomy is approximately 6% (Hillman et al., 1998). However, depending on the speech rehabilitation practices involving the recommendation of clinicians for alaryngeal speech, this number can be as

² The present work is concerned with electronic, artificial larynges. For a thorough review of the history regarding artificial larynges, the reader is referred to Keith, Shanks, and Doyle (2005).

low as 0% (Ward et al., 2003). Hillman et al. (1998) suggest that the decline in ES may be attributed to the growing use of EL speech and to the introduction of TE speech in the 1980s (Singer & Blom, 1980). Doyle and Eadie (2005). However, have commented that medical advancements since the 1980s have led to an improvement in PE segment function postlaryngectomy (e.g., surgical reconstruction techniques), and as a result, those who desire to learn ES may have an increased likelihood of producing it.

Failure to produce speech postlaryngectomy remains a potential scenario for individuals opting to use TE speech as well. In addition to potential problems with PE segment function following TL, air leakage around or through the prosthesis due to candida albicans (a yeast), formation of a fistula or granulation tissue around the fistula, and general inward or outward movement of the prosthesis within the surgically-created fistula are some potential reasons for TE failure (Lewin, 2005; Singer & Blom, 1980; Ward et al., 2003). Together, ES and TE speech failure provide a clear example of the importance of the EL to act as both a primary communication option and as a dependable standby. Collectively, all three alaryngeal communication options are identified as being perceptually different than normal, laryngeal speech. However, EL devices continue to pose unique auditory-perceptual limitations due to the non-biologic, electronic nature of the signal produced (Doyle & Eadie, 2005; Meltzner & Hillman, 2005).

EL speech is often identified by listeners as having a sound quality that is unnatural and mechanical (Bennett & Weinberg, 1973; Doyle & Eadie, 2005; Hillman et al., 1998; Meltzner & Hillman, 2005). Historically, this has led to ES and TE often being the relatively preferred speaking methods when judged by naïve listeners and laryngectomees. This general preference also has been a central theme in the controversy

and divided opinion surrounding EL use between speech-language pathologists (SLP) and medical professionals (Berry, 1978; Doyle, 1994; Duguay, 1978; Gates et al., 1982; Lauder, 1968). In response to these concerns, several authors have continued to uphold that the most important consideration for speech rehabilitation following TL is that all individuals should be exposed to multiple alaryngeal speech options and have the right to choose the option that best suits their needs and lifestyles (Berry, 1978; Diedrich & Youngstrom, 1966; Doyle, 1994; Hillman et al., 1998; Lauder, 1968; Salmon, 1978). This is based on the premise that verbal communication is essential following TL.

Furthermore, McCroskey and Mulligan (1963) argued that it is important to prevent an outright bias against EL device use because this form of alaryngeal communication can provide the majority of laryngectomees with sufficient speaking ability. Still, some laryngectomees view the EL as an inferior alaryngeal communication option because they do not enjoy listening to the EL device and report that listeners may have greater difficulty understanding their speech (McCroskey & Mulligan, 1963).

Laryngectomees often identify speech as an important concern following surgery. However, no significant or consistent link has been found between alaryngeal speech outcomes (e.g., speech intelligibility and/or speech acceptability) and quality of life (QOL) (Eadie, Day, Sawin, Lamvik, & Doyle, 2012; Eadie & Doyle, 2005; Stewart, Chen, & Stach, 1998; Vilaseca, Chen, & Backscheider, 2005). Danker et al. (2010), however, found that there is a strong potential for TL and postlaryngectomy voice and speech to impact psychosocial functioning. In their study, 218 laryngectomees were asked to complete a total of six, validated questionnaires related to social activity (e.g., The European Organisation for Research and Treatment of Cancer Core questionnaire -

EORTC QLQ-C30), speech intelligibility (SI) (e.g., Postlaryngectomy Telephone Test – PLTT), mental well-being (e.g., Hospital Anxiety and Depression Scale), and perceived stigmatization (Questionnaire of Psychosocial Adjustment after Laryngectomy – FPAL) (Danker et al., 2010). Results indicated that the majority of laryngectomees surveyed (i.e., 87%) felt they were stigmatized as a result of their postlaryngectomy voice. This led 54% of laryngectomees to report that they talked less after TL, 40% refused to go anywhere they knew they had to speak, and only a third continued to go to restaurants, meetings, or public events (Danker et al, 2010). In addition, a significant, negative correlation ($r = -0.634, p < 0.01$) suggests that laryngectomees' often withdraw from talking as a result of, amongst other things, a self-perceived reduction in their SI (Danker et al., 2010). This correlation is stronger than findings related to objective SI and withdrawal, which were noted by Danker et al. (2010) to also have a significant (albeit weaker) negative relationship, $r = -0.367, p < 0.01$. These findings are important when discussing SI following TL, considering EL speakers are often reported to have lower SI scores when compared to ES and TE speakers (Barney et al., 1959; Clark & Stemple, 1982; Hillman et al, 1998; Shames, Font, & Matthews, 1963). In addition, out of all three alaryngeal communication options, EL speakers have reported the lowest voice-related quality of life when compared to ES and TE speakers (Moukarbel et al., 2010). Therefore, reduced SI in EL speakers, for example, could account for reduced psychosocial functioning.

Due to the unique nature of postlaryngectomy communication, the influence of alaryngeal method on laryngectomees' voice-related QOL, the research examining SI and

auditory-perceptual characteristics of EL speech, and the strategies that seek to improve these aspects of EL speech will be discussed in the following section.

Alaryngeal Speech and Voice-Related Quality of Life

Loss of the larynx has significant consequences for an individual's physical, psychological, and social functioning (Desanto, Olsen, Perry, Rohe, & Keith, 1995; Doyle, 1999; Eadie, 2003; Hillman et al., 1998; Terrell, Fisher, Wolf, 1998). Further, the acoustic and perceptual changes in one's voice following TL negatively impact quality of life (Cox & Doyle, 2014; Doyle & Eadie, 2005; Moukarbel et al., 2010). The World Health Organization (WHO) (2001) defines QOL as:

...individuals' perceptions of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. (p. 3).

To understand the impact of voice use on QOL, Hogikyan and Sethuraman (1999) created a questionnaire to index the degree to which an individual's voice (and voice disorder) impacts their daily QOL. The Voice-Related Quality of Life (V-RQOL) was originally standardized using individuals with laryngeal-based voice disorders, but has more recently been applied to alaryngeal populations (Bornbaum, Day, & Doyle, 2014; Moukarbel et al., 2010).

Moukarbel et al. (2010) studied V-RQOL scores from 75 laryngectomees: 18 EL speakers, 15 ES speakers, and 42 TE speakers. Data revealed that EL speakers had the lowest self-perceived V-RQOL score while no significant differences were noted between ES and TE speakers. This is supported by previous findings from Clements, Rassekh, Seikaly, Hokanson, and Calhoun (1997) who indicated that TE speakers report the highest satisfaction with their QOL postlaryngectomy when compared to those who

use other alaryngeal speech modes. These increases were attributed by Clements et al. (1997) to TE speakers having reported better self-perceived alaryngeal voice quality and the ability to communicate effectively over the telephone. More recent research examining the V-RQOL in a group of 40 EL speakers found wide-ranging variability in scores (Cox & Doyle, 2014). While a majority of EL speakers were found to have ‘good’ or better V-RQOL scores, approximately 25% of these speakers exhibited ‘poor’ or ‘fair’ V-RQOL scores. This speaks to the varied response from EL speakers and provides support for the idea that not all who use the EL experience a significant communication disability (Cox & Doyle, 2014). Similarly, it may also suggest that individual data are critical when examining a variety of speech outcomes in those who use any method of alaryngeal speech.

Taken together, research using the V-RQOL suggests that higher levels of voice-related QOL are reported by laryngectomees’ whose voice and speech do not interfere with their daily activities. However, although data suggest that EL speakers have lower V-RQOL group scores when compared to ES and TE speakers, not every EL speaker reports a similar level of disability. Thus, it is important to investigate possible factors that can account for EL speakers’ variability in relation to voice-related functioning. Therefore, the following sections will first examine the SI of EL speakers. This will be followed by a review of the acoustic features that comprise the EL voice and speech signal (e.g., intensity, frequency, etc.), and will conclude with a description of findings from listeners’ auditory-perceptual evaluation of EL voice and speech.

Speech Intelligibility

Kent, Weismer, Kent, and Rosenbek (1989) defined SI as, “the degree to which the speaker's intended message is recovered by the listener” (p. 483). Schiavetti (1992) adds that, “any measure of speech intelligibility is a measurement of the interaction between a speaker, a transmission system, and a listener.”(p. 12). Interestingly, SI has been labelled as the most important aspect of speech production, and speech produced using an EL is no exception (Goldstein, 1978).

Since the earliest investigations on EL voice and speech, this communication method has consistently been shown to produce the lowest SI when compared to ES, TE, and normal, laryngeal speakers (Barney et al., 1959). In their study, Barney et al. (1959) compared SI ratings of laryngeal and alaryngeal speakers, including ES and EL speakers. Two experienced ES speakers read words from the Harvard Phonetically-Balanced Word lists (Egan, 1948) and again using neck-type ELs. Based on transcriptions from seven listeners, the EL was judged to have a word intelligibility score of 58.1%, compared to 79% for ES speech and 97.3% for laryngeal speech (Barney et al., 1959). Similar results were found by Shames et al. (1963) who examined the intelligibility of 118 ES and 35 EL speakers. Recordings of words, sentences, and passage stimuli from both speaker groups were orthographically transcribed by a group of five undergraduate students. A statistically significant difference was found for several variables between ES and EL speakers; more specifically, a higher number of correctly articulated consonants by the ES ($M=66\%$) relative to the EL group ($M=58\%$) and a higher word intelligibility score for ES ($M = 54.9\%$) compared to the EL speakers ($M = 35.5\%$) (Shames et al., 1963).

Weiss, Yeni-Komshian, and Heinz (1979) examined word intelligibility for five normal speakers trained to use EL speech. A group of eight listeners identified 90% of word stimuli from the Modified Rhyme Test (House, Williams, Hecker, & Kryter, 1965) when presented in a closed-response format. A group of seven listeners provided correct phonetic transcriptions for 57% of these stimuli. Reduced intelligibility was attributed to the loss of voicing characteristics specific to stop consonants (e.g., voiced for voiceless confusions in word-initial position), in addition to vowel confusions amongst listeners. Thus, consideration of both phoneme and word scoring, as well as the phonetic position of stimuli, are of importance to measures of SI in EL speakers.

To investigate the influence of individual speaker characteristics, Kalb and Carpenter (1981) compared the intelligibility of 5 EL, 5 ES and 5 laryngectomees who used both forms of alaryngeal speech. ES and EL speakers were recorded as they read 50 phonetically-balanced words, while the 5 speakers who were proficient in both ES and EL speech produced the words using both alaryngeal modes. Thirty listeners evaluated recordings of stimuli from all 15 speakers in addition to a sample from normal speakers. Although no alaryngeal group had mean SI scores as high as normal speakers (98.4%), ES speakers had a mean SI score of 78.55% (*range* = 60-96%) compared to 61.81% for EL speakers (*range* = 24- 90%). Interestingly, the speakers proficient in both modes demonstrated intelligibility scores of 67.33% (*range* = 32-96%) and 70.73% (*range* = 28-94%) in ES and EL, respectively. While Kalb and Carpenter (1981) acknowledged that differences exist between ES and EL speakers, they highlighted that individual speaker characteristics rather than speaking mode may account for the differences noted for speakers that used both methods.

Weiss and Basili (1985) examined the intelligibility of six EL speakers who used different EL devices (e.g., Western Electric and Servox). Each speaker read a list of 66 words with each device. Recordings were rated by five SLPs who phonetically transcribed the words. Weiss and Basili (1985) reported that the transcribers identified 33% (*range* = 16-54%) of words recorded with the Western Electric and 36% (*range* = 19-55%) of words using the Servox, but differences between devices were not statistically significant. Although SI scores for EL speakers vary and may appear relatively low when compared to ES and TE speakers, research involving competing noise produced surprising results. In fact, EL speakers have been shown to be more intelligible when competing noise is present. In one of the first studies to compare all three alaryngeal methods in noise, Clark and Stemple (1982) analyzed the SI of synthetic sentences produced by laryngeal, EL, ES, and TE speakers. Twenty adult listeners identified stimuli presented at message-to-competition ratios of 0, -5 and -10dB relative to speech. No significant differences were found between the four speech modes at 0 dB. However, results indicated that the EL speakers were judged to be the most intelligible in both competing noise scenarios. This finding would suggest that aspects of the source signal relative to its own acoustic characteristics must also be considered in the context of alaryngeal speech. Additionally, listeners might understand EL speakers more than ES or TE speakers in realistic communication environments (e.g., social gatherings) where background noise is present.

Numerous studies have also investigated the relationship between listener training and experience on EL speech ratings. McCroskey and Mulligan (1963) studied SI of ES and EL three separate groups of listeners including SLPs, SLP students and naïve

speakers. Five ES and five EL speakers produced stimuli from multiple-choice intelligibility tests (Black, 1944) and listeners provided judgments of SI using a three word closed-set option. Results indicated that SLPs and students comprehended more words from ES speakers (62% and 62.8%, respectively) than EL speakers (57.9% and 56.2%, respectively); naïve listeners comprehended 60.3% of EL speakers' words and 58.2% of ES speakers' words. McCroskey and Mulligan (1963) concluded that although professionals and students might find ES speech more intelligible, those who have not received training or had previous exposure to alaryngeal voice might better understand EL speech. This highlights the role that exposure or training to alaryngeal communication can have on the listener (McCroskey & Mulligan, 1963). Merwin, Goldstein, and Rothman (1985) compared SI of sentences spoken by eight laryngectomees using EL speech before TE puncture *and* TE speech after TE puncture. Twenty-five undergraduate and graduate students identified words heard from one of four options and correct word scores were generated. Results indicated that listeners preferred TE speech, and the authors suggested that EL device noise could have impacted EL intelligibility (Merwin et al., 1985).

Williams and Watson (1985) compared judgments from naïve, 'informed', and expert listeners on TE, ES, and EL speakers' rate of speaking, extraneous noise during speech, intelligibility, and overall communicative effectiveness. They found that naïve (e.g., undergraduate students not exposed to alaryngeal speech), 'informed' (e.g., graduate students who learned about alaryngeal speech through coursework) and expert judges (e.g., SLPs who treated laryngectomees) all rated TE speakers to have significantly better SI than EL speakers, while ES speakers were not different from EL

speakers. In a follow-up study, Watson and Williams (1987) had naïve, informed, and expert judges *and* laryngectomees rate TE, ES, and EL speech. Laryngectomees rated intelligibility of EL speakers significantly different from informed judges, but similar to naïve and expert judges (Watson & Williams, 1987).

The research investigating the SI of EL speech is often reduced relative to other alaryngeal communication methods and, of course, normal, laryngeal speech. In addition, while some naïve listeners might understand EL speech to a lesser degree than TE and ES speech, there is a possibility that group differences may not necessarily be comparable between EL speech and other alaryngeal communication methods. The acoustic characteristics of EL speech, the reliance on an external, electronic voicing source (i.e., ES and TE are considered ‘intrinsic’ methods of alaryngeal communication), and wide-variability of resulting speech demonstrate how complicated such comparisons can become. Further, there are many acoustic factors that may impact SI ratings of EL speech that are dissimilar to other laryngeal and alaryngeal speaking methods. For example, Merwin et al. (1985) found that the device noise produced by EL devices can impact EL speakers’ communication with listeners. To further understand how such factors may directly impact judgments of the EL signal, more research is required to explain the expected acoustic and temporal characteristics of EL voice and speech.

Summary

Within the preceding section, the SI of EL speakers was discussed in addition to an analysis of comparative data between EL, ES, and TE speech. Generally, research has shown that EL speakers have varied SI scores that can range from 16% to 90%. Further, SI scores for EL speakers have consistently been reported to be lower than those for ES

and TE speakers; for example, ES speakers have reported SI scores within a general reported range of 60% to 96%. The reasons often cited for reduced SI in EL speakers include device noise, voicing characteristics of stop consonants, and vowel confusions between listeners. All three of these examples are in part the result of the unique acoustic properties that characterize EL speech; that is, due to its electronic nature, acoustic aspects of the signal must be considered. An understanding of the acoustic features of EL voice and speech permits a greater appreciation for how this alaryngeal communication option specifically can impact communication between EL speakers and their partners. Therefore, the following section will provide a discussion of additional factors that may influence judgments of the EL speech, namely, intensity, signal-to-noise ratio, frequency, and speaking rate.

Acoustic Properties of EL Speech

Intensity and signal-to-noise ratio. Barney et al. (1959) investigated the intensity of the first transcervical EL. They reported sound-pressure levels (SPLs) of approximately 70-75dB when laryngectomees produced vowels. Weiss et al. (1979) reported an average intensity level of 74 dB, although this is based on normal, laryngeal speakers using EL devices to generate speech. Goldstein and Rothman (1976) investigated the speech intensity of ‘good’ and ‘poor’ EL speakers (as cited in Rothman 1978, 1982). First, groupings of ‘good’ and ‘poor’ speakers were formed after six SLPs rated sentences read by 15 EL speakers and then rated ‘speech proficiency’. SLPs used their professional experience to self-define communication proficiency, using an equal-appearing scale ranging from 1 (least proficient) to 7 (most proficient). Five EL speakers with the highest ratings were classified as ‘good’ and five EL speakers with the lowest

ratings were classified as 'poor'. Goldstein and Rothman (1976) found that 'good' EL speakers were able to maintain overall speech intensity, while 'poor' EL speakers exhibited a large variability in intensity levels (as cited in Rothman 1978, 1982). More proficient speakers were credited with properly using their EL devices, which contributed toward improved intensity levels. Specifically, the 'good' EL speakers powered on and shut off their EL devices at the appropriate times during speech and maintained a consistent amount of contact pressure against the neck. The latter is particularly important when considering that good contact must be established and maintained between the vibrating portion of the EL device and the neck.

When parallel contact of the EL device is made between the EL device and neck, a majority (if not all) of the vibratory energy is directed into neck tissue. However, if such contact is not achieved or maintained, the EL signal can radiate into the environment. This resulting device noise has the potential to interrupt communication between EL speakers and their partners. Barney et al. (1959) reported that when an EL device is pressed against the neck and the mouth is closed, the intensity level of externally-radiated EL noise interference is approximately 20-25 decibel (dB) lower than when the vowel 'ah' is produced. Knox and Anneberg (1973) noted that there is a minimum signal-to-noise ratio (SNR) threshold for EL speech, below which device noise can begin to reduce SI. More specifically, naïve and sophisticated listeners achieve higher SI scores when SNRs are a minimum of 4 dB higher than device noise. No significant differences in SI were found when this increased to 9 dB SNL (Knox & Anneberg, 1973). These levels are achieved by appropriate placement of the EL device

against neck tissues to ensure sufficient energy transfer, thereby, minimizing competition of device noise on communication (Knox & Anneberg, 1973).

Several years later, Weiss et al. (1979) reported a mean SNR of 9 dB (*range* = 4 – 15 dB) above device noise for EL speech produced by five laryngeal speakers trained to use the device. Results from their study were similar to Knox and Anneberg (1973) whereby SI scores were the lowest as speakers approached an SNR ratio of 4 dB above EL device noise. However, no predictive relationship was found between SNR measures and intelligibility (Weiss et al., 1979). Interestingly, Weiss et al. (1979) concluded that radiated device noise had minimal impact on overall intelligibility of EL speech. More current research has found improvements in specific phonemic classes (e.g., correct identification of word initial non-nasal sounds) or degradations (e.g., word-final nasals) that can be achieved by filtering EL device noise (Espy-Wilson, Chari, Huang, & Walsh, 1998). Further, while noise levels are believed to have a masking effect on phonemes (i.e., voicing and manner features), the steady-state nature (i.e., lack of frequency variation) of EL devices may permit speaker adjustments to specific acoustic characteristics of the EL signal. Weiss et al. (1979) indicate that the frequency and formant characteristics of the EL signal are one such example. Therefore, the following section will describe the frequency characteristics and their impact of frequency on SI and listener perception of EL voice and speech.

Frequency. Previous research has indicated that “[t]he ideal electronic larynx should produce periodic energy at least throughout the speech range (i.e., up to approximately 4,000 Hz)” (Rothman, 1978, p.104). This should also include “strong low-frequency components” (Barney et al., 1959, p.9). Because neck-type EL devices provide

an extrinsic sound source that must transmit a vibratory signal through tissue, the efficiency of signal transfer is of importance. The process of how sound energy is transmitted through neck tissues is referred to as the 'neck frequency response function' (NFRF) (Meltzer, Kobler, & Hillman, 2003). Briefly, neck tissues following TL and/or radiation therapy can be asymmetric, fibrotic and/or inflamed, which presents a significant challenge for maintaining the vibratory signal energy. Although limited research has been conducted on the impact of EL signal transmission across neck tissues, low-frequency energy deficits and a general lack of frequency range and variation in the EL signal are thought to contribute to its poor quality (Meltzner & Hillman, 2005; Nagle, Eadie, Wright, & Sumida, 2012; Qi & Weinberg, 1991; Watson & Schlauch, 2009; Weiss et al., 1979). Regarding the low-frequency deficits in EL speech, Goldstein and Rothman (1976), Weiss et al. (1979) and Qi and Weinberg (1991) have reported on the decreased spectral energy below 500Hz. Goldstein and Rothman (1976) found that when a Servox device was coupled to neck tissue, sound energy was strongest above 700 Hz, compared to an uncoupled device, which produced strong energy bands below 300 Hz (as cited in Rothman 1978, 1982). Similarly, a coupled Western Electric No. 5 produced energy that was strongest above 600 Hz, but produced strong energy below 385Hz when uncoupled. In essence, all of this information highlights the energy losses in the various frequencies bands when EL devices are coupled to neck tissues. Thus, the electroacoustic characteristics of EL devices and characteristics of a speaker's neck will have a direct impact on the speech produced. Qi and Weinberg (1991) have indicated that the reduction in spectral energy below 500Hz was significantly lower than normal speakers. They found that by enhancing the low-frequency energy of EL speech, listeners reported

improvements in overall voice quality compared to the unenhanced signal. This reduction in low-frequency EL energy when added to the artificial sound quality also impacted the energy spectra of vowels (Qi & Weinberg, 1991). However, Meltzner and Hillman (2005) found that low-frequency energy alone is not the only contributing factor for poor EL speech quality. In their study, Meltzner and Hillman (2005) compared listener ratings across EL speech samples involving numerous acoustic manipulations, including low-frequency enhancement, noise reduction, and frequency variation. These were compared to unmodified EL speech and several samples from normal speakers. Their findings indicate that while low-frequency enhancement was better than unmodified EL speech, the best voice quality was achieved when samples included low-frequency enhancement, device noise reduction, and frequency variation.

Due to the lack of EL frequency variation, Cole, Sridharan, Moody, and Geva (1997) noted that this acoustic characteristic contributes to its perceived mechanical, monotone quality. In addition, Goldstein and Rothman's (1976) study of 'good' and 'poor' EL speakers, they reported that 'good' speakers typically have a mean frequency range of 16.10 Hz ($SD = 2.45$; $range = 13.06-20.26$ Hz), while 'poor' speakers had a mean frequency range of 11.10 Hz ($SD = 3.42$ Hz; $range = 6.61-15.32$ Hz) (as cited in Rothman 1978, 1982). The 'good' EL speakers had a greater range, suggesting that they had more variation in their EL speech signals. This would allow more proficient EL speakers to better approximate the frequency variation patterns of normal speakers, which has been shown to result in higher SI ratings compared to those who lack variation (Laures & Weismer, 1999).

Research investigating the relationship between frequency variation and SI in EL speech has also been conducted by Watson and Schlauch (2009). In their study, one male laryngectomee read a series of sentences (Revised List of Phonetically Balanced Sentences, IEEE, 1969). A total of 60 sentences were recorded. However, 40 sentences were recorded with a device equipped with a pressure-sensitive variable frequency control (i.e., a TruTone EL set at a base frequency of 50 Hz with a range of 300 Hz) and 20 sentences were recorded using the device with a fixed frequency (i.e., a base frequency of 65 Hz without using the tone control). SI evaluations of 20 naïve listeners for sentences in both conditions were compared. Results indicated that intelligibility was at least 10% higher when speakers used an EL device with variable frequency compared to a flattened frequency. Watson and Schlauch (2009) suggested that the improvement in SI with variable intonation may be consistent with the work of Laures and Weismer (1999), which suggests that variable intonation may improve SI. Further, Watson and Schlauch (2009) suggested that listeners might have difficulty identifying speech when there is a drastic departure in frequency, which limits use of certain cues from rising or falling intonation patterns. More recently, Nagle et al. (2012) conducted several experiments to examine the impact of EL frequency on SI, speech acceptability (ACC), and perceived gender. The first of these experiments investigated the impact of three EL frequencies on SI; 34 normal speakers read sentences and a reading passage using EL devices set at 75 Hz, 130Hz, and 175Hz. Results indicated that SI was highest for the speakers using EL devices set at 75 Hz (Nagle et al., 2012). The second experiment investigated listener ratings of gender and ACC for normal speakers using EL devices set at the same three frequency levels. Stimuli included the reading passages from 22

speakers included in the first study. Results indicate that listeners rated male and female speakers as being 'more male' as EL device frequency decreased from 175Hz to 75Hz, while female speakers were rated as 'more female' as device frequency moved from 75Hz to 175Hz. When the same listeners were told that the speakers using devices set at the lowest frequency (e.g., 75 Hz) were female, ACC ratings decreased. Lastly, judgments of ACC were more favourable as SI improved (Nagle et al., 2012). When viewed together, research on EL device frequency confirms the interrelatedness between device frequency, the impact of neck tissues on the transmission of EL signal energy, and the general lack of frequency variation. This body of research also highlights how these characteristics can impact listener ratings of overall EL voice quality and SI. However, in addition to the influences of intensity and frequency characteristics of the perception of the EL signal, speaking rate also must be considered. The following section will discuss speaking rate in EL speakers and how alterations in rate can facilitate the overall changes in listener perception.

Speaking rate. Alaryngeal speech generally requires a slower rate than that of normal speech (i.e., 149.5 to 196.1 words per minute) (Doyle & Eadie, 2005). Research on alaryngeal speech rates suggests that ES speakers' have a speaking rate of 99.1 to 114.3 words per minute (wpm) (Hoops & Noll, 1969; Robbins, Fisher, Blom, & Singer, 1984; Snidecor & Curry, 1959). Male and female TE speakers typically have a speech rate of 127 and 138 wpm, respectively (Robbins et al., 1984; Trudeau & Qi, 1990). Finally, Hillman et al. (1998) have reported a speech rate of 130 wpm for EL speakers. Of all three speech modes, TE speech has been found to be the closest to normal speakers, primarily due to TE speakers' access to pulmonary air for speech (Doyle et al.,

1988; Hillman et al., 1998). The speech rate of ES speakers, however, is negatively impacted by their need to regularly insufflate air. For EL speakers, their rate reduction may be secondary to changes in articulation (i.e., over-articulate) while using an EL device (Doyle & Eadie, 2005).

In a study investigating the speech rate of EL speakers, Goldstein and Rothman (1976) found that those rated as 'good' alaryngeal speakers had a mean speech rate of 3.86 seconds ($SD = 0.36$) when reading 12-word sentences. 'Poor' speakers, however, had a mean rate of 6.48 seconds ($SD = 2.23$). Further analysis of the data indicated that 'poor' EL speakers often paused more during speech and had EL activation/deactivation issues (as cited in Rothman, 1978, 1982). Based on their analyses, speech rate was found to be the greatest predictor of EL speech proficiency.

Williams and Watson (1985) compared listener ratings of speaking rates across all three alaryngeal communication modes. Naïve, graduate student, and SLP listener groups made judgments of speech rate from videotaped samples of 33 alaryngeal speakers (11 EL, 12 ES, 10 TE) who completed four different speech tasks (e.g., automatic speech, reading, picture description, and conversation). Based on a 7-point rating scale (e.g., 1 indicating 'excellent' and 7 indicating 'poor'), SLPs judged the rate of TE speakers more favourably than EL and ES users, while 'informed' judges rated TE and EL speakers similarly (and more favourable than ES speakers). Naïve judges rated TE speaker's rate of speech more favourably than ES, but not significantly different from that of EL speakers. Watson and Williams (1987) explored this further by including laryngectomees as judges alongside naïve, informed, and expert listeners. Findings indicated that

laryngectomees rated EL speakers similar to naïve and expert judges, but significantly different than the informed listeners on rate and SI (Watson & Williams, 1987).

Taken together, the above findings suggest a significant relationship between ratings of a faster (rather than slower) rate of speech and an increase in effectiveness (Hoops & Noll, 1969; Snidecor & Curry, 1959). The importance of this finding is one that recognizes that alaryngeal speakers who are judged to be more ‘effective’ may produce speaking rates that approximate those of normal speakers. Therefore, the closer alaryngeal speakers are to producing normal speaking rates, the more listeners might deem them as ‘effective’ communicators. However, regardless of alaryngeal speech mode, more rapid speech rates may also result in altered articulation with its potential to negatively impact SI. Consequently, interaction between multiple factors must be considered when addressing concerns regarding alaryngeal voice quality and/or SI. This would appear to be of particularly importance in the context of speakers who use the EL. Alongside the above mentioned discussion of intensity, frequency, and speaking rate, it is important to also consider the contribution of suprasegmental features to EL speech production and perception. Therefore, the discussion to follow will focus on the importance of intonation, stress, rhythm and word juncture on the perception of speech.

Suprasegmental Features of EL speech

While there is a paucity of research on the suprasegmental features of EL speech, the suprasegmental features of intonation, stress, rhythm and word juncture can impact EL speech production and its perception.

Intonation. Gandour and Weinberg (1983) refer to intonation as, “the pitch changes that occur during a sentence” (p. 142). The ability to raise pitch contours during

speech is often perceived as a question, whereas a falling pitch contour is often perceived as a statement (Gandour & Weinberg, 1983). More objectively, it is the fundamental frequency (F_0) contour that provides cues to listeners in order to discern a statement from a question. Several experiments compared listener perceptions of intonational contrasts produced by normal, ES, TE, and EL speakers (Gandour & Weinberg, 1983; 1984). Forty naïve listeners rated the sentence 'Bev loves Bob' produced with varying intonational contrasts, resulting in a statement and a question. Listeners were required to identify if they heard statements or questions while listening to stimuli. Findings indicate that all speaker groups except EL speakers were able to achieve a high degree of intonational contrast. Further, the data indicate that EL speakers are generally unable to code intonational contrasts; only one of three speakers using a Western Electric No. 5 with variable intonation control was able to adequately achieve intonational contrasts. Further acoustic analysis of the data indicate that the inability for many EL speakers to control F_0 remains the primary reason EL speakers cannot realize intonation patterns (Gandour & Weinberg, 1983, 1984). Other suprasegmental features, such as stress, appear to be more complex acoustically; although stress can be realized through changes in F_0 , it appears that cues related to stress can be provided by intensity and durational changes.

Stress. Although there is no single acoustic parameter that clearly identifies stress, this feature is often realized by normal laryngeal speakers through the use of a higher F_0 , greater intensity, and longer durations of syllables (Lehiste, 1976). Several types of stress include contrastive, lexical, and syntactic. Briefly, contrastive stress refers to an individual's ability to increase F_0 in order to produce a question rather than a statement. Lexical stress occurs within words in order to change the syntactic category of

the word. Lastly, syntactic stress enables speakers to choose between the production of compound nouns and noun phrases.

Several studies have investigated the ability for EL speakers to realize all of these types of stress (Gandour & Weinberg, 1982, 1984, 1985; Gandour, Weinberg, & Grazione, 1983; Gandour, Weinberg, & Kosowsky, 1982). Gandour et al. (1982) discovered that EL speakers were better able to realize contrastive, lexical, and syntactic stress when compared to intonation. Acoustic analyses of stress patterns in EL speech suggests that the majority of EL speakers are only able to vary the durational properties of speech with no consistent ability to vary F_0 or intensity (Weinberg & Gandour, 1985). If provided with an EL device that enables F_0 , then EL speakers might vary frequency and duration, but not intensity (Weinberg & Gandour, 1985). Findings indicate that the realization of stress is more complex acoustically than intonation, especially for the EL speaker. While the lack of F_0 variability lead to poor intonational contrasts in EL speech (Weinberg & Gandour, 1984), EL users are able to sufficiently produce contrastive, lexical, and contrastive stress patterns (Weinberg & Gandour, 1982, 1983, 1985; Weinberg et al., 1982, 1983). Further, in the absence of F_0 and intensity changes, further support is provided for the notion that stress is not determined by any single acoustic parameter. Rather, it is driven by frequency, intensity, and duration during EL speech production.

Rhythm. Weinberg, Gandour, Petty, and Dardarananda (1986) define rhythm according to the timing of syllables and the timing of the space between them. In addition, Martin (1972) identified rhythm as the pattern of stress on a series of syllables. Over an entire speech utterance then, rhythm would involve numerous accented or

stressed portions that "...occur with some regularity, regardless of regardless of tempo (fast, slow) or tempo *changes* within the pattern (accelerate, retard)." (Martin, 1972, p. 490). Given that EL speakers have a speech rate that typically falls within the range for normal speakers (Hillman et al., 1998), in addition to their ability to produce stress patterning (Gandour & Weinberg, 1985), EL speakers should be able to produce a relatively normal rhythm during speech. However, there are no data to support this conclusion at present. This supposition is based on data pertaining to EL speakers' speech rate and the realization of stress. It is important to note that part of EL speech rehabilitation involves the use of a slower rate of speech and over-articulation while speaking (Doyle, 1994). The combination of reducing speech rate and over-articulating introduces more pauses between words and lengthens individual speech sounds (Picheny, Durlach, & Braida, 1986). Therefore, EL speakers might be able to properly accent various syllables throughout an utterance, which further acts to separate these syllables within or between word junctures.

Juncture. Juncture refers to the relationship between sounds within words or between words within continuous speech. Two common way for realizing juncture is through pauses and word boundaries (Skandera & Burleigh, 2005). Interestingly, when actively attempting to slow speech rate, over-articulate, and make speech clearer, research indicates that individuals insert more pauses and increase the duration of individual speech sounds (Picheny et al., 1986). These productive aspects of speech can lengthen the juncture between sounds within words and between words within sentences. Further, Skandera and Burleigh (2005) indicate that there is the possibility that "different types of juncture are blurred in rapid speech" (p. 62). Therefore, the consequences of

modifying the productive aspects of speech (e.g., slowing down rate, over-articulation) could potentially facilitate improved retrieval of the spoken message by the receiver. Given that SI is one of the most important aspects of human communication, junctural cues are an important consideration, particularly if listeners are required to discern the type of message being communicated (e.g., statement versus question).

Summary

Within the previous section, several acoustic, temporal and suprasegmental features of EL speech and their role in SI have been highlighted. Previous research indicates that the SNR of EL speech must be at least 4 dB or higher than EL noise to be most efficient (Knox & Anneberg, 1973), and that EL speakers are relatively intelligible in low environmental noise (Verdolini et al., 1985). Furthermore, SI improves (even if only slightly) when the low-frequency energy of the EL speech signal is enhanced (Meltzner & Hillman, 2005; Qi & Weinberg, 1991). When EL speakers are able to vary the frequency during speech, SI has been reported to improve to levels of at least 10% (Watson & Schlauch, 2009). In addition, SI improves when EL devices are set at lower fundamental frequencies (e.g., 75 Hz) and this effect decreases as EL frequencies increase (e.g., to 175 Hz) (Nagle et al., 2012). Speaking rates of EL speakers are often reduced and rated inferior to TE and normal speech (Williams & Watson, 1985). Several studies have even attempted to assess acoustic characteristics (e.g., frequency, intensity) of EL speech using subjective listeners' ratings in order to describe their potential impact on the perception EL voice and speech (Nagle et al., 2012; Williams & Watson, 1985). Further, EL speakers' inability to vary F_0 or change the intensity of their speech during conversation greatly impacts their ability to realize some suprasegmental features of EL

speech. Unfortunately, there continues to be a relative lack of comprehensive research focusing on listener judgments of communication postlaryngectomy, including EL voice and speech (Doyle, 1994). Therefore, the following sections will discuss auditory-perceptual features of EL voice and speech in greater detail.

Perceptual Features of EL Speech

Speech acceptability. The perceptual dimension of ACC refers to an assessment of speech in which listeners are asked to make collective judgments based on pitch, rate, understandability, and voice quality (Bennett & Weinberg, 1973). This is a broad perceptual dimension that has been extensively described in early investigations of ES speech (Berlin, 1965; Shipp, 1967). Berlin (1965) used a seven-point rating scale for ES speakers ranging from 1 indicating ‘highly acceptable’ speech to 7 indicating ‘unacceptable’ speech. Shipp (1967) had participants rate alaryngeal speakers (using unspecified alaryngeal speech methods) using a five-point scale of 1 (least acceptable) to 5 (most acceptable) without actually defining ACC and found that fundamental frequency was correlated with ACC ratings. In addition, ACC has been used in a similar manner to assess ‘speech quality’, which also is concerned with the “acceptability of the speech to listeners” (Meltzner & Hillman, 2005, p. 767). Bennett and Weinberg (1973), however, were the first to provide comparative data regarding ACC across normal, ES, and neck-type EL speakers. Eighteen alaryngeal speakers read the second sentence of the Rainbow Passage (Fairbanks, 1960) and were rated by 37 naïve adults with little familiarity of alaryngeal voice. Listeners were asked:

In making your judgments about the speakers you are about to hear, give careful consideration to the attributes of pitch, rate, understandability, and voice quality. In other words, is the voice pleasing to listen to, or does it cause you some discomfort as a listener? (Bennett & Weinberg, 1973, p. 610)

ACC ratings ($1 = \text{low}$, $7 = \text{high}$) were noted to be 5.48 for normal speakers, 2.54 for ES speakers, and 1.59 EL speakers. At least half of listeners indicated that their low ACC ratings of EL speech was due to speech sounding “mechanical”, quality not sounding “normal” and being “monotonous” (Bennett & Weinberg, 1973, p. 615). These findings highlight the potential impact that the acoustic variables of EL speech, which create an unnatural, mechanical and monotone sound, can have on listener perception.

There are many benefits in using ACC as a rating tool, not only to understand speech rehabilitation outcomes, but to understand listener perceptions across a series of variables related to the speech signal (i.e., pitch, rate, understandability, and voice quality). This is supported by the perceptual work of O’Brian et al. (2003), who indicated that, “... a more important outcome of treatment, at least to the client, is the extent to which treatment increases the social acceptability of speech.” (p. 504). While no definition was provided for ACC in their work, O’Brian et al. (2003) stressed the importance of listener judgments for measuring treatment outcomes.

While some might argue that ratings of ACC should focus on individual components separately, alaryngeal speech is multidimensional in nature and numerous variables contribute to ACC in a collective manner. For example, ACC ratings of EL ‘voice quality’, which is concerned with the ‘acceptability of the speech to the listener’, were found to be improved when they included low-frequency enhancement, EL device noise reduction, and frequency variation (Meltzner & Hillman, 2005). Alongside acoustic variables, even laryngectomees’ ratings of self-esteem and general well-being have been found to be correlated with listener judgments of ACC (Blood, Luther, & Stemple, 1992). Additional concerns for EL speakers are based on the highly visual nature of EL device

use and its unique sound. For example, Doyle (1994) commented that the visual nature of EL device use could negatively impact ACC judgments of EL speech, especially when compared to the other intrinsic forms such as ES and TE speech.

Given the multidimensional nature of EL speech, it is reasonable that ratings of ACC must remain a perceptual composite across numerous variables in order to provide a rich understanding of listener perceptions of EL speakers in various communication contexts. Furthermore, ACC ratings not only highlight the potential consequences of both the acoustic and visible nature of EL speech, but may provide support for the use of other auditory-perceptual ratings of EL voice and speech as well. For example, Bennett and Weinberg's (1973) use of ACC included the requirement for listeners to consider, "...is the voice pleasing to listen to, or does it cause you some discomfort as a listener?" (p. 610). The introduction of the term 'discomfort' is closely linked to the more recent use of listener comfort (LC) scales (O'Brian et al., 2003; Susca & Healey, 2001; 2002). While assessments of LC have been often used to investigate pre and post-treatment speech outcomes in persons who stutter, there is potential utility for them in other communication disorders, including EL speakers. Therefore, the following paragraph will outline the use of LC as a perceptual feature of importance in the evaluation of EL voice and speech.

Listener comfort. Clinically, LC is a perceptual dimension that has the potential to capture "...the sense of listeners' feelings of what it would be like to communicate with a speaker." (O'Brian et al., 2003, p.504). LC was originally used by Susca and Healey (2001, 2002) to examine listener perceptions of simulated, fluent and disfluent speech samples of an adult speaker. O'Brian et al. (2003) extended this work to measure how

comfortable listeners are when communicating with persons who stutter. O'Brian et al.

(2003) defined LC to listeners as,

...how comfortable you would feel listening to the person's speech in a social situation. Your response should reflect your feelings about the way the person was speaking (i.e., how comfortable you would feel listening to them), not what the person was saying or how their personality affected you. (p. 509).

This perceptual construct has also been applied to populations with voice disorders (Eadie et al., 2007). In essence, Eadie et al. (2007) suggest that LC might be useful measure for determining the impact of voice disorders on communication partners; more specifically, how comfortable these individuals are while communicating with someone who has a voice disorder. Unfortunately, to date, only one study has examined LC relative to alaryngeal speech, specifically in TE speakers. This lack of exploration exists even though LC appears to provide a broad understanding of listeners' feelings toward voice and speech disorders in a similar fashion to ACC (Doyle, Day, Dzioba, Bornbaum, & Sleeth, 2011). This is of vital importance when considering the multidimensional nature of alaryngeal speech rather than a sole focus on individual (or, 'unidimensional') parameters, such as intensity or speech rate (Doyle & Eadie, 2005). Furthermore, Eadie (2003) stated that if impairments "are associated with undesirable deviation, discontinuity, or discomfort, then they give rise to a need for corrective actions" (p. 11). This speaks to the fundamental nature of EL speech whose auditory-perceptual characteristics have been repeatedly rated as inferior to other alaryngeal modes of speech. Thus, a need for 'corrective action' could imply therapeutic interventions attempting to improve aspects of EL speech to reduce the potential for making listeners 'uncomfortable', or deeming EL voice and speech as 'unacceptable'.

Summary

Within this section, perceptual features of EL speech have been discussed. The definition of ACC has evolved to include a broad perceptual composite involving speaking rate, pitch, understandability and voice quality, and has been applied to EL speech (Bennett & Weinberg, 1973). EL speech is often perceived to be ‘less acceptable’ than ES or TE speech due to the unnatural, monotonous, and mechanical quality that separates it from these other alaryngeal communication options. The highly visible nature of EL device use also has been suggested to reduce ACC ratings with potential implications to impact the communication exchange between an EL speaker and listener. Within the dyadic interaction between the EL speaker and his or her communication partner, auditory-perceptual ratings are suggested to be important for assessing treatment outcomes.

Since research has shown that listener perceptions can negatively impact laryngectomees’ psychological and social functioning (Blood et al., 1992), it is important to understand the potential social consequences of negative listener perceptions on EL speakers. Therefore, the following section will investigate the potential social consequences laryngectomees might face while using EL speech and how it can impact their physical and psychosocial functioning.

Social Consequences of EL Speech

It is well-documented that society has specific expectations for all of its members and they must adhere to these expectations to avoid being stigmatized (Goffman, 1963). Doyle (1994, 2012) has discussed the potential impact of the noise created by EL devices and the visible nature of its use. Salmon (2005) has commented that the EL can be considered a visual distractor that has the ability to “...divert listeners’ attention from

what the [alaryngeal] speaker is attempting to communicate” (p. 64). The previously described deficits in the acoustic signal of EL speech and how they are perceived by listeners is of particular importance when considering the potential violations of societal expectations. The abnormal nature and mechanical sound of EL speech may place laryngectomees at a greater risk for experiencing more communicative challenges than non-EL users (Doyle, 1994; 1999). Any of these challenges increases the potential for EL users to experience restrictions to social participation (Doyle, 1999).

Cox and Doyle (2014) analyzed Voice-Related Quality of Life (V-RQOL) questionnaire data from 40 laryngectomees who used EL as their primary alaryngeal communication method. Approximately 25% of EL users reported a “poor/fair” voice-related quality of life that included vocal challenges such as not being heard in noisy situations or having trouble communicating on the telephone (Cox & Doyle, 2014). The remaining EL users reported that they generally experienced fewer challenges in daily life, while indicating some similar difficulties reported by EL speakers with less satisfactory VRQOL. Relative to social-emotional functioning, some of these EL speakers reported increased anxiety or depression because of their speech (Cox & Doyle, 2014). These findings highlight that the EL can have a wide-ranging impact on EL speakers’ physical, social and emotional functioning postlaryngectomy. However, the exact reasons related to EL device function are not identified. It appears that the majority of difficulties experienced by laryngectomees are related to acoustic and perceptual characteristics of EL speech (e.g., SI, ACC, etc.). Therefore, it is important to seek information on how EL speakers can improve their speech and how listeners perceive them.

Experimental Attempts to Improve Acoustic Characteristics of EL Speech

Several attempts have been pursued in order to improve the acoustic characteristics EL devices and the resulting speech produced. A substantial limitation of most neck-type EL devices is the fact that they generate an electronic background noise that can be heard during EL speech. Sound energy escapes into the surrounding environment while the vibratory head of the EL makes contact with the speaker's neck. Espy-Wilson et al. (1998) investigated the impact of the extraneous noise generated by EL devices on listener preference and SI. Through the use of adaptive filtering to remove the background noise, Espy-Wilson and her colleagues (1998) compared unmodified EL speech to modified EL speech signals with the noise component removed. They found that the removal of background noise can lead to a significant improvement in listener preferences, but interestingly, had no significant impact on SI (Espy-Wilson et al., 1998). This finding highlights the complex relationship between the EL source quality and SI; that is, an improvement in one feature or signal parameter does not necessarily lead to improvement in other(s). In fact, work by Wong (2003) has shown that SI may be sacrificed with attempts to make the EL signal more acceptable to the listener. Since there are numerous variables that contribute to voice quality and SI (e.g., articulation, device noise, speaking rate, etc.), it appears that attempts to improve EL speech should seek to address several of these parameters.

To date, one study has experimentally manipulated several acoustic variables of EL speech and made comparisons of these modifications via listener ratings. Meltzner and Hillman (2005) examined three EL signals with enhancement of the low-frequency energy, reduced device noise, and pitch variations to mimic normal 'laryngeal' speech.

Their findings indicated that listeners favour an EL speech signal when all three modifications are present. However, Meltzner and Hillman (2005) concluded that the improved EL speech signal did not fully meet listeners' expectations when compared to normal speech. Meltzner and Hillman (2005) indicated further that there are other unexplored factors that may contribute to reduced quality for EL speech.

First, while experimental attempts to advance the electro-acoustic characteristics of neck-type EL devices have improved frequency and noise related-aspects that are perceptually salient, current EL devices continue to produce a noisy and abnormal signal that impact listeners' judgments of overall quality and SI (Meltzner & Hillman, 2005). Therefore, a more feasible approach to address the perceived shortfalls of EL speech might rely upon advancing speech rehabilitation through the use of existing or novel therapeutic techniques. This may provide the EL user with more immediate enhancements that influence how their speech is perceived by others. In order to verify this possibility, the proposed study will examine the use of *clear speech* (CS) as a therapeutic technique aimed at improving several characteristics related to communication by EL speakers.

The underlying premise of CS seeks to slow a speaker's rate, in addition to encouraging the speaker to over-articulate. By doing so, these changes may permit improvements in SI and auditory-perceptual dimensions (e.g., speech acceptability). Given that over-articulation and speaking rate are central to producing effective and understandable EL speech, CS is able to address both of these aspects at once. Therefore, the following section will detail the key features of CS and its application for improving the SI and auditory-perceptual characteristics of EL speech.

Clear speech. CS is a style of speaking that requires speakers to produce speech as clearly as possible (Krause & Braida, 2002; Picheny, Durlach, & Braida, 1985; Picheny et al., 1986; Smiljanić & Bradlow, 2009; Uchanski, 2005). The concept of CS can be traced back to the work of Snidecor, Malbry, and Hearsey (1944) which focused on improving communication over military radio systems. Snidecor et al. (1944) found that improved SI was facilitated by instructing participant speakers to produce louder speech and by increasing mouth opening, speaking at a slower rate, and making a deliberate effort to speak more clearly (as cited in Picheny et al., 1985). Several years later, Tolhurst (1957) investigated the impact of speaking rate on word intelligibility and listener preferences. Recordings of one adult male who read words using three speech rates: normal, prolonged, and staccato (Tolhurst, 1957). Twelve panels of listeners ranging from 12 to 15 listeners per panel identified 86% of prolonged words, 84% of words in the normal delivery, and 77% in the staccato delivery. In addition, listeners preferred normal and prolonged conditions over the staccato delivery, but no significant differences were found between these preferred conditions (Tolhurst, 1957). This may suggest that CS, which involves a reduced rate of speech and over-articulation, could improve EL speakers' SI without negatively impacting listener judgments.

To date, CS has been used to improve the SI and reception of verbal communication for individuals living with a variety communication disorders. Picheny et al. (1985, 1986) found that CS improved the SI of speech spoken to hearing-impaired listeners when compared to speech spoken in a standard, conversational manner. Specifically, five listeners with sensorineural hearing loss listened to nonsense sentences that were recorded by three male talkers in a typical, conversational manner and while

using CS; listeners either orthographically transcribed or spoke their responses. Results indicate that listeners found sentences recorded in CS more intelligible than the conversational sentences. In fact, SI improved 17% for speakers using CS and increases were found across all phoneme classes (Picheny et al., 1985). In a follow-up study, Picheny et al. (1986) examined acoustic aspects of conversational and CS. Fifty nonsense sentences were used in both conditions to measure speaking rate, pause time, fundamental frequency distribution, and long-term spectra. Results showed that CS was significantly slower than conversational speech (e.g., 90 to 100 wpm versus 160 to 200 wpm, respectively) (Picheny et al., 1986). This change was accounted for by lengthening individual speech sounds, as well as an increase in the number of pauses added between individual words (Picheny et al., 1986). Other findings for CS suggest that vowels are less likely to be reduced (or, become 'schwa-like'), there were fewer eliminations of stop bursts (i.e., 15% of the time in CS compared to 60% of the time in conversational speech), and durational changes occurred with tense vowels and plosives (Picheny et al., 1986). The findings from these studies indicate that there is an advantage of CS over conversational speech by an average of 17 percentage points for sentence intelligibility. In addition, CS has the ability to impact phonological (e.g., vowel modification) and phonetic-level (e.g., segmental durations) aspects of speech. However, Picheny et al. (1986) suggest that speaking rate alone cannot account for the acoustic modification of vowels, durational properties of segments and overall improvement in intelligibility at the phoneme and word-level.

In a third study, Picheny, Durlach, and Braida (1989) investigated variables contributing to the speaking rates of CS and conversational speech. Speaking rates were

modified so that the original CS speaking rate of 100 wpm was doubled (e.g., 200 wpm) and the original conversational rate of 200 wpm was halved (e.g., 100 wpm). Five hearing-impaired speakers with sensorineural hearing loss listened to the original sentences presented in unmodified, modified, and restored versions. Results indicated that after modifying sentences recorded in a conversational manner to match the speech rate of CS, word intelligibility scores could not be improved, and in fact, actually decreased (e.g., $M=53\%$ for unmodified versus $M=40\%$ for modified). While this body of research provides an in-depth analysis of the acoustic and temporal characteristics of CS, there is a general consensus is that CS improves overall SI. Furthermore, while the abovementioned research focused on speech reception by hearing-impaired listeners, the manipulation of EL speaker's speech rate and articulatory patterns might provide a similar 'CS benefit' to listeners.

Alongside improvements for individuals with hearing impairment, CS has been shown to improve SI for individuals with dysarthria (Beukelman, Fager, Ullman, Hanson, & Logemann, 2002), and more recently, shows promise for individuals living with Parkinson's disease (PD) and multiple sclerosis (MS) (Tjaden, Sussman, & Wilding, 2014). Beukelman et al. (2002) compared the SI of four different speech supplementation strategies (i.e., strategies involving cueing to assist communication). The four speaking conditions involved habitual speech, CS, alphabet supplementation, and topic supplementation. Nine individuals with dysarthric speech secondary to traumatic brain injury (TBI) read sentences using all of the supplementation strategies. Results indicated that the overall SI of sentences was greatest in alphabet supplementation (100%), followed by topic supplementation (96.8%), CS (95.1%), and HS (87.1%) (Beukelman et

al., 2002). Aside from the improvements in the alphabet and topic supplementation strategies, CS improved SI by 8% when compared to habitual speech (Beukelman et al., 2002).

Hanson, Beukelman, Fager, and Ullman (2004) followed-up this research in an effort to examine listener preferences (e.g., ACC and effectiveness) toward speech supplement strategies; the same participant speakers were used. Speakers were videotaped while speaking 12 sentences, three sentences for each of the previous speech supplementation strategies. Sixty participant listeners comprised of 15 naïve listeners, 15 SLPs, 15 allied health professionals, and 15 family members viewed the videotapes. Each listener rated sentences based on how “acceptable” and “effective” the speakers’ communication was throughout all conditions. Each participant was encouraged to use his or her own interpretation of the terms “acceptable” and “effective” while rating speakers’ communication (Hanson et al., 2004; Richter et al., 2003). Results indicated that speech using alphabet supplementation was the most preferred strategy, followed by topic supplementation, CS, and habitual speech. Hanson et al. (2004) noted that, while alphabet supplementation was the most preferred strategy, there were significant negative correlations between listener ratings and SI. This indicates that listeners could find a strategy unacceptable even in the presence of improvements of SI (Hanson et al., 2004). This differed for CS, however, whereby significant correlations of 0.63 and 0.73 were found between listener ratings and SI scores for SLPs and family members, respectively. Overall, speech strategies that are the most preferred are not always those that correspond to the greatest levels of SI. When listeners judged the acceptability and effectiveness of

individuals using CS, however, listener ratings correlated with SI scores, further supporting the assumed relationship of greater preference for improved SI.

Tjaden et al. (2014) investigated the impact of reduced speaking rate, increased intensity, and CS in speakers with MS and PD. Seventy-eight individuals, including 32 healthy, normal-speaking controls, 16 individuals with PD, and 30 individuals with MS read sentences in habitual, CS, loud, or slow conditions. Speakers were instructed to speak at half of the rate of their normal speech for the slow condition, which was achieved by prolonging words and producing stimuli on a single breath. For the CS condition, speakers were specifically asked to say each sentence more clearly (e.g., twice as clear compared to normal speech). This was achieved by speakers exaggerating their speech movements as though they were speaking in a noisy environment or to someone with a hearing impairment. Fifty listeners made SI judgments of sentences presented in multi-talker babble and another group of 50 listeners judged sentences in multi-talker babble for speech severity using visual analogue scales (VAS). SI was defined as how well listeners understood the sentences and it was scaled along a continuum that ranged from '*understand everything*' to '*cannot understand anything*'. The severity of their speech was based on judgments that crossed voice, resonance, articulatory precision, ranging from '*no impairment*' to '*severely impaired*'. Results indicated that SI scores improved by 7-11% in the CS condition for both speaker groups. The loud and CS conditions resulted in significantly better SI scores than habitual condition, but SI scores did not significantly differ between loud and CS. Speech severity ratings were to be less severe when PD speakers used loud and CS relative to habitual speech, while severity ratings were improved in the loud and habitual conditions for individuals with MS.

Finally, SI and severity scores were found to be significantly correlated for both MS and PD groups (0.66 and 0.63, respectively). Together, these findings indicate that CS provides a significant improvement in SI, especially in challenging communication contexts (e.g., multi-talker babble). Given the documented success of CS at improving the speech characteristics in individuals with a variety of communication disorders, in addition to the relative ease of implementing this strategy through simple instructions, CS also may be a viable option for improving SI and global listener assessments of EL speakers.

Summary

CS has been used to improve communication for over 70 years. While it began as a style of speaking to improve the speech of military personnel over radio telecommunications, it has evolved to become a viable therapeutic technique to improve SI for both normal hearing and hearing impaired individuals. The therapeutic application of CS has resulted in SI increases that have ranges from 17% to 26% for individuals with hearing impairment (Payton, Uchanski, & Braida, 1994; Picheny et al., 1985) and approximately 7% to 11% for individuals with dysarthria (Beukelman et al., 2002; Hanson et al., 2004; Tjaden et al., 2014). It has been suggested that individuals with either a speech or hearing impairment primarily benefit from features associated with the slower-rate-of-speech and over-articulation due to CS (Picheny et al., 1985, 1986; Tjaden et al., 2014). These hallmark features of CS may also assist EL speakers to coordinate the productive/articulatory aspects of speech alongside the timing of the on/off operation of EL devices. Interestingly, a slow rate of speech, over-articulation, and device timing are of central importance when laryngectomees receive initial instruction on EL device use

(Doyle, 1994). Since CS has been applied to diverse populations of individuals with speech and hearing disorders, the series of studies to follow are the first to investigate the application of CS in EL speakers.

Rationale for the Present Studies

EL devices have remained relatively similar in both design and speech quality since their development in the 1950s. In addition, research has found that EL speech is inferior with respect to its acoustic (e.g., frequency, intensity, and rate) and auditory-perceptual characteristics when compared to ES and TE speech (Bennett & Weinberg, 1973; Kalb & Carpenter, 1981; McCroskey & Mulligan, 1963; Snidecor, 1968; Williams & Watson, 1985; Williams & Watson, 1987). Unfortunately, the majority of previous research regarding EL speech may not be entirely applicable to current devices. For example, Pindzola and Moffett (1988) acknowledged that previous work which addressed SI of EL speech were completed with devices that were no longer in use (e.g., Western Electric 5A), or that the exact type of device was not identified. Further, research supports investigations into improving postlaryngectomy speech (i.e., frequency variation, reduction of device noise, etc.) through manipulation of EL signals (Espy-Wilson et al., 1998; Meltzner & Hillman, 2005). However, while aspects of EL signal “quality” have been observed experimentally, SI has not improved significantly. In fact, Wong (2003) showed that attempts to enhance EL signal quality can have a negative impact on SI. In Wong’s (2003) study, the SI of a commercially available EL device (i.e., Servox) and a modified EL device using adaptive filtering (the Prototype Electro-Larynx Massachusetts Eye and Ear Infirmary) were compared. While previous research indicated improved listener preference for the modified EL device (Beaudin, 2002), Wong’s (2003)

results revealed that speakers using the unmodified device were judged to be more intelligible (66%) than when using the prototype EL (59%). Thus, concerns related to global assessment of the EL signal by listeners and the resulting influence on SI remain of importance. Applying the documented success of CS in improving SI with other clinical populations (Beukelman et al., 2002; Hanson et al., 2004; Payton et al., 1994; Picheny et al., 1985; Tjaden et al., 2014) would appear to support its potential benefit for EL speakers. Yet, in addition to interest in SI, the potential impact of changes secondary to the introduction of CS in EL speakers must also be considered. Therefore, investigation of both SI and composite auditory-perceptual characteristics of EL speech would appear to be warranted. This justification may be of even greater value clinically, especially when the EL has been considered an indispensable mode of alaryngeal speech that must be introduced to all individuals following TL (Doyle, 1994, 2005; Salmon, 1978).

Based on information provided within the preceding review of literature and the potential value of applying CS in the context of EL speech, questions specific to its evaluation by listeners emerge. Thus, the main objective of the series of three investigations to follow was guided by a desire to understand how EL speakers are perceived by normal hearing, naïve listeners when EL speakers are provided with guided instructions to make their speech as “understandable” as possible (i.e., using CS). First, the potential influence of how CS impacts SI in EL speakers were explored. Second, questions regarding the influence of CS on the acoustic characteristics of EL speech were addressed given that such changes may influence SI. Finally, the potential impact of CS on listeners’ auditory-perceptual evaluation of EL speech also warranted

consideration. Collectively, auditory-perceptual evaluation of EL speech by normal-hearing, naïve listeners may provide an ideal means of characterizing differences between the EL speakers while using habitual speech (HS) or CS. It is anticipated that findings from the experimental questions proposed below may identify the potential therapeutic value of CS for those who undergo laryngectomy and use EL speech. Consequently, the following three experimental questions will be addressed:

When compared to habitual EL speech:

- (1) Does CS facilitate improved word intelligibility of EL speakers? (Chapter 2)
- (2) Does CS alter the acoustic characteristics of words and vowels in EL speech? (Chapter 3)
- (3) Does CS result in altered auditory-perceptual ratings by listeners, namely ACC and LC, for EL speakers? (Chapter 4)

References

- American Cancer Society. (2015). *Cancer facts & figures 2015*. Retrieved from <http://www.cancer.org/acs/groups/content/@editorial/documents/document/acspc-044552.pdf>
- American Joint Committee on Cancer (AJCC) (2010). *AJCC cancer staging manual (7th ed.)*. S. Edge, D.R. Byrd., C.C. Compton, A.G. Fritz, F.L. Greene, A. Trotti. (Eds.). New York, NY: Springer-Verlag.
- Bailey, B. J. (1971). Partial laryngectomy and laryngoplasty. *Laryngoscope*, *81*, 1742–1771.
- Barney, H. L., Haworth, F. E., & Dunn, H. K. (1959). An experimental transistorized artificial larynx. *Bell system technical Journal*, *38*(6), 1337-1356.
- Beaudin, P.G. (2002). *A paired-comparison evaluation of electronic artificial laryngeal device preference for vowel stimuli*. Unpublished master's thesis, Western University, Ontario, Canada.
- Bennett, S., & Weinberg, B. (1973). Acceptability ratings of normal, esophageal, and artificial larynx speech. *Journal of Speech and Hearing Research*, *16*, 608-615.
- Berlin, C. I. (1965). Clinical measurement of esophageal speech: III. Performance of non-biased groups. *Journal of Speech and Hearing Disorders*, *30*, 174-182.
- Berry, W. R. (1978). Attitudes of speech language pathologists and otolaryngologists about artificial larynges. In S.J. Salmon and L. P. Goldstein (Eds). *The artificial larynx handbook*. New York: Grune & Stratton.

- Beukelman, D. R., Fager, S., Ullman, C., Hanson, E., & Logemann, J. (2002). The impact of speech supplementation and clear speech on the intelligibility and speaking rate of people with traumatic brain injury. *Journal of Medical Speech-Language Pathology, 10*, 237–242.
- Biacabe, B., Crevier-Buchman, L., Hans, S., Laccourreye, O., and Brasnu, D. (1999). Vocal function after vertical partial laryngectomy with glottic reconstruction by false vocal fold flap: Durational and frequency measures. *Laryngoscope, 109*, 698–704.
- Blood, G. W., Luther, A. R., & Stemple, J. C. (1992). Coping and adjustment in alaryngeal speakers. *American Journal of Speech-Language Pathology, 1*(2), 63-69.
- Blom, E., Singer, M., & Hamaker, R. (1986). A prospective study of tracheoesophageal speech. *Archives of Otolaryngology - Head and Neck Surgery, 112*, 440-447.
- Bornbaum, C. C., Day, A. M., & Doyle, P. C. (2014). Examining the construct validity of the V-RQOL in speakers who use alaryngeal voice. *American Journal of Speech-Language Pathology, 23*(2), 196-202.
- Canadian Cancer Society (2015). *Laryngeal cancer statistics*. Retrieved from: <https://www.cancer.ca/en/cancer-information/cancer-type/laryngeal/statistics/?region=on>
- Clark, J. C., & Stemple, J. C. (1982). Assessment of three modes of alaryngeal speech with a synthetic sentence identification (SSI) task in varying message-to-competition ratios. *Journal of Speech, Language, and Hearing Research, 23*, 333-338.

- Clements, K.S., Rassekh, C.H., Seikaly, H., Hokanson, J.A., and Calhoun, K.H. (1997). Communication after laryngectomy: An assessment of patient satisfaction. *Arch Otolaryngol Head Neck Surg*, 123(5), 493-496.
- Cole, D., Sridharan, S., Moody, M., & Geva, S. (1997). Applications of noise reduction techniques for alaryngeal speech enhancement. *Proceedings of the IEEE Telecon 1997 Region 10 Annual Conference, Speech and Image Technologies for Computing and Telecommunications*, 2, 491-494.
- Cox, S.R., & Doyle, P.C. (2014). The influence of electrolarynx use on postlaryngectomy voice-related quality of life. *Otolaryngology - Head and Neck Surgery*, 150(6), 1005-1009.
- Damste, P. H. (1979). Some obstacles in learning esophageal speech. In R.L. Keith and EL. Darley (Eds.), *Laryngectomy rehabilitation*. San Diego: College-Hill Press.
- Danker, H., Wollbruck, D., Singer, S., Fuchs, Brahler, E., & Meyer, B. (2010). Social withdrawal after laryngectomy. *European Archives of Oto-Rhino-Laryngology*, 267(4), p. 593-600.
- Department of Veterans Affairs Laryngeal Cancer Study Group. (1991). Induction chemotherapy plus radiation compared with surgery plus radiation in patients with advanced laryngeal cancer. *N Engl J Med*, 324(24), 1685-1690.
- DeSanto, L. W., Olsen, K. D., Rohe, D. E., Perry, W. C., & Keith, R. L. (1995). Quality of life after surgical treatment of cancer of the larynx. *Annals of Otolology, Rhinology & Laryngology*, 104(10), 763-769.
- Diedrich, W. M. (1968). The mechanism of esophageal speech. *ANN NY ACAD*, 155(1), 303-317.

- Diedrich, W. M., & Youngstrom, K. A. (1966). *Alaryngeal speech*. Springfield, IL: Charles C. Thomas.
- Doyle, P.C. (in press). *Communication challenges and laryngeal cancer*. In L. Vinnery and J. Frieberg (Eds). *Laryngeal cancer: An interdisciplinary resource for practitioners*. Thorofare, NJ: Slack Publishing.
- Doyle, P.C. (1994). *Foundations of voice and speech rehabilitation following laryngeal cancer*. San Diego: Singular.
- Doyle, P.C. (1999). Postlaryngectomy speech rehabilitation: Contemporary considerations in clinical care. *Journal of Speech-Language Pathology and Audiology*, 23(3), 109-116.
- Doyle, P. C. (2005). Clinical procedures for training use of the electronic artificial larynx. In P. C. Doyle & R. L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing*. Austin, TX: Pro-Ed.
- Doyle, P.C. (2012). Seeking to better understand factors that influence postlaryngectomy speech rehabilitation. *SIG 3 Perspectives on Voice and Voice Disorders*, 22, 45-51.
- Doyle, P.C., Day, A.M.B., Dzioba, A, Bornbaum, C.C. & Sleeth, L. (November, 2011). Relationships between listener comfort & voice severity in tracheoesophageal speech. Paper presented at the Annual American Speech-Language Hearing Association convention, San Diego, CA.

- Doyle, P.C., & Eadie, T.L. (2005). The perceptual nature of alaryngeal voice and speech. In P.C. Doyle, & R.L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing*. Austin, TX: Pro-Ed.
- Duguay, M. J. (1978). Why not both? In S.J. Salmon and L. P. Goldstein (Eds), *The artificial larynx handbook*. New York: Grune & Stratton.
- Eadie, T. L., & Doyle, P. C. (2004). Auditory-perceptual scaling and quality of life in tracheoesophageal speakers. *Laryngoscope*, *111*(4), 753-759.
- Eadie, T. L., & Doyle, P. C. (2005). Quality of life in male tracheoesophageal (TE) speakers. *J Rehabil Res Dev*, *42*(1), 115-124.
- Eadie, T. L., Nicolici, C., Baylor, C., Almand, K., Waugh, P., & Maronian, N. (2007). Effect of experience on judgments of adductor spasmodic dysphonia. *Annals of Otology, Rhinology and Laryngology*, *116*, 695–701.
- Espy-Wilson, C. Y., Chari, V. R., MacAuslan, J. M., Huang, C. B., & Walsh, M. J. (1998). Enhancement of electrolaryngeal speech by adaptive filtering. *Journal of Speech, Language, and Hearing Research*, *41*, 1253–1264.
- Fairbanks, G. (1960). *Voice and articulation drillbook (2nd ed.)*. New York: Harper & Row.
- Finizia, C., Hammerlid, E., Westin, T. and Lindström, J. (1998). Quality of life and voice in patients with laryngeal carcinoma: A posttreatment comparison of laryngectomy (salvage surgery) versus radiotherapy. *Laryngoscope*, *108*, 1566–1573.

- Forastiere, A. A., Goepfert, H., Maor, M., Pajak, T. F., Weber, R., Morrison, W., Glisson, B., ...Cooper, J. (2003). Concurrent chemotherapy and radiotherapy for organ preservation in advanced laryngeal cancer. *N Engl J Med*, 349(22), 2091-2098.
- Gandour, J., & Weinberg, B. (1982). Perception of contrastive stress in alaryngeal speech. *Journal of Phonetics*, 10, 347-359.
- Gandour, J., & Weinberg, B. (1983). Perception of intonational contrasts in alaryngeal speech. *Journal of Speech and Hearing Research*, 26, 142-148.
- Gandour, J., & Weinberg, B. (1984). Production of intonation and contrastive stress in electrolaryngeal speech. *Journal of Speech and Hearing Research*, 27, 605-612.
- Gandour, J., & Weinberg, B. (1985). Production of intonation and contrastive stress in esophageal and tracheoesophageal speech. *Journal of Phonetics*, 13, 313-395.
- Gandour, J., Weinberg, B., & Garziona, B. (1983). Perception of lexical stress in alaryngeal speech. *Journal of Speech and Hearing Research*, 26, 418-424.
- Gandour, J., & Weinberg, B., & Kosowsky, A. (1982). Perception of syntactic stress in alaryngeal speech. *Language and Speech*, 25, 299-304.
- Gandour, J., Weinberg, B., Petty, S. H., & Dardarananda, R. (1986). Rhythm in Thai esophageal speech. *Journal of Speech, Language, and Hearing Research*, 29(4), 563-568.
- Gates, G., Ryan, W., Cooper, J., Lawlis, G., Cantu, E., Hyashi, T.,...Hearne, E. (1982). Current status of laryngectomy rehabilitation: I. Results of therapy. *American Journal of Otolaryngology*, 3, 1-17.
- Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. Englewood Cliffs, NJ: Prentice-Hall.

- Goldstein, L. P., & Rothman, H. B. (1976, November). Analysis of speech produced with an artificial larynx. Presented at the American Speech and Hearing Association Convention, Houston, TX.
- Goldstein, L. P. (1978). Part D: Intelligibility. In S.J. Salmon and L. P. Goldstein (Eds), *The artificial larynx handbook*. New York: Grune & Stratton.
- Hanna, E., Sherman, A., Cash, D., Adams, D., Vural, E., Fan, C. Y., & Suen, J. Y. (2004). Quality of life for patients following total laryngectomy vs chemoradiation for laryngeal preservation. *Arch Otolaryngol Head Neck Surg*, *130*(7), 875-879.
- Hanson, E., Beukelman, D. R., Fager, S., Ullman, C. (2004). Listener attitudes and speaking rate of people with traumatic brain injury. *Journal of Medical Speech-Language Pathology*, *12*(4), 161-166.
- Hillman, R. E., Walsh, M. J., Wolf, G. T., Fisher, S. G., & Hong, W. K. (1998). Functional outcomes following treatment for advanced laryngeal cancer. Part I - Voice preservation in advanced laryngeal cancer. Part II - Laryngectomy rehabilitation: The state of the art in the VA system. Research Speech-Language Pathologists. Department of Veterans Affairs Laryngeal Cancer Study Group. *Annals of Otolaryngology, Rhinology and Laryngology Supplement*, *172*, 1-27.
- Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *J Voice*, *13*(4), 557-569.
- Hoops, H.R., & Noll, J.D. (1969). Relationship of selected acoustic variables to judgments of esophageal speech. *Journal of Communication Disorders*, *2*, 3-13.

- Institute of Electrical and Electronics Engineers (1969). *IEEE recommended practice for speech quality measurements. IEEE Audio and Electroacoustics, 17, 227-246.*
- Kalb, M.B., & Carpenter, M.A. (1981). Individual speaker influence on relative intelligibility of esophageal speech and artificial larynx speech. *Journal of Speech and Hearing Disorders, 46, 77-80.*
- Keith, R. L., & Darley, F. L. (1986). *Laryngectomy rehabilitation*. San Diego, CA: College-Hill Press.
- Keith, R. L., Shanks, J. C., & Doyle, P. C. (2005). Historical highlights: Laryngectomy rehabilitation. In P.C. Doyle and R. L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer; Voice, speech and swallowing*. Autsin, TX: Pro-Ed. pp. 17-58.
- Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility in dysarthria. *Journal of Speech and Hearing Disorders, 54, 482-499.*
- Knox, A. W., & Anneberg, M. (1973). The effects of training in comprehension of electrolaryngeal speech. *Journal of Communication Disorders, 6, 110- 120.*
- Krause, J.C., & Braida , L.D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *Journal of Acoust Soc. Am, 112 (5), 2165-2172.*
- Lauder, E. (1968). The laryngectomy and the artificial larynx. *Journal of Speech and Hearing Disorders, 33(2), 147-157.*

- Laures, J., & Weismer, G. (1999). The effect of flattened F0 on intelligibility at the sentence-level. *Journal of Speech, Language, and Hearing Research*, 42, 1148–1156.
- Lehiste, I. (1976). Suprasegmental features of speech. In N. Lass. (Ed.), *Contemporary issues in experimental phonetics*. New York, NY: Academic Press.
- Lewin, J. (2005). Problems associated with alaryngeal speech. In P. C. Doyle & R. L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing*. Austin, TX: Pro-Ed.
- Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological review*, 79(6), 487-509.
- McCroskey, R. L., & Mulligan, M. (1963). The relative intelligibility of esophageal speech and artificial-larynx speech. *Journal of Speech and Hearing Disorders*, 28(1), 37-41.
- Meltzner, G. S., & Hillman, R.E. (2005). Impact of aberrant acoustic properties on the perception of sound quality in electrolarynx speech. *Journal of Speech, Language, and Hearing Research*, 48(4), 766-769.
- Meltzner, G. S., Kobler, J. B., & Hillman, R. E. (2003). Measuring the neck frequency response function of laryngectomy patients: Implications for the design of electrolarynx devices. *Journal of the Acoustical Society of America*, 114(2), 1035-1047.

- Mendenhall, W. M., Morris, C. G., Stringer, S. P., Amdur, R. J., Hinerman, R. W., Villaret, D.B., & Robbins, K. T. (2002). Voice rehabilitation after total laryngectomy and postoperative radiation therapy. *Journal of Clinical Oncology*, *20*, 2500–2505.
- Merwin, G. E., Goldstein, L. P., & Rothman, H. B. (1985). A comparison of speech using artificial larynx and tracheoesophageal puncture with valve in the same speaker. *Laryngoscope*, *95*, 730-734.
- Moukarbel, R.V., Doyle, P.C., Yoo, J., Franklin, J.H., Fung, K., & Day, A. (2011). Voice-related quality of life (V-RQOL) outcomes in laryngectomees. *Head and Neck*, *33*(1), 31-36.
- Nagle, K. F., Eadie, T. L., Wright, D. R., & Sumida, Y. A. (2012). Effect of fundamental frequency on judgments of electrolaryngeal speech. *American Journal of Speech-Language Pathology*, *21*(2), 154-166.
- National Cancer Institute. (2014). *Laryngeal cancer treatment (PDQ)*. Retrieved from http://www.cancer.gov/cancertopics/pdq/treatment/laryngeal/Patient/page4#_57
- O'Brian, S., Packman, A., Onslow, M., Cream, A., O'Brian, N., & Bastock, K. (2003). Is listener comfort a viable construct in stuttering research? *Journal of Speech, Language, and Hearing Research*, *46*, 503-509.
- Payton, K.L., Uchanski, R.M., & Braida, L.D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America*, *95*, 1581–1592.

- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing. I. Intelligibility differences between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 28, 96–103.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 29, 434–446.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1989). Speaking clearly for the hard of hearing. III. An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 32(3), 600-603.
- Pindzola, R. H., & Moffet, B. (1988). Comparison of ratings of four artificial larynges. *Journal of communication disorders*, 21(6), 459-467.
- Qi, Y. Y., & Weinberg, B. (1991) Low-frequency energy deficit in electrolaryngeal speech. *Journal of Speech, Language, and Hearing Research*, 34(6), 1250-1256.
- Richter, M., Ball, L., Beukelman, D., Lasker, J., & Ullman, C. (2003). Attitudes toward communication modes and message formulation techniques used for storytelling by people with amyotrophic lateral sclerosis. *Augmentative and Alternative Communication*, 19(3), 170-186.
- Rinkel, R.N.P.M., Verdonck-de Leeuw, I. M., van den Brekel, N., de Bree, R., Eerenstein, S. E. J., Aaronson, N., & Leemans, C. R. (2014). Patient-reported symptom questionnaires in laryngeal cancer: Voice, speech and swallowing. *Oral Oncology*, 50(8), 759 - 764

- Robbins, J., Fisher, H.B., Blom, E.C., & Singer, M.I. (1984). A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production. *Journal of Speech and Hearing Disorders*, 49(2), 202–210.
- Robertson, S. M., Yeo, J. C. L., Sabey, L., Young, D. & MacKenzie, K. (2013). Effects of tumor staging and treatment modality on functional outcome and quality of life after treatment for laryngeal cancer. *Head Neck*, 35, 1759–1763.
- Rothman, H. B. (1978). *Analyzing artificial electronic larynx speech*. In S. J. Salmon and L. P. Goldstein. (Eds.), *The artificial larynx handbook*. New York, NY: Grune. pp. 87-111.
- Rothman, H. B. (1982). *Acoustic analysis of artificial electronic larynx speech*. In A. Sekey (Ed.), *Electrical analysis and enhancement of alaryngeal speech*. Springfield, IL: Charles C. Thomas.
- Saikachi, Y., Stevens, K. N., & Hillman, R. E. (2009). Development and perceptual evaluation of amplitude-based F0 control in electrolarynx speech. *Journal of Speech, Language, and Hearing Research*, 52(5), 1360-1369.
- Salmon, S. J. (1978). *Acoustic analysis of artificial electronic larynx speech*. In S. J. Salmon and L. P. Goldstein. (Eds.), *The artificial larynx handbook*. New York, NY: Grune. pp. 145-147.
- Salmon, S. J. (2005). Commonalities among alaryngeal speech methods. In P.C. Doyle and R. L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing* (pp. 59-74). Austin, TX: Pro-ed.

- Salmon, S.J., & Goldstein, L. P. (Eds.) (1978). *The artificial larynx handbook*. New York, NY: Grune.
- Schiavetti, N. (1992). *Scaling procedures for the management of speech intelligibility*. In R. Kent (Ed.) *Intelligibility in speech disorders: Theory, measurement, and management*. Philadelphia, PA: John Benjamins Publishing Co.
- Shames, G. H., Font, J., & Matthews, J. (1963) Factors related to speech proficiency of the laryngectomized. *Journal of Speech and Hearing Disorders*, 28, 273-287.
- Shipp, T. (1967) Frequency, duration and perceptual measures in relation to judgments of alaryngeal speech acceptability. *Journal of Speech, Language, and Hearing Research.*, 10, 417-427.
- Siegel, R. L., Miller, K. D., & Jemal, A. (2015). Cancer statistics, 2015. *CA: a cancer journal for clinicians*, 65(1), 5-29.
- Silver, C.E., Beitler, J.J., Shaha, A.R., Rinaldo, A., & Ferlito, A. (2009). Current trends in initial management of laryngeal cancer: the declining use of open surgery, *Oto-Rhino-Laryngology*, 266(9), 1333-1352.
- Silver, C.E., & Ferlito, A. (1996). *Surgery for cancer of the larynx and related structures* (2nd ed.). Philadelphia, PA: Saunders.
- Singer, M. I., & Blom, E. D. (1980). An endoscopic technique for restoration of voice after laryngectomy. *Annals of Otolaryngology, Rhinology, and Laryngology*, 89, 529–533.
- Skandera, P., & Burleigh, P. (2005). *A manual of English phonetics and phonology: twelve lessons with an integrated course in phonetic transcription*. Gunter Narr Verlag.

- Smiljanić, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Lang Linguist Compass*, 3(1), 236-264.
- Snidecor, J. C. (1968). Speech rehabilitation of the laryngectomized. Springfield, IL: Thomas.
- Snidecor, J. C., & Curry, E.T. (1959). Temporal and pitch aspects of superior esophageal speech. *Ann Otol Rhinol Laryngol*, 68, 623-629.
- Stewart, M.G., Chen, A.Y., Stach, C.B. (1998). Outcomes analysis of voice and quality of life in patients with laryngeal cancer. *Arch Otolaryngol Head Neck Surg*, 124, 43–148.
- Surveillance, Epidemiology, and End Results (SEER) Program. (2014). *SEER cancer statistics factsheets: Larynx cancer*. Bethesda, MD: National Cancer Institute. Retrieved from <http://seer.cancer.gov/statfacts/html/laryn.html>.
- Susca, M., & Healey, E. C. (2001). Perceptions of simulated stuttering and fluency. *Journal of Speech, Language, and Hearing Research*, 44, 61–72.
- Susca, M., & Healey, E. C. (2002). Listener perceptions along a fluency-disfluency continuum: A phenomenological analysis. *Journal of Fluency Disorders*, 27, 135–161.
- Terrell, J. E., Fisher, S. G., & Wolf, G. T. (1998). Long-term quality of life after treatment of laryngeal cancer. *Archives of Otolaryngology–Head & Neck Surgery*, 124(9), 964-971.

- Timmermans, A. J., de Gooijer, C. J., Hamming-Vrieze, O., Hilgers, F. J.M. and van den Brekel, M. W.M. (2014). T3-T4 laryngeal cancer in The Netherlands Cancer Institute; 10-year results of the consistent application of an organ-preserving/-sacrificing protocol. *Head Neck*. doi: 10.1002/hed.23789.
- Tjaden, K., Sussman, J.E., & Wilding, G. E. (2014). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research*, 57, 779-792.
- Tolhurst, G. (1957). Effects of duration and articulation changes on intelligibility, word reception and listener preference. *Journal of Speech and Hearing Disorders*, 22, 328-334.
- Trudeau, M. D., & Qi, Y. Y. (1990). Acoustic characteristics of female tracheoesophageal speech. *Journal of Speech and Hearing Disorders*, 55, 244–250.
- Uchanski, R. M. (2005). Clear speech. In D. B. Pisoni and R. E. Remez (Eds.), *The handbook of speech perception*. Malden, MA: Blackwell.
- Union for International Cancer Control (UICC). (2009). *TNM classification of malignant tumours (7th Ed.)*. L. H. Sobin , M. K. Gospodarowicz, C. Wittekind. (Eds.). West Sussex, UK: Wiley-Blackwell.
- Vilaseca, I., Chen, A. Y., & Bakscheider, A. G. (2005). Long-term quality of life after laryngectomy. *Head and Neck*, 28(4), 313-320.

- Ward, E. C., Koh, S. K., Frisby, J., & Hodge, R. (2003). Differential modes of alaryngeal communications and long-term voice outcomes following pharyngolaryngectomy and laryngectomy. *Folia Phoniatrica et Logopaedica*, 55 (1), 39-49.
- Watson, P. J., & Williams, S. E. (1987). Laryngectomees' and nonlaryngectomees' perceptions of three methods of alaryngeal voicing. *J Comm Disord.*, 20, 295-304.
- Watson, P. J., & Schlauch, R. S. (2009). Fundamental frequency variation with an electrolarynx improves speech understanding: A case study. *American Journal of Speech-Language Pathology*, 18(2), 162–167.
- Weinberg, B., & Gandour, J. (1986). Prosody in alaryngeal speech. *Seminars in Speech and Language*, 7, 95–107.
- Weiss, M S., & Basili, A. M. (1985). Electrolaryngeal speech produced by laryngectomized subjects: Perceptual characteristics. *Journal of Speech and Hearing Research*, 28, 294-300.
- Weiss, M. S., Yeni-Komshian, G. H., & Heinz, J. M. (1979). Acoustical and perceptual characteristics of speech produced with an electronic artificial larynx. *The Journal of the Acoustical Society of America*, 65(5), 1298-1308.
- Williams, S. E., & Watson, J. B. (1985). Differences in speaking proficiencies in three laryngectomee groups. *Archives of Otolaryngology—Head & Neck Surgery*, 111(4), 216.
- Williams, S. E., & Watson, J. B. (1987). Speaking proficiency variations according to method of alaryngeal voicing. *The Laryngoscope*, 97(6), 737-739.

Wong, J.H.C. (2003). *An experimental evaluation of speech intelligibility using a prototype electrolaryngeal device*. Unpublished master's thesis, Western University, Ontario, Canada.

World Health Organization. (2012). *The World Health Organization Quality of Life (WHQOL)*. Retrieved from:
http://www.who.int/mental_health/publications/whoqol/en/

Chapter 2

The Impact of Clear Speech on Word Intelligibility of Electrolaryngeal Speakers

Current evidence-based practice guidelines in speech-language pathology indicate that clinicians must provide communication options for individuals who seek voice and speech rehabilitation following head and neck cancer treatment (American Speech-Language-Hearing Association, 2015; Royal College of Speech-Language Therapists, 2005). Common communication options following total laryngectomy (TL) include esophageal speech (ES), tracheoesophageal puncture (TEP) voice restoration, and the use of the electronic artificial larynx, or what is more commonly referred to as the electrolarynx (EL)³. While all three alaryngeal speech modes may provide an effective means of postlaryngectomy verbal communication, they vary considerably when compared to normal laryngeal speech, particularly relative to speech intelligibility (SI).

SI refers to how well a speaker's message is understood by a listener (Kent, Weismer, Kent, & Rosenbek, 1989). Accordingly, comprehensive assessment of SI should focus on its component parts: the speaker, the method of transmission, and the listener (Schiavetti, 1992). In this regard, individuals who use any alaryngeal communication method provide a unique clinical population for SI research. First, those who undergo TL lose their primary voicing source (i.e., larynx), and consequently, must attempt to regain functional verbal communication. Second, alaryngeal speech may rely on either intrinsic biological sources of vibration (i.e., the pharyngoesophageal segment)

³ It is beyond the scope of this work to provide detailed descriptions for each postlaryngectomy communication option. Therefore, the reader is referred to texts by Doyle (1994) and Doyle and Keith (2005).

for ES and TEP speech, or rely on the use of an extrinsic source of voicing for EL speech. Each of these methods possesses unique acoustic and auditory-perceptual features that directly impact communication. For example, EL speakers often have difficulty communicating over the telephone and in certain levels of environmental noise due to deficits in the frequency and intensity of the voice (Qi & Weinberg, 1991; Saikachi, Stevens, & Hillman, 2009; Verdolini, Skinner, Patton, & Walker, 1985). Such acoustic features serve to explain in part why ES and TE speech have been reported to be relatively more intelligible than EL speech (Barney, Haworth, & Dunn, 1959; Eadie et al., in press; Kalb & Carpenter, 1981; McCroskey & Mulligan, 1963; Shames, Font, & Matthews, 1963; Williams & Watson, 1985). While research has shown that EL speech presents difficulties to listeners in various communication contexts, attempts to improve the SI of EL speech remain.

The EL is a hand-held, battery operated device that is most often placed against the neck (transcervical or transcutaneous), although the speech signal also can be introduced into the mouth (transoral or intraoral). Research has indicated that approximately 50% of individuals use an EL at two years postlaryngectomy (Hillman, Walsh, Wolf, Fisher, & Hong, 1998; Mendenhall et al., 2002; Ward, Koh, Frisby, & Hodge, 2003). Even when the EL is not a primary mode of alaryngeal communication, it is a reliable back-up mode of alaryngeal speech (Doyle, 1994; Hillman et al., 1998). Barney et al. (1959) were the first to provide SI data on the neck-type EL in comparison to normal and ES using phonetically-balanced word stimuli (Egan, 1948). Barney et al. (1959) found that listeners correctly transcribed 58.1% of words spoken with the EL when compared to normal (97.3%) and ES speakers (79%). Research has consistently

shown that EL speech is less intelligible than both ES and TE speech. General findings indicate that the SI of EL speakers can range from 35.5% to 60.3% for words (Bennett & Weinberg, 1973; Shames et al., 1963; McCroskey & Mulligan, 1963; Weiss, Yeni-Komshian, & Heinz, 1979). A portion of listeners' errors are directly related to voicing confusions for consonants (Weiss & Basili, 1985; Weiss et al., 1979). For example, word-initial (WI) voiceless plosives tend to exhibit reduced intelligibility due to the constantly voiced nature of the EL source. This often results in listeners mistaking a voiceless stop as its voiced cognate (Weiss & Basili, 1985; Weiss et al., 1979). Yet, other factors can also influence SI of EL speakers.

Meltzner and Hillman (2005) acknowledged that previous research has often addressed EL signal deficits in isolation, and therefore, has not yielded a collective approach to improve SI for EL speakers. For example, the lack of low frequency energy below 500Hz in EL speech has been suggested to contribute to an inferior and artificial sound quality that may impact the noise spectra of consonants (Qi & Weinberg, 1991; Weiss et al., 1979). This is an important consideration, considering data from Black (1946) demonstrated that more acoustical power is required to generate increased intensity at lower frequencies. Another example of a prominent acoustic difficulty associated with EL speech is related to the simultaneous noise that radiates from the device into the communication environment, which then competes with the speech signal. Attempts to remove this noise were conducted by Espy-Wilson, Chari, Huang, and Walsh (1998) who compared an unmodified EL speech signal to one that was filtered to remove noise. Although naïve listeners and laryngectomees preferred the filtered EL speech signal, no significant differences in SI were reported between these two signals (Espy-

Wilson et al., 1998). In a more recent investigation, Watson and Schlauch (2009) examined the effect of EL frequency variation on SI after an EL speaker read sentences with and without variable frequency control, and an increase in intelligibility of at least 10% was observed.

In response to studies that focused on improving specific acoustic deficits of EL speech, Meltzner and Hillman (2005) identified that a combination of low-frequency enhancement, a reduction in device noise, and the ability of speakers to vary frequency contribute to the best overall voice quality ratings by listeners. This research was supported by Beaudin (2002), who indicated that acoustically modifying the EL signal can lead to improved listener preference. A follow-up study by Wong (2003), however, found that these voice quality improvements occurred at the expense of reduced SI. Specifically, Wong (2003) found that modified EL devices using adaptive filtering techniques had a negative impact on SI. While direct modification of the EL source may benefit listener judgments of signal quality, the negative influence on SI that occurs poses an ongoing challenge. Thus, the present study sought to explore potential changes in the SI of EL speakers through application of a therapeutic modification termed clear speech (CS).

CS was first introduced with the purpose of improving communication over radio and telecommunication systems (Picheny, Durlach, & Braida, 1985). Briefly, CS is a style of speaking that attempts to improve the understandability of a speaker's message (Picheny et al., 1985; Krause & Braida, 2002; Smiljanić & Bradlow, 2009; Uchanski, 2005). CS attempts to improve the understanding of speech by the listener through the speaker's deliberate use of a slower speech rate, increased speaking volume, and over-

articulation involving increased mouth opening (Picheny et al, 1985). Since its introduction, CS has been used in an effort to improve speech production and understandability in individuals with communication disorders including dysarthria secondary to traumatic brain injury (Beukelman, Fager, Ullman, Hanson, & Logemann, 2002), Multiple Sclerosis and Parkinson's disease (Tjaden, Sussman, & Wilding, 2014), in addition to those with hearing impairment (Picheny et al., 1985; Picheny, Durlach, & Braida, 1986). When compared to typical conversational speech, CS has consistently been shown to improve SI for the clinical populations noted with reports indicating an improvement in SI of up to 26% (Payton, Uchanski, & Braida, 1994; Picheny et al., 1985). This increase has been partly attributed to the slowed rate *and* over-articulation that evolves from CS. However, while the application of CS has been shown to be effective for improving SI for individuals with an array of speech and hearing disorders, no reports of the application of CS to alaryngeal speakers has been pursued. Given that EL speech consistently has been found to be less intelligible than other alaryngeal speech methods, and that previous attempts to experimentally control acoustic aspects of EL speech have failed to improve SI, research into the application of CS in EL speakers would appear to be warranted.

Coincidentally, the production aspects of the CS method are part of the instructions provided to all laryngectomees when they begin to use an EL device (Doyle, 1994; 2005). In this regard, CS seems well-suited for improving SI of EL speakers. For the EL speaker, a wider mouth opening secondary to over-articulation may permit greater vibratory sound energy from the EL device to be resonated within the oral cavity. With more sound energy in the oral cavity, alongside the conscious effort to over-articulate

each speech sound and reducing one's rate of speech, improvements in a SI may emerge. In addition, CS might result in improvements in the production of consonant features such as voicing and manner, which in turn may lead to improved SI of specific sounds and words. Therefore, the purpose of this study sought to determine the impact of CS on the SI of words and to assess consonant SI by phonetic position for EL speech.

Method

Participant Speakers

Ten adult men who had undergone TL and who used EL speech as their primary method of communication served as speakers for this study. All reported to be native English speakers. Participants ranged in age from 59 to 87 years ($M_{age} = 74$ years). All speakers were self-reported to be in good general health at the time of the study with no known neurological, medical or psychological conditions. This included self-reports of no known hearing difficulties. However, given the age and medical treatment related to laryngeal cancer, some level of hearing loss cannot be ruled out.

Time using an EL device postlaryngectomy was reported to range from 24 to 300 months ($M_{time} = 133$ months). Seven speakers had a neck dissection as part of their TL. Speakers received radiation therapy (RT) either before ($n=4$), after ($n=5$), or both before and after laryngectomy ($n=1$). Two speakers received combined chemoradiotherapy, one prior to and one after surgery. As part of their participation, each speaker was asked to bring their own EL device to the experimental recording session; this included an equal representation of five individuals who used the Servox Digital EL (Servona GmbH, Troisdorf, Germany) and five who used the TruTone (Griffin Laboratories, Temecula, CA) device.

Speech Stimuli

The stimuli used in this investigation were comprised of 17 monosyllabic consonant-vowel-consonant (CVC) and one consonant-vowel (CV) English words selected from a larger 66-item word list first described by Weiss and Basili (1985) (see Appendix A). The goal of stimuli selection was to ensure equal representation of consonants in both WI and word-final (WF) positions. Specifically, these stimuli included three sets of cognate pairs including six plosives (/p/, /b/, /t/, /d/, /k/, and /g/), seven fricatives (/f/, /v/, /s/, /z/, /ʃ/, /θ/, and /ð/), two affricates (/tʃ/ and /dʒ/), and two nasal consonants (/m/ and /n/). Of the 18 stimulus words, 16 words represented target consonants in both word initial (WI) and word final (WF) positions. However, two additional words (i.e., ‘know’ and ‘loathe’) were included to represent the WI nasal (/n/) and the WF voiced fricative (ð).

Acquisition of Speech Stimuli

All recordings were gathered in a quiet room free of background noise as judged by a v. Recording of speaker stimuli occurred after informed consent was obtained from all speakers (Western University Research Ethics Board Approval #105382) (see Appendices B and C). Demographic information and a brief medical history also were obtained from each participant in advance of recording (see Appendix D). A microphone (Shure PG-81, Niles, IL) attached to a desktop microphone stand was placed approximately 15cm above each participant speaker and directed at each speaker’s mouth at a 45 degree angle. All speaker stimuli were recorded onto a laptop computer (Dell, Round Rock, TX) at a sampling rate of 44.1kHz using the SonaSpeech II software employing the Multidimensional Voice Profile application (Kay Pentax, Lincoln Park, NJ). Volume input

levels were adjusted for each speaker at the beginning of each session and were monitored during the recordings using a volume unit (VU) metre in SonaSpeech II to avoid any under- or over-driving of the input signal.

To begin each recording session, participant speakers were provided with a printed copy of the Rainbow Passage (Fairbanks, 1960) and given the following verbal instructions: “Please take a moment to look over the following paragraph. Once you are ready, please read it aloud. If you make a mistake, I will ask you to repeat the sentence(s) once you finish reading”. Once each participant speaker finished reading, they were provided with a printed copy of the 18-item word list and the following instructions: “Please take a moment to look over the words. Once you are ready, please read each word. If you make a mistake, I will ask you to repeat the word(s) once you finish reading”.

Once the HS recording task was completed, the investigator provided each participant with instructions on how to produce clear speech (CS) for the second phase of the recording procedure. Similar to the instructions used by Picheny et al. (1985), participants were asked, “Now I would like you to re-read the words and the reading passage by speaking as clearly as possible. This will involve slowing down while speaking and over-articulating” (Picheny et al., 1985). Each speaker quietly reading stimuli using this style of speaking prior to recording. Participant speakers always began the recording session in the HS condition, followed by the CS condition. This order was used to control for any carryover effects from the experimental speaking condition (i.e., CS) had that been recorded first. All recording sessions lasted approximately 20 minutes.

Editing speech stimuli. After all recordings were completed, 36 audio files containing words (18 HS and 18 CS audio files) were edited using Audacity 2.0.5

(Mazzoni & Dannenberg, 2013). Audible recording noise on each audio file was removed using the ‘Noise Removal’ tool within Audacity. A small window was highlighted at the beginning of each audio file (e.g., not involving speech stimuli) to capture a profile of track noise. The track noise was analyzed and then removed, while leaving speech stimuli unaltered in the process. Finally, to extract each stimulus, words were highlighted, copied, and pasted into new audio tracks and saved as individual audio files in .wav format.

Across all 10 speakers, there were a total of 360 words [18 words x 10 speakers x 2 speaking conditions]. Additionally, 60 words (~16.67%) were randomly selected and duplicated and then included in the master stimuli lists; these additional samples served as reliability samples for the evaluation of the listeners’ auditory-perceptual judgments. Therefore, each listener was presented with a series of 420 experimental stimuli. In addition to primary stimuli, five running speech samples were selected in advance for presentation to listeners at the start of the formal session. These five samples were included to limit the potential for a naïve listener to be surprised by the unusual nature of the EL signal and, consequently, potentially disrupt their level of attention to the primary samples that they would be requested to transcribe. Finally, each stimulus list was presented to listeners in a unique, randomized presentation using software created specifically for this project (Failla, 2014).

Evaluation of Intelligibility

Participant Listeners

Twelve adult women ranging in age from 21;0 to 29;09 years ($M_{\text{age}} = 23;09$ years) served as participant listeners for this study. All participants were undergraduate or

graduate students who responded to class announcements or postings regarding the study. All were self-reported to be in good health and indicated that they did not have any history of speech, voice, language, and/or hearing difficulties, and all were native English speakers. Listeners were not reimbursed for their time or participation.

Participant listeners were considered to be naïve after indicating that they had no formal training in and/or experience with voice or speech disorders, especially postlaryngectomy ‘alaryngeal’ speech. Research has suggested that naïve listeners are able to provide important data related to the general populations’ assessment and perception of individuals with voice disorders, including alaryngeal speakers (Eadie & Doyle, 2004; Kreiman, Gerratt, Kempster, Erman, & Berke, 1993; Tardy-Mitzell, Andrews, & Bowman, 1985). Further, laryngectomees are more likely to encounter individuals who lack an understanding or prior exposure to alaryngeal speech (Eadie & Doyle, 2004; Tardy-Mitzell et al., 1985). Lastly, research indicates that naïve listeners are able to provide similar judgments related to speech rate and SI of EL speakers as expert listeners (Watson & Williams, 1987). Therefore, naïve listeners were deemed appropriate to understand the effect of CS on EL speakers.

Listening Procedure

Each listener participated in a single listening session within the Voice Production and Perception Laboratory at Western University. At the beginning of each session, listeners were provided with a letter of information for the study and any questions they may have had were answered, and informed consent was obtained (Western University Research Ethics Board Approval #105884) (see Appendices E and F). Each participant was then seated in front of a desktop computer (Dell, Round Rock,

TX) within a listening laboratory free of ambient noise and provided with stereo headphones (Sony MDRV-150). Prior to the formal word transcription task, all listeners were first presented with the five initial “exposure” samples of EL speech noted previously with the knowledge that these samples were presented in order to briefly familiarize them with the types of stimuli that would follow. Upon completing the presentation of these exposure samples, the principal investigator then opened a master experimental list located in a single Microsoft Office document with a randomized list of the 420 word stimuli; listeners were also provided with a printed copy of a document that represented the exact information they viewed on the computer in order to directly record their perceptual responses. The following instructions were provided to listeners prior to beginning the transcription task:

You are about to hear a series of words. Please write the word you hear in the space provided on the score sheet provided to you. If you cannot understand the word, please draw a line through the space for that word.

Each participant listener began the task by clicking on a computer icon that identified the exposure file, and once completed, they proceeded to listen to and make their perceptual judgments of each sample by clicking on individual stimulus icons. Listeners were allowed to listen to individual items as many times as they desired prior to providing their response, but they were instructed to not change their transcriptions once it was written on their score sheets. Additionally, listeners were instructed to not return to any prior sample, but to continue sequentially through the randomized list until all judgments were completed. The primary investigator remained in the testing area to monitor participant progress and answered questions if further clarification was required. Immediately after each listener completed the transcription task, the researcher reviewed the data sheets for

any misspellings. In the event of a misspelling (e.g., 'lothe' or 'loath' for 'loathe'), the researcher asked the participant to confirm the intended word. Overall, individual listening sessions required an average of 81 minutes (*range* = 55-113 minutes) with the entire task completed in a single session.

Data Analyses

Listener transcriptions for all words were scored by an independent transcriber in two ways. First, a word SI score was calculated by dividing the number of correctly identified words by the total number of words presented. For the second analysis, the independent transcriber considered transcription errors specific to WI and WF consonants. These data were then used to generate individual listener confusion matrices for both WI and WF consonants for each of the 10 speakers. Thus, both whole word and consonant scores by word position were generated for each individual listener and speaker. Finally, individual speaker matrices were collapsed across the group of speakers into a master confusion matrix for both WI and WF consonants.

Statistical Analyses

Word-level analyses. A repeated measures analysis of variance (ANOVA) was used to assess the influence of speaking conditions on word SI. Post-hoc testing with a Bonferroni correction was used to compare overall word SI for each speaking condition (e.g., HS vs. CS). This was followed by comparisons of word SI scores within each device group according to speaking condition (e.g., Servox Digital users' HS vs. Servox Digital users' CS, TruTone users' HS vs. TruTone users' CS), and then between device groups and speaking conditions (e.g., Servox Digital HS vs. TruTone HS, etc.). The magnitude of effect for speaking condition was determined by calculating Partial Eta

Squared. Interpretation of effect size followed guidelines by Cohen (1988), including 0.01 for a small effect, 0.06 for a medium effect, and 0.14 for a large effect. An a priori significance level was set at $p < .05$ for all statistical analyses.

Word-position analyses. Analyses of consonant voicing, manner, and omissions in WI and WF stimuli were conducted using a repeated measures ANOVA. Similar to word-level analyses, assessment of device grouping and speaking condition was conducted for voicing, manner, and omissions for WI and WF positions (e.g., overall WI/WF scores, Servox WI/WF scores in HS vs. Servox WI/WF scores in CS, Servox WI/WF scores in HS vs. TruTone WI/WF scores in HS, etc). Similar to word-level analyses, effect size was determined by calculating Partial Eta Squared and interpreted according to Cohen (1988). A Bonferroni correction was used for post-hoc testing and an a priori significance level was set at $p < .05$ for all statistical analyses.

Relationships between SI scores. Assessments of the degree of relationship in speaker performance between the HS and CS conditions were also undertaken using Pearson product-moment correlations.

Agreement and Reliability. Measures of agreement and reliability based on each listener's responses to the 60 duplicated stimulus words were determined for whole words, as well as for WI and WF consonants. This measure was based on the consistency of a listener's response to the first presentation of a stimulus item to that of the second duplicate sample of the same item. Thus, regardless of whether the response at any comparative level (word and WI or WF position) was correct or incorrect, agreement served to index the overall consistency of the listener's response. In total, agreement within listeners ranged from 55% to 83% ($M = 70%$) for whole-word stimuli, and from

67% to 92% ($M = 80\%$) for WI consonants and 67% to 90% ($M = 80\%$) for WF consonants. Intra-class correlation coefficients (ICC) (Shrout & Fleiss, 1979) were used to analyze inter-rater reliability. The average group ICC was 0.981. Therefore, intra- and inter-rater reliability were in agreement for judgments of SI across words, WI phonemes and WF phonemes.

Results

Word Intelligibility

Word SI scores for the group of listeners were based on 2,160 perceptual ratings in each speaking condition (18 words x 10 speakers x 12 listeners). Individual speaker word scores (raw and percentages) were grouped according to EL device used are shown in Table 2.1. A mean word intelligibility score of 51.7% ($Mdn = 55.3\%$, $range = 29.2-69.9\%$) was observed for HS and 53.0% ($Mdn = 57.4\%$; $range = 29.2-67.1\%$) for CS. When raw listener data are collapsed across speakers, remarkable similarities can be noted between scores in the HS condition ($SD = 29.3$; $range = 63-151$) and the CS condition ($SD = 28.8$; $range = 62-145$). Thus, overall word scores across the two experimental conditions differed by only 1.3%. Results from a repeated measures ANOVA indicated that there was no significant effect of speaking condition on word SI.

Table 2.1*Individual Speaker Raw and Percentage Scores for Overall Words*

Speaker	Habitual Speech		Clear Speech	
	N*	%	N*	%
1	151	69.9	133	61.6
2	96	44.4	108	50.0
3	148	68.5	145	67.1
4	119	55.1	103	47.7
5	63	29.2	69	31.9
Servox Total	577**	53.4	558**	51.7
6	126	58.3	119	55.1
7	70	32.4	63	29.2
8	102	47.2	137	63.4
9	120	55.6	138	63.9
10	121	56.0	129	59.7
TruTone Total	539**	49.9	586**	54.3
Overall	1116***	51.7	1144***	53.0
*216 words for each speaker				
**1,080 words for each device group				
***2,160 words in each speaking condition				

Comparison of Word Intelligibility by Device Group

Table 2.1 also provides a comparison of word SI scores between speaking condition and the EL device used. Servox Digital users had a mean word SI score of 53.4% (*Mdn* = 55.1%; *range* = 29.2-69.9%) in HS and a mean score of 51.7% (*Mdn* = 50%; *range* = 31.9-67.1%) in CS. TruTone users had a word SI score of 49.9% (*Mdn* = 55.6%; *range* = 32.4-58.3%) in HS and a mean score of 54.3% (*Mdn* = 59.7%, *range* = 29.2-63.9%) in CS. These data indicate that Servox users had a word score that was 3.5% greater than those who used the TruTone during HS. However, results from the repeated measures ANOVA indicated that there was no significant effect of EL device on word SI score in HS. For CS, TruTone users achieved a word score that was 2.6% greater than the Servox users. However, the repeated measures ANOVA indicated that there was no significant effect of EL device on word SI score in CS.

Relationship Between Speaking Conditions

The relationship between word SI scores in HS and CS is illustrated in Figure 2.1. Overall, there was a strong, statistically significant correlation between word SI scores in HS and CS scores, $r = 0.842$, $p < .01$), thus, accounting for slightly more than 70% of the variance.

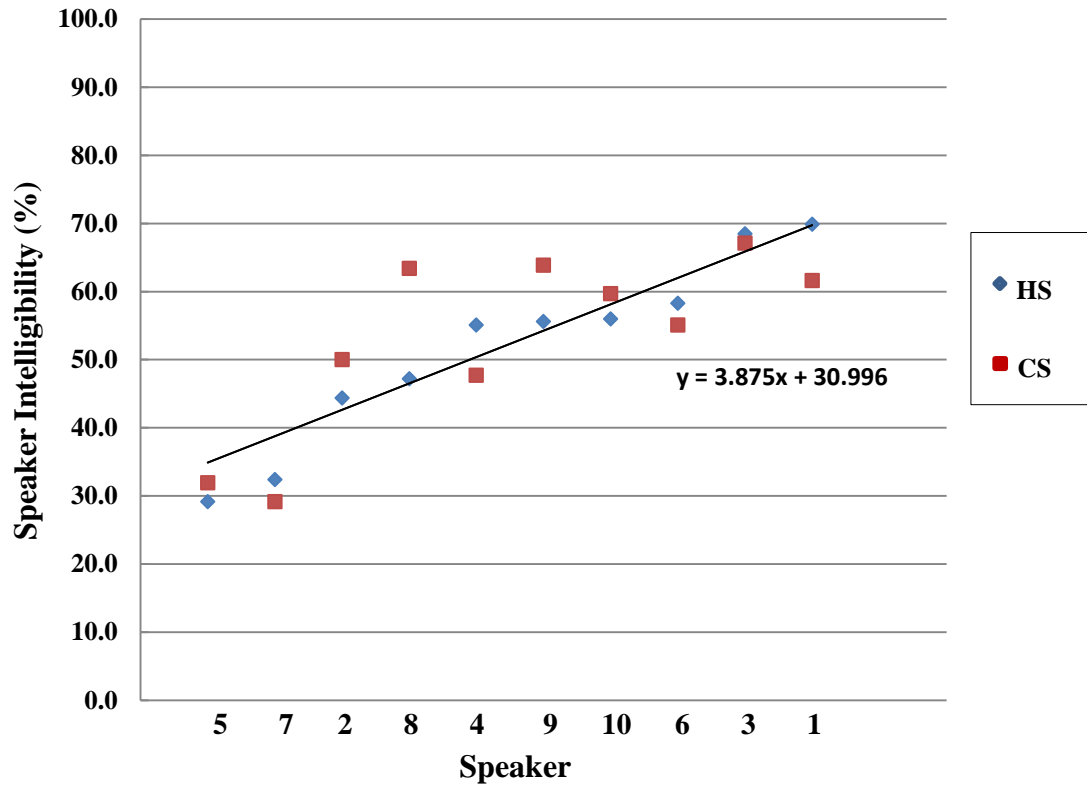


Figure 2.1. Relationship between individual speaker intelligibility in habitual speech (HS) and clear speech (CS). Speaker intelligibility is arranged from lowest to highest.

Intelligibility by Consonant Position: WI and WF

WI and WF position data are summarized in the confusion matrices shown in Tables 2.2a and 2.2b, respectively. In total, word-position SI scores (WI and WF) were based on 2,040 perceptual ratings in each word-position (17 consonants x 10 speakers x 12 listeners). Individual speaker scores by position (raw and percentages) were grouped by EL device and summarized in Table 2.3.

WI position. In total, 1541 out of 2,040 consonants were correctly identified ($SD = 16.5$; $range = 122-169$) in HS and 1573 out of 2,040 consonants were correctly identified in the CS condition ($SD = 16.8$; $range = 126-176$). Thus, the overall WI SI score of 75.5% ($Mdn = 78.4\%$; $range = 59.8-82.8\%$) was observed for HS and 77.1% ($Mdn = 78.9\%$; $range = 61.8-85.8\%$) for CS. Results from a repeated measures ANOVA indicated that there was a significant effect of speaking condition on consonant scores in WI position, $F(1,8) = 6.954$, $p < .05$, $partial \eta^2 = .465$. Further, the magnitude of the effect revealed that speaking condition had a large effect on WI consonant SI (Cohen, 1988). Post-hoc testing indicated that SI scores in WI position were significantly greater when EL speakers used CS compared to HS ($p < .05$).

WF position. For the HS condition, 1656 out of 2,040 consonants were correctly identified ($SD = 23.9$; $range = 127-198$) compared to 1674 out of 2,040 consonants being correctly identified during CS ($SD = 23.1$; $range = 125-194$). Thus, an overall WF consonant SI score of 81.2% ($Mdn = 84.8\%$; $range = 62.3-97.1\%$) was observed for HS and a score of 82.1% ($Mdn = 86.5\%$; $range = 61.3-95.1\%$) was noted for CS. Results from the repeated measures ANOVA indicated no significant effect of speaking condition on consonant SI scores in WF position.

Table 2.2a

Overall Perceptual Confusion Matrix for Word-Initial Consonants Spoken with Habitual Speech/Clear Speech

Response Target	p	b	t	d	k	g	f	v	θ	ð	s	z	ʃ	tʃ	dʒ	m	n	ŋ	l	r	w	j	h	NR	Total
p	3/5	87/91		0/1			3/0	1/1		6/1						11/18				0/2				9/1	120/120
b	0/1	88/99		0/1	1/0		2/0	3/0								19/16	0/1						1/0	6/2	120/120
t	0/1	1/2	79/84	12/13	1/1	1/2	1/0			3/1	4/2		3/1	1/2		0/2	4/5		1/1					9/3	120/120
d		4/2		84/83		0/5				17/16						9/4	6/9							0/1	120/120
k		13/15		4/2	85/86	15/10	1/1	0/1		1/0					1/1	0/1								2/3	120/120
g	2/0		0/1	5/3	1/1	108/114				1/0			0/1	1/0	1/0	1/0	1/0								120/120
f		4/5		2/3		1/0	93/82	3/12		3/10	2/0					5/5	1/0		0/1					6/2	120/120
v	0/2	30/27		5/5		0/1	10/14	20/23		10/13	0/2	0/1				19/21	1/0		20/7		1/4			4/0	120/120
θ		3/1	3/3	3/4		4/3	6/12	4/4	87/71	1/2	2/2					0/3	1/0		2/2					4/3	120/120
ð	5/0	0/2	0/1	2/4			2/1			97/96							5/4		6/12					3/0	120/120
s		6/4						2/0	4/4	1/2	83/81	14/18	1/0	2/1			3/2		1/0		1/0		1/0	2/8	120/120
z		2/3		2/1			2/2	1/1	1/0	3/9	63/57	30/40	0/1		1/0	1/1			8/2					6/3	120/120
ʃ		1/3		1/2		13/0				2/4			86/103	5/0	9/6		0/1							3/1	120/120
tʃ			3/4		4/1	5/0							17/8	77/92	8/11	0/1								6/3	120/120
dʒ		0/3	6/4	8/12	64/55	26/34	1/0	1/0	1/2	1/0	2/2			1/0	3/5		1/0		1/1		1/0	1/0	2/2	120/120	
m																119/115	1/4							0/1	120/120
n																									120/120
Total	10/9	237/257	91/97	128/134	156/144	173/169	121/112	35/42	93/77	146/154	156/156	44/59	107/113	86/96	23/23	184/187	144/146	0/0	39/26	0/2	1/4	2/0	2/0	62/33	2040/2040

Note. NR = No response.

Table 2.2b

Overall Perceptual Confusion Matrix for Word-Final Consonants Spoken with Habitual Speech/Clear Speech

Response Target	p	b	t	d	k	g	f	v	θ	ð	s	z	ʃ	ʒ	dʒ	m	n	ŋ	l	r	w	j	h	NR	Total
p	81/ 78	3/4		1/0				7/ 11	4/3							15/ 19	3/3	1/0		0/1				5/1	120/120
b	2/3	84/ 81		9/5		1/1		0/2								21/ 24	3/1							0/3	120/120
t	0/2	4/4	64/ 72	17/ 15	4/3	2/6			2/0				0/1	3/2	2/0	1/4	8/10			2/0				5/1	120/120
d		1/0	1/5	93/ 97	0/2	4/3			2/1	0/1					1/1	3/0	13/3	0/2	0/2					2/3	120/120
k	0/3	1/1	1/4	7/3	75/ 73	24/ 24										0/4	2/3	2/2						2/3	120/120
g			1/0	0/ 14	19/ 17	75/ 67							1/0	0/1		0/1	2/6	1/0	1/0					14/ 14	120/120
f	7/8						74/ 88	22/ 9	2/6			0/1				5/0	2/6							2/2	120/120
v	14/7	1/0		2/0			1/0	75/ 93		1/3		0/1				1/8	3/1		1/0					21/7	120/120
θ	0/4			0/6	1/0		1/1	11/ 7	73/ 85	1/1		3/7					0/5			1/0				17/4	120/120
ð			3/0	8/9	1/0		0/3	2/1	2/2	96/ 92		3/6					2/3			0/1				3/3	120/120
s			2/0	5/4	15 /14	1/1			3/6		76/ 67	1/5	13/8	0/11	0/1		3/3		1/0						120/120
z				5/5				0/2		0/1		114/ 107	0/1	0/1		1/0	0/2			0/1					120/120
ʃ				0/1	4/1	0/8		2/0		2/2	1/0	94/ 80	2/0	7/3	0/2	0/ 12								2/ 11	120/120
ʒ		2/0	4/3	7/0		7/5								74/ 85	20/ 22	1/1	4/1							1/3	120/120
dʒ				0/2	1/2	5/5						1/0	2/0	10/ 14	86/ 94		3/3								120/120
m	0/1			4/0	1/0											106/ 107	5/5	0/3	0/1	1/0				3/3	120/120
n	0/1	1/0			1/0		0/1	8/4								21/ 26	80/ 85		1/0					8/3	120/120
Total	122/ 107	97/ 90	76/ 84	176/ 161	122/ 112	125/ 120	76/ 93	127/ 129	88/ 103	98/ 98	78/ 69	129/ 127	110/ 90	95/ 114	116/ 121	175/ 196	133/ 152	4/7	4/3	4/3	0/0	0/0	0/0	85/ 61	2040/ 2040

Note. NR = No response.

Table 2.3

Overall Individual Speaker SI Raw and Percentage Scores for Word-Initial and Word-Final Positions Across Habitual Speech (HS) and Clear Speech (CS) Conditions

Speaker	Word-Initial				Word-Final			
	HS		CS		HS		CS	
	N	%	N	%	N	%	N	%
1	160	78.4	167	81.9	189	92.6	186	91.2
2	165	80.9	159	77.9	153	75.0	153	75.0
3	169	82.8	163	79.9	198	97.1	181	88.7
4	169	82.8	169	82.8	171	83.8	172	84.3
5	122	59.8	131	64.2	127	62.3	140	68.6
Servox Total	785	77.0	789	77.4	838	82.2	832	81.6
6	160	78.4	151	74.0	161	78.9	154	75.5
7	129	63.2	126	61.8	127	62.3	125	61.3
8	149	73.0	173	84.8	175	85.8	194	95.1
9	165	80.9	175	85.8	179	87.7	186	91.2
10	153	75.0	159	77.9	176	86.3	183	89.7
TruTone Total	756	74.1	784	76.9	818	80.2	842	82.5
Overall	1541	75.5	1573	77.1	1656	81.2	1674	82.1

When comparing SI scores between word positions, speakers achieved a higher SI score in WF compared to WI position (81.2% vs. 75.5%, respectively) when EL users used HS. Repeated measures ANOVA revealed a significant effect of word position on consonant SI scores in HS, $F(1, 8) = 5.515, p < .05$, partial $\eta^2 = .408$. Post-hoc testing indicated that consonant SI scores in WF position were significantly higher than WI position when EL users spoken in HS ($p < .05$). Similarly, speakers achieved a higher SI score in WF compared to WI position in CS compared to HS (82.1% vs. 77.1%, respectively). Repeated measures ANOVA revealed a significant effect of word position on consonant SI scores in CS, $F(1, 8) = 8.969, p < .05$, partial $\eta^2 = .529$. Post-hoc testing indicated that consonant SI scores in WF position were significantly higher than WI position when EL users spoken in CS ($p < .05$). Overall, word position was found to have a large effect on consonant SI scores for both speaking conditions (Cohen, 1988).

Comparison of Word Position by Device Group

Table 2.3 also provides a comparison of consonant scores by device. Servox users had a mean WI consonant score of 77.0% for HS ($Mdn = 80.9\%$; $range = 59.8-82.8\%$) and 77.4% for CS ($Mdn = 79.9\%$; $range = 64.2-82.8\%$). In WF position, Servox users achieved mean scores of 82.2% during HS ($Mdn = 83.8\%$; $range = 62.3-97.1\%$) and 81.6% during CS ($Mdn = 84.3\%$; $range = 68.6-91.2\%$). Repeated measures ANOVA testing indicated no significant effect of speaking condition on word position scores for Servox users.

In the WI position, TruTone users had a mean intelligibility score of 74.1% ($Mdn = 75\%$; $range = 63.2-80.9\%$) for HS and 76.9% ($Mdn = 77.9\%$; $range = 61.8-85.8\%$) for CS. In WF position, a score of 80.2% ($Mdn = 85.8\%$; $range = 62.3-87.7\%$) in HS and

82.5% (*Mdn* = 89.7%; *range* = 61.3-91.2%) in CS was identified. Overall, TruTone demonstrated slight increases in scores when using CS (2.8% in WI and 2.3% in WF), but there was no significant effect of speaking condition on word position scores for TruTone users.

Overall, Servox users had a consonant score that was 3.4% greater in WI position and 1.1% higher in WF position than those who used the TruTone across both speaking conditions. However, EL device did not have a significant effect on word position scores. For WI position, Servox users achieved a 2.9% increase in HS and 0.5% increase in CS compared to TruTone users. For WF position, Servox users achieved a consonant score that was 2.0% greater than TruTone users during HS, but TruTone users saw a slight benefit (0.9%) during CS compared to Servox users. Repeated measures ANOVA testing indicated that there was no significant effect of EL device on word position scores.

Voicing Feature

Voiced consonants. In total, voicing analyses were conducted on a total of 4320 voiced consonants across speaking conditions with 2160 in each speaking condition (HS and CS). Listeners correctly identified 1846 (*SD* = 9.9; *range* = 74-104) of 2160 voiced consonants in HS (*M* = 85.5%; *Mdn* = 86.6%; *range* = 68.5-96.3%) and 1892 of 2160 consonants (*SD* = 7.3; *range* = 78-104) in CS (*M* = 87.6%; *Mdn* = 88.4%; *range* = 72.2-96.3%). Although listeners identified 46 (2.1%) more voiced consonants in the CS condition, repeated measures ANOVA revealed no significant effect of speaking condition on voiced consonant scores.

WI voiced consonants. Table 2.4a shows the individual speaker scores for WI voiced consonants. Listeners correctly identified 887 (*SD* = 7.7; *range* = 74-103)

Table 2.4a

Overall Individual Speaker SI Raw and Percentage Scores for Word-Initial Phonemic Voicing Features Across Habitual Speech (HS) and Clear Speech (CS) Conditions

Speaker	WI Voicing							
	Voiced				Voiceless			
	HS		CS		HS		CS	
	N	%	N	%	N	%	N	%
1	85	78.7	87	80.6	75	78.1	80	83.3
2	88	81.5	87	80.6	77	82.0	72	75.0
3	84	77.8	78	72.2	85	88.5	85	88.5
4	92	85.2	96	88.9	77	78.1	73	76.0
5	84	77.8	94	87.0	38	39.6	37	38.5
Servox Total	433	80.2	442	81.9	352	65.2	347	64.3
6	93	86.1	92	85.2	67	69.8	59	77.6
7	103	95.3	102	94.4	26	27.1	24	25.0
8	74	68.5	95	88.0	75	78.1	78	81.3
9	90	83.3	96	88.9	75	78.1	79	82.2
10	94	87.0	90	83.3	59	61.5	69	71.9
TruTone Total	454	84.1	475	88.0	302	56.0	309	57.2
Overall	887	82.1	917	84.9	654	68.1	656	68.3

or 82.1% of voiced WI consonants ($Mdn = 82.4\%$; $range = 35-100\%$) in the HS condition compared to 917 ($SD = 6.6$; $range = 78-102$), or 85.5% ($Mdn = 86.1\%$; $range = 72.2-94.4$) in CS. Repeated measures ANOVA indicated no significant effect of speaking condition on WI voiced consonant scores.

Closer examination of the data in Table 2.4a revealed that Speaker 7 had the highest score for WI voiced consonants (95.3%) in both HS and CS (e.g., 103 and 102, respectively), while Speaker 8 had the lowest WI in HS (68.5%) and Speaker 3 had the lowest score WI in the CS condition (72.2%). This resulted in a difference in the HS condition of 26.8% between the best and lowest scores achieved, while a smaller difference of 22.2% was noted between these speakers for CS, a difference of 4.6% across conditions.

WF voiced consonants. Table 2.4b shows the individual raw and percentage intelligibility scores for WF voiced consonants. Listeners correctly perceived 959 ($SD = 10.9$; $range = 74-101$) or 88.8% ($Mdn = 93.5\%$; $range = 68.5-93.5\%$) of voiced consonants in HS and 975 ($SD = 7.0$; $range = 84-104$) or 90.3% ($Mdn = 92.1\%$; $range = 77.8-96.3\%$) in CS. No significant effect of speaking condition was found on SI scores of WF voiced consonants.

Closer examination of the data indicate that listeners correctly identified more voiced consonants for Speakers 3 and 10, who had raw SI scores of 96.3% in HS. This SI score was 27.8% greater than that of Speaker 2, who had the lowest HS score (68.5%). For CS, the highest scores were achieved by Speakers 7 and 9, also with 96.3% of consonants correctly identified. These two speakers had 18.5% more consonants correctly identified by listeners than Speaker 2, who also had the lowest SI score in CS (77.8%).

Table 2.4b

Overall Individual Speaker SI Raw and Percentage Scores for Word-Final Phonemic Voicing Features Across Habitual Speech (HS) and Clear Speech (CS) Conditions

Speaker	WF Voicing							
	Voiced				Voiceless			
	HS		CS		HS		CS	
	N	%	N	%	N	%	N	%
1	97	89.9	103	95.4	92	95.8	83	86.5
2	74	68.5	84	77.8	79	82.3	69	71.9
3	104	96.3	98	90.7	94	97.9	83	86.5
4	95	88.0	91	84.3	76	79.2	81	84.4
5	78	72.2	90	83.3	49	51.0	50	52.1
Servox Total	448	83.0	466	86.3	390	72.2	366	67.8
6	101	93.5	101	93.5	60	62.5	53	55.2
7	103	95.4	104	96.3	24	25	21	21.9
8	101	93.5	103	95.4	74	77.1	91	94.8
9	102	94.4	104	96.3	77	80.2	82	85.4
10	104	96.3	97	89.8	72	75.0	86	89.6
TruTone Total	511	94.6	509	94.3	307	56.9	333	61.7
Overall	959	88.8	975	90.3	697	72.6	699	72.8

Overall, there was an SI difference of 27.8% between the best and lowest scores for WF voiced consonants spoken in HS, while a smaller difference of 18.5% was noted between EL speakers for CS. Further, there was a difference of 9.3% when comparing speaking conditions.

Voiceless consonants. Analyses of voiceless consonants across HS and CS conditions were conducted and the data are presented in Tables 2.4a and 2.4b. In total, analyses were conducted on 3840 voiceless consonants with 1920 presented in both HS and CS conditions.

Overall, listeners correctly identified 1351 ($SD = 19.6$, $range = 24-94$) out of a total of 1920 voiceless consonants in HS condition and 1388 ($SD = 17.0$; $range = 24-85$) in CS condition. When converted into percentages, listeners correctly identified 70.4% ($Mdn = 78.1\%$; $range = 25.0-97.9\%$) of voiceless consonants in HS and 72.3% ($Mdn = 78.6\%$; $range = 25.0-88.5\%$) in the CS condition. Thus, listeners identified only 1.9% more voiceless consonants in CS than HS. A repeated measures ANOVA did not reveal a significant effect of speaking condition on voiceless consonant scores.

WI voiceless consonants. Table 2.4a illustrates that listeners correctly identified 654 ($SD = 19.1$, $range = 26$ to 85) of voiceless consonants in HS ($M = 68.1\%$; $Mdn = 78.1\%$; $range = 27.1-88.5\%$) and 656 ($SD = 20.0$; $range = 24-85$) in CS ($M = 68.3\%$; $Mdn = 75.5\%$; $range = 25-88.5\%$) in CS. There was no significant effect of speaking condition on voiceless consonant scores in WI position.

Individual speaker data indicated that Speaker 3 had the highest SI scores in both HS and CS (i.e., 85, or 88.5%, respectively), while Speaker 7 obtained the lowest SI score in HS (e.g., 26, or 27.1%) and in CS (e.g., 24, or 25%). This indicates that listeners

identified 59 (or, 61.4%) more voiceless consonants in WF position in HS and 61 (or, 63.5%) more in CS when comparing the highest and lowest speakers for each condition. Interestingly, Speaker 7 had the highest score for WI voiced consonants, yet was judged to have the lowest score for WF voiceless consonants. Similarly, Speaker 3 had the highest score in WI voiceless consonants, but the lowest score for WF voiced consonants. In addition, CS resulted in a greater difference in the correct identification of WI voiceless consonants between the best and worst speakers (e.g., Speaker 3 and Speaker 7).

WF voiceless consonants. Table 2.4b illustrates the findings for WF voiceless consonants. Listeners correctly identified of 697 ($SD = 20.8$, $range = 24-94$) voiceless consonants in HS and 732 ($SD = 13.4$; $range = 50-83$) voiceless consonants in CS. When these raw scores are converted to percentages, listeners correctly identified 72.6% ($Mdn = 78.1%$; $range = 25.0-97.9%$) of voiceless consonants in HS and 76.3% ($Mdn = 84.4%$; $range = 52.1-86.5%$) in CS. No significant effect of speaking condition was found on WF voiceless consonant scores.

Individual speaker data indicate that Speaker 3 achieved the highest score for WF voiceless consonants in HS (97.9%) and Speaker 7 obtained the lowest score (25%), a difference of 72.9% between the best and poorest speakers. For CS, Speakers 1, 3, 6, and 8 achieved the highest scores WF (86.5%) and Speakers 5 and 10 achieved the lowest (52.1%), thus a 34.4% difference.

Comparison of Voiced and Voiceless Consonants

Listeners correctly identified the most voiced and voiceless consonants for Speaker 3 in WF position in HS. Speaker 7 achieved the second highest SI score (i.e., 103

vs. 104) for WF voiced consonants in HS and the highest rating (i.e., 104) in CS.

Interestingly, Speaker 7 achieved the lowest SI score for voiceless consonants in WF position in HS and CS condition. Speaker 10 achieved a similar high score as Speaker 3 for WF voiced consonants, but had the lowest SI score for WF voiceless consonants. Therefore, it is suggested that CS has the potential to negatively impact SI consonant scores for some EL speakers' WF voiced consonants, but not voiceless consonants.

Comparison of Voiced and Voiceless Consonants by Speaking Condition

A comparison of how well listeners were able to correctly identify voiced versus voiceless consonants was conducted across speaking conditions. Overall, listeners correctly identified significantly more voiced (15.2%) than voiceless consonants (85.5% vs. 70.3%) in HS. Results from repeated measures ANOVA revealed a significant effect of voicing on consonant scores produced in HS, $F(1,18) = 7.974, p < .05$, partial $\eta^2 = 0.307$. Post-hoc testing with a Bonferroni correction revealed that voiced consonant scores were significantly higher than voiceless consonants in HS ($p < .05$). Similarly, listeners correctly identified 16.2% more (87.6% vs. 71.4%) voiced in the CS condition. A repeated measures ANOVA revealed a significant effect of voicing on consonant scores produced in CS, $F(1,18) = 8.720, p < .01$, partial $\eta^2 = 0.326$. Post-hoc testing with a Bonferroni correction revealed that voiced consonant scores were significantly higher than voiceless consonants in CS ($p < .05$). The magnitude of the above effects revealed that voicing had a large effect on listeners' identification of phonemes in both speaking conditions and word positions (Cohen, 1988).

Comparison of Voiced and Voiceless Consonants by Word Position

Relative to WI consonants, listeners correctly identified 14.0% more voiced than voiceless consonants (82.1% vs. 68.1%) in HS. For the CS condition, listeners correctly perceived voiced consonants 15% more often (84.9% vs. 69.9%) than voiceless consonants. Voicing was not found to have a significant effect on consonants scores in WI position when EL speakers used HS or CS.

In WF position, listeners perceived 16.2% (i.e., 88.8% vs. 72.6%) more voiced than voiceless consonants in HS. Repeated measures ANOVA revealed that voicing had a significant effect on WF consonant scores in HS, $F(1,8) = 7.288$, $p < .05$, partial $\eta^2 = 0.477$. The magnitude of the effect revealed that voicing had a large effect on listeners' identification of phonemes in WF position. Post-hoc testing with a Bonferroni correction revealed that voiced consonant scores were significantly higher than voiceless consonant scores in WF position ($p < .05$).

For the CS condition, listeners perceived voiced consonants 17.5% (i.e., 90.3% vs. 72.8%) more than voiceless consonants in WF position. Results indicate that voicing had a significant effect on WF consonant scores, $F(1,8) = 2.751$, $p < .05$, partial $\eta^2 = 0.448$. The magnitude of the effect revealed that voicing had a large effect on listeners' identification of phonemes in WF position while EL users' spoke using CS (Cohen, 1988). Post-hoc testing revealed that WF voiced consonant scores were significantly higher than voiceless consonants in WF position ($p < .05$).

Manner Feature

An analysis of manner in both WI and WF positions was conducted on 2040 stimuli for each word position for each speaking condition. Overall raw and percentages

scores for segmented by manner features are presented in Tables 2.5a and 5b. Listeners correctly identified 61.3% WI consonants in HS and 63.7% in CS. Similarly, for WF consonants, listeners correctly identified 69.5% in HS and 70.4% in CS.

The most accurately perceived manner class of WI consonants in both HS and CS conditions was found for nasals, followed by plosives, fricatives, and affricates. The largest improvements observed for WI consonants in the CS condition were noted for plosives and affricates which improved by a raw score of 27 (3.8%) and 27 (11.2%), respectively. For WF consonants, the most accurately perceived manner class in both HS and CS were nasals, affricates, fricatives, and plosives.

Nasals. No significant effects were found in the identification of nasals when speaking condition and word position were considered.

Plosives. No significant effects were found in the analysis of plosives when speaking condition and word position were considered.

Fricatives. No significant effects were found in the identification of fricatives when speaking condition and word position were considered.

Affricates. The effect of speaking condition on the SI score of affricates according to word position approached significance, $F(1,8) = 5.206$, $p = .052$, partial $\eta^2 = .394$. This p -value was close to the a-priori significance value of .05.

Table 2.5a

Individual Speaker Raw and Percentage SI Scores for Word-Initial Consonants By Manner Class Across Habitual Speech (HS) and Clear Speech (CS) Conditions.

Speaker	Habitual Speech								Clear Speech							
	Plosives		Fricatives		Affricates		Nasals		Plosives		Fricatives		Affricates		Nasals	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1	56	77.8	61	72.6	11	45.8	24	100	57	79.2	59	70.2	14	58	24	100
2	40	55.6	50	59.5	11	45.8	24	100	46	63.9	52	61.9	10	41.7	24	100
3	54	75.0	68	81.0	11	45.8	24	100	57	79.2	60	71.4	11	45.8	24	100
4	49	68.1	61	72.6	10	41.7	24	100	44	61.1	50	59.5	13	54.2	24	100
5	22	30.6	35	41.7	6	25.0	24	100	30	41.7	30	35.7	10	41.7	19	79
Servox Totals	221	61.4	275	65.5	49	40.8	120	100	234	65.0	251	59.8	58	48.3	115	95.8
6	46	63.9	55	65.5	6	25.0	24	100	48	66.7	50	59.5	6	25.0	24	100
7	38	52.8	18	21.4	1	4.2	24	100	36	50.0	16	19.0	1	4.2	24	100
8	38	52.8	50	59.5	1	4.2	24	100	50	69.4	64	76.2	12	50.0	24	100
9	54	75.0	54	64.3	9	37.5	24	100	57	79.2	62	73.8	10	41.7	24	100
10	49	68.1	43	51.2	3	12.5	24	100	48	66.7	52	61.9	9	37.5	24	100
TruTone Totals	225	62.5	220	52.4	20	16.7	120	100	239	66.4	244	58.1	38	31.7	120	100
Overall	446	61.9	495	58.9	69	28.8	240	100	473	65.7	495	58.9	96	40.0	235	97.9

Table 2.5b

*Individual Speaker Raw and Percentage SI Scores for Word-Final Consonants
By Manner Class Across Habitual Speech (HS) and Clear Speech (CS) Conditions*

Speaker	Habitual Speech								Clear Speech							
	Plosives		Fricatives		Affricates		Nasals		Plosives		Fricatives		Affricates		Nasals	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1	65	90.3	72	85.7	21	87.5	21	88	53	73.6	69	82.1	22	92	20	83
2	39	54.2	61	72.6	11	45.8	15	63	39	54.2	57	67.9	18	75.0	12	50.0
3	61	84.7	68	81.0	21	87.5	23	96	60	83.3	79	94.0	23	95.8	24	100
4	41	56.9	72	85.7	11	45.8	18	75	36	50.0	69	82.1	15	62.5	17	71
5	23	31.9	39	46.4	15	62.5	8	33	36	50.0	37	44.0	12	50.0	6	25
Servox Totals	229	63.6	312	74.3	79	65.8	85	70.8	224	62.2	311	74.0	90	75.0	79	65.8
6	48	66.7	57	67.9	19	79.2	24	100	45	62.5	53	63.1	20	83.3	24	100
7	31	43.1	25	29.8	10	41.7	19	79	28	38.9	25	29.8	6	25.0	21	88
8	52	72.2	62	73.8	22	91.7	23	96	61	84.7	71	84.5	23	95.8	23	96
9	58	80.6	68	81.0	22	91.7	12	50	58	80.6	74	88.1	24	100	12	50
10	53	73.6	66	78.6	18	75.0	23	95.8	57	79.2	67	79.8	16	66.7	24	100
TruTone Totals	242	67.2	278	66.2	91	75.8	101	84.2	249	69.2	290	69.0	89	74.2	104	86.7
Overall	471	65.4	590	70.2	170	70.8	186	77.5	473	65.7	601	71.5	179	74.6	183	76.3

The magnitude of the effect of speaking condition on the production of affricates was deemed to be a large effect (Cohen, 1988). Post-hoc testing revealed that SI of affricates in WI position was marginally higher when EL spoke in CS compared to HS ($p = .052$). No significant effect of speaking condition on the SI of affricates in WF position was observed.

Comparison of Individual Speaker Scores by Word Position

Overall, Speaker 3 had the highest SI for WI consonants in all manner classes for HS (77.0%), while Speaker 5 had the lowest score WI (42.6%), a difference of 34.4% between these speakers. These figures remained similar when EL speakers used CS; that is, a difference of 37.8% between Speaker 1 (who achieved the highest score of 75.5%) and Speaker 7 (who achieved the lowest score of 37.7%) was observed. This reveals a relatively similar performance in the range of scores when word position and speaking condition are considered.

The biggest improvement in WI consonant scores across all manner classes was achieved by Speaker 8 while using CS; that is, listeners identified 18.1% more consonants in CS compared to Speaker 8's consonant productions using HS. Meanwhile, Speaker 3 experienced the highest (although small) decrease of 2.5% in his SI scores moving from HS to CS.

Overall, Speaker 3 had the highest intelligibility for WI consonants in all manner classes for HS (77.0%), while Speaker 5 had the lowest score WI (42.6%), a difference of 34.4% between these speakers. These figures remained similar when EL speakers used CS; that is, a difference of 37.8% between Speaker 1 (who achieved the highest score of 75.5%) and Speaker 7 (who achieved the lowest score of 37.7%) was observed. This

reveals a relatively similar performance in the range of scores when word position and speaking condition are considered.

The biggest improvement in WI consonant scores across all manner classes was achieved by Speaker 8 while using CS; that is, listeners identified 18.1% more consonants in CS compared to Speaker 8's consonant productions using HS. Meanwhile, Speaker 3 experienced the highest, although small decrease of 2.5% in his intelligibility scores between HS and CS conditions.

The largest improvement in WF consonant scores was also achieved by Speaker 8 for whom listeners identified 9.3% more WF consonants in CS compared to HS. Meanwhile, Speaker 1 experienced the greatest reduction in his WF score with a decrease of 7.4% between HS and CS conditions. These values are two to three times less than the values achieved by the best (and worst) EL speaker in WI position. For example, Speaker 8 improved by 9.3% in WF position using CS, but this was almost half of the 18.1% improvement obtained in WI position.

Overall, all speakers achieved relatively higher scores across all manner classes in WF when compared to WI. For example, Speaker 1 achieved a score of 87.7% across all manner classes in HS, while Speakers 5 and 7 achieved the lowest SI scores with 41.7%. This difference between the most and least intelligible speakers in WF and WI position was 45.3%, indicating a substantially larger difference between scores for WF versus WI positions.

Omissions

The overall number of omissions classified by manner feature is summarized in Table 2.6a. Listeners omitted a total of 147 consonants in the HS condition and a total of 94 consonants in the CS regardless of word positions. Listeners omitted 31 and 33 fewer consonants in WI and WF positions, respectively, when speakers used CS. Across all manner classes, listeners consistently omitted more fricatives in WI position for both the HS and CS conditions. This was followed closely by omissions of plosives (26), with fewer noted for affricates and nasals. In WF position, listeners omitted fricatives and plosives most often, followed by nasals and affricates in both HS and CS conditions.

Table 2.6b provides a further breakdown of consonant omissions by listeners according to individual speaker data. The range was 0 to 37 WI consonants omitted across all individual EL speakers in HS, and 0 to 16 WI consonants across all individual EL speakers in CS. The range in WF position was comparable; 0 to 35 consonants were omitted in HS and 0 to 23 consonants were omitted in CS.

Overall, repeated measures ANOVA testing indicated that there was no significant effect of speaking condition on omissions according to manner or word position.

Table 2.6a

Total Number of Omissions By Manner Feature for Habitual Speech (HS) and Clear Speech (CS) Conditions

Target Consonant	HS			CS		
	WI	WF	Total	WI	WF	Total
Plosives	26	28	54	10	25	35
Fricatives	28	45	73	17	27	43
Affricates	8	1	9	5	3	9
Nasals	0	11	11	1	6	7
Totals	62*	85*	147**	33*	61*	94**
*2,040 possible targets in each word-position						
**4,080 total targets across word-positions in each speaking condition						

Table 2.6b

Individual Speaker Omissions By Manner Feature for Habitual Speech (HS) and Clear Speech (CS) Conditions

Speaker	WI								WF							
	Plosives		Fricatives		Affricates		Nasals		Plosives		Fricatives		Affricates		Nasals	
	HS	CS	HS	CS	HS	CS	HS	CS	HS	CS	HS	CS	HS	CS	HS	CS
1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
2	5	1	1	2	0	0	0	0	8	9	8	2	0	0	1	0
3	3	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0
4	0	1	0	0	0	0	0	0	3	6	3	5	0	1	0	0
5	9	4	21	8	7	3	0	1	14	4	13	14	0	1	8	4
Servox Totals	17	6	23	10	8	3	0	1	26	20	24	21	0	2	9	4
6	1	1	1	1	0	1	0	0	1	2	6	1	1	0	0	0
7	3	2	1	5	0	1	0	0	1	3	10	4	0	1	1	1
8	5	1	3	1	0	0	0	0	0	0	5	1	0	0	1	1
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
TruTone Totals	9	4	5	7	0	2	0	0	2	5	21	6	1	1	2	2
Overall	26	10	28	17	8	5	0	1	28	25	45	27	1	3	11	6

Discussion

Given that CS has consistently been shown to improve SI for individuals with communication disorders up to 26% (Payton et al., 1994; Picheny et al., 1985), this study sought to identify impact of CS on the SI of words and word position of consonants in EL speech. Comparisons between HS and CS indicated that EL speakers achieve similar SI scores for words and word positions of consonants. These results will be discussed more thoroughly in the following sections.

Word Intelligibility

Relative to word SI, results revealed no significant effect of speaking condition on listeners' transcription words. There was, however, a small improvement observed in SI when these speakers used CS compared to HS (i.e., 53.0% vs. 51.7%). These data are comparable to word scores previously reported in the literature that range from 35.5% to 60.3% (Bennett & Weinberg, 1973; McCroskey & Mulligan, 1963; Shames et al., 1963; Weiss & Basili, 1985; Weiss et al., 1979). While previous attempts have been made to improve both SI and overall signal quality of EL speech, the present findings support the notion that EL devices have remained relatively similar since their introduction in 1959. To assist in generalizing the present findings, Egan (1948) provided evidence to indicate a relationship between the intelligibility of words and sentences (or, "articulation") scores; that is, sentence intelligibility scores are often higher than those generated from isolated word scores. As stated by Barney et al. (1959):

...it has been found that a 60 per cent articulation from such isolated words corresponds to a sentence intelligibility of more than 95 per cent, and that even 40% in the word score means that more than 90 per cent of sentences would be understood. (p. 1355).

This may suggest that the ~53.0% word intelligibility achieved by the present EL speakers while using CS could correspond to more than 90% SI if sentence stimuli were used. It is important to note that the isolated words used in this study presented a highly decontextualized communication context to participant listeners (i.e., transcribing isolated single words). Yet, it was the intention of the present work to more fully understand the impact of CS when listeners were asked to identify words in such a context (i.e., single CVC words only). Therefore, the present study represents the first documented study to address the therapeutic application of CS in EL speech.

Second, the negligible 1.3% improvement in word intelligibility observed in the present work is lower than the previously reported benefits of CS when used with other clinical populations. Some have considered smaller improvements in SI secondary to CS (e.g., an 8% improvement for individuals with dysarthria) to be “clinically meaningful” in challenging contexts (Tjaden et al., 2014, p. 780). While the increase observed in our speakers is minimal, it is important to consider the difficulty in directly comparing speech produced by an EL speaker to individuals with neuromuscular conditions. Research to date indicates that individuals with neurologically-based speech disorders may benefit from using a slower rate of speech (Yorkston, Hakel, Beukelman, & Fager, 2007). However, the acoustic deficits inherent in EL speech (i.e., device noise, low fundamental frequency cut-off, lack of variable frequency) might be too complex to overcome through the simple modification of speech rate and over-articulation. Furthermore, EL speakers are reliant on an externally-based, electronic voice source, whereas speakers with dysarthria continue to use a laryngeal-based voice source when using CS.

The general lack of differences between word SI scores and speaking conditions is in part accounted for by the significant, positive correlation that exists between them. This correlation accounts for nearly 71% of the variance between word SI scores and speaking conditions. It is important to recall that each EL speaker was provided with general instructions to make their speech clearer, slow down their rate of speech, and over-articulate. This suggests that the EL speakers might have been already speaking as clearly as possible using a reduced rate of speech alongside over-articulation (i.e., from previous EL training). Further, research has indicated that a reduction of speech rate is necessary in CS, but not the only factor that can account for changes in SI during CS (Lam & Tjaden, 2013; Picheny et al., 1989; Uchanski, Choi, Braida, Reed, & Durlach, 1996). Picheny et al. (1989) and Uchanski et al. (1996) suggest that improvements in SI cannot be observed without a reduction in speech. However, Lam and Tjaden (2013) sought to examine the best set of instructions leading to improved SI. The researchers found that telling speakers to ‘over-enunciate’ each word lead to the highest SI (Lam & Tjaden, 2013). Overall, it appears that the changes in SI in the present experiment, although not significant, are accounted for by either by EL speakers’ training and maximal level of performance already being met, or a failure for all speakers to utilize the numerous (and required) productive changes to produce CS.

Comparison of Word Intelligibility by Device Group

There was no significant effect of EL device used on word SI scores. However, substantial variability was observed. For example, Speaker 5 (Servox) and Speaker 7 (TruTone) exhibited the lowest overall word scores in HS (29.2%) and CS (29.2%), respectively. Speaker 5’s word SI score only improved by 2.7% for CS. In contrast,

Speaker 7's CS score was 3.2% lower than his score in HS (32.4%). The highest overall HS word score (69.9%) was achieved by Speaker 1, but his score was reduced by 8.3% during CS.

Further analysis of the individual speaker data revealed that five speakers (two Servox and three TruTone users) demonstrated improved word scores in the CS condition. The mean improvement in overall SI for these speakers during CS was found to be 7.3% (*Mdn* = 5.3%; *range* = 2.8-16.2%). In contrast, for the remaining five speakers who exhibited reductions in their scores when using CS, a mean change of 3.8% (*Mdn* = 3.2%; *range* = 1.4- 8.3%) was observed. Thus, patterns of speaker performance were highly individualized.

Intelligibility by Consonant Position: WI and WF

Results from the overall analyses of listener data indicated that only WI consonants were significantly different when EL speakers used CS compared to HS. However, relatively small improvements of 1.6% and 0.9% were observed for WI and WF consonants, respectively.

Previous studies have described the difficulties that listeners have in correctly perceiving both WI and even some WF consonants in EL speech; that is, while perceptual errors in voicing commonly arise, some errors related to manner can occur (Weiss & Basili, 1985; Weiss et al., 1979). The present data provide support for the general differences in EL speakers' scores relative to phonetic position. In the HS condition, for example, EL speakers' consonants were 11.4% more intelligible in the WF compared to WI position. For CS, EL speakers as a group improved their consonant SI by 8.5% in WF position when compared to WI. These comparisons were found to be significantly

different, and are in agreement with previous research indicating that the identification of WF consonants is often higher than WI consonants in EL speech (Weiss & Basili, 1985; Weiss et al., 1979). Research has suggested that this improved WF consonant SI scores are likely the result of the durational properties of the preceding vowel in normal and EL speech (Raphael, 1972; Weiss et al., 1979).

Interestingly, similar SI values were observed between the most and least intelligible speakers in WI and WF positions (e.g., most intelligible in HS and most intelligible in CS). This indicates that some EL speakers will not necessarily derive further benefit from using CS, especially if they achieve a relatively higher SI in HS. Moreover, EL speakers who begin with a lower SI score may, in fact, be more likely to improve their SI through the use of CS. This finding was observed for at least half of the EL speakers in each phonetic position. This speaks to the wide-variability found in the individual speaker data in the identification of both WI and WF consonants. Overall, EL speaker scores in HS ranged from 42.2% to 77.5% for WI position to 41.7% to 91.7% for WF position. These values are similar for the WI and WF consonants in CS (e.g., 42.6% to 75.5% and 39.2% to 91.2%, respectively).

Comparison of Word Position by Device Group

While WF consonants were identified 5.7% more than WI consonants in HS and 5% more in CS, there was further variability noted between device groups (although a similar range was relatively found between devices). For example, speakers using Servox devices ranged from 42.2% to 77.5% in WI consonant intelligibility. This differed slightly from TruTone users, whose WI consonant SI ranged from 45.6% to 68.6%. This

demonstrates similar performance between device groups, even though TruTone users had a slightly narrower range (i.e., variability) in performance.

Overall, greater variability and reduced overall perception of WI consonants support previous research data indicating that listeners more accurately perceive WF consonants in EL speech (Weiss & Basili, 1985; Weiss et al., 1979). Considering that the EL is a constantly voiced source, these findings highlight how more WI voiceless consonants are misperceived as voiced in WI position. Therefore, the following section will further elaborate on the present study's findings related to voicing features.

Voicing Feature

Listener perception of EL speech has been consistently met with difficulty due to the continuous voiced nature of this alaryngeal communication option. The present findings support previous data indicating that EL speakers have difficulty maintaining voicing characteristics of individual consonants (Weiss & Basili, 1985; Weiss et al., 1979). Overall, the perception of voiced consonants was significantly better than voiceless consonants in HS and CS. According to phonetic position, only voiced consonants in WF position were significantly between HS and CS. Closer examination of the listeners' data indicate that 31.9% WI voiceless consonants in HS were misperceived as voiced, while 31.7% WI voiceless consonants were noted with CS. Similarly, 27.1% of voiceless consonants in WF position were incorrectly perceived as voiced in HS, with 27.2% noted for CS. Given that EL speech is characterized by continuous voicing (Weiss et al., 1979), the present data indicate that use of CS may not serve to overcome these inherent acoustic EL signal limitations.

Previous research has indicated that CS has the potential to increase sound energy in the 1000 to 3000Hz range (Krause & Braida, 2004). Weiss et al. (1979) indicated that EL speech is typically 5-10 dB higher than normal, laryngeal speech between 2k to 4k Hz. If an increase is observed in CS, then CS could result in more ‘voiced’ sound energy that is transferred from the neck directly into a larger-than-normal oral cavity (in part, due to the exaggerated articulatory productions). This is an over-simplified outcome, given the host of surgical and other issues that can impact the neck-transfer function of the EL speech signal (Meltzner, Kobler, & Hillman, 2003).

The listeners’ difficulty in perceiving the voicing characteristics of consonants is demonstrated by individual speaker data, especially in the WI position. Across both speaking conditions, EL speakers obtained higher scores when producing voiced consonants in WI and WF positions when compared to voiceless consonants. For example, EL speakers achieved SI scores that were 15.2% (i.e., 85.5% vs. 70.3%) greater for voiced compared to voiceless consonants while using HS and 16.2% (87.6% vs. 71.4%) greater while using CS.

EL speakers using both HS and CS were able to achieve a statistically significant increase in the overall SI of voiced consonants compared to voiceless consonants in both phonetic positions. This general finding is likely to be the result of the continuous voicing provided by an EL device, in addition to the articulatory precision afforded by the instructions for EL speakers to over-articulate and slow their rate of speech. Finally, the improvements observed in voiced consonants in WF position may be the result of the acoustic cues of the preceding vowel; that is, vowel duration can play a large role in maintaining the voicing characteristics of following consonants (Raphael, 1972; Weiss et

al., 1979). A commensurate change was not noted with voiceless sounds. However, this is likely attributed to the fact that EL speakers are unable to ‘turn-off’ the voiced nature of their speech regardless of the rate-enhancing technique used. Therefore, due to the electronically-based nature of EL speech, the correct identification of voiceless sounds was reduced in both WI and WF position.

Overall, EL speakers maintained the voicing characteristics of voiced consonants when compared to voiceless sounds. Data from the present study indicate also that voiced-for-voiceless errors were more common than voiceless-for-voiced in WI position. Considering that the EL is a continuously voiced source, this finding is not surprising across word-position and speaking condition. For example, voiced-for-voiceless confusions occurred 14% more frequently than did voiceless-for-voiced confusions in HS and 16.6% more than voiceless-for-voiced confusions in CS. In WF position, voiced-for-voiceless confusions occurred 16.2% more than voiceless-for-voiced confusions in HS and voiced-for-voiceless confusions occurred 7.5% more than often in CS.

The present study is the first to provide evidence suggesting that the use of CS in EL speakers might further facilitate improvements in listener identification of voicing characteristics for EL speech. For example, WI voicing characteristics were correctly identified 2.2% more when EL speakers used CS. This resulted in 6.9% less voicing confusions in WF position. EL speakers were able to increase their SI scores for voiced in WI position by 2.8% and 1.5% in WF position when they used CS. Overall, the present data are similar to those previously reported for EL speakers with respect to the voiced-for-voiceless confusion (Weiss & Basili, 1985; Weiss et al., 1979).

Manner Feature

The present data reveal that listeners identified a similar number of nasals, plosives, and fricatives according to word position and speaking condition. However, affricates were significantly different according to word position and/or speaking condition (see Tables 5a and b). Similar to other analyses of WI intelligibility in this study, listeners exhibited greater variability in the perception of WI consonants according to manner feature. Listeners had the greatest accuracy identifying nasals followed by plosives, fricatives, and affricates in WI position. In WF position, listeners accurately identified nasals, affricates, plosives, and fricatives.

Plosives. Overall, results demonstrated that listeners accurately perceived 2.8% more WI plosives when speakers used CS compared to HS. As seen in Table 2.5a, however, results indicate that speaking condition did not have a significant effect on maintaining the manner feature, particularly for voiceless plosives.

Overall, more voiced-for-voiceless confusions compared to voiceless-for-voiced confusions occurred for plosives. In fact, the only voiceless-for-voiced error listeners made in WI position involved /g/ for /k/ confusions. The most prominent voiced for-voiceless confusion involved /p/ for /b/ in WI position. Data revealed that this was often attributed to listeners incorrectly perceiving the word ‘pad’ as ‘bad’. This /b/-/p/ confusion occurred 3.3% more in CS. Similar voiced-for-voiceless confusions were also observed with /t/ being confused with /d/, but these were negligible across conditions.

Affricates. It is important to mention that affricates (and nasals) accounted for the lowest represented consonant group. In WI position, 28.8% of affricates were identified

in HS and 40.0% in CS. This was drastically reduced when compared to WF affricates, which were perceived 70.8% of the time in HS and 74.6 % of the time in CS.

Previous data have indicated increases of 22% to 27% when comparing WI and WF affricates in EL speech (Weiss & Basili, 1985). The high-level of SI for WF affricates (and fewer errors) is in agreement with Weiss et al. (1979) and Weiss and Basili (1985). These studies reported SI scores for WF affricates to as high as ~93% (Weiss & Basili, 1985; Weiss et al., 1979). It is, however, important to note that WI affricates were perceived more accurately when EL speakers used CS (i.e., 40.0% in CS vs. 28.8% in HS) with smaller increases of 3.8% in WF position (i.e., 74.6% in CS vs. 70.8% in HS).

Improvements in listener identification of WF affricates are potentially the result of the blended stop and fricative components. Fricatives (or, the fricative component in affricates) in EL speech may benefit from the durational properties of preceding vowels (Weiss et al., 1979). This is especially helpful in maintaining their voicing characteristics since vowel cues are generally well-preserved in EL speech (Raphael, 1972; Weiss et al., 1979). In addition, the use of CS seeks to improve intelligibility through a slowed-rate of speech and over-articulation, which in turn, may then serve to lengthen durational aspects of the affricate (Picheny et al., 1985; 1986; Krause & Braida, 2004).

Overall, manner confusions could imply that over-articulation of voiceless phonemes (e.g., plosives) and a reduced rate of speech during CS can actually work against EL speakers in specific phonemic contexts. First, consider that EL speakers have difficulty with voiced-for-voiceless distinctions due to their continuously voiced signal. Second, EL speakers produce voice without a reliance on direct pulmonary support, and

therefore, may not produce a substantial release burst that is characteristic of plosives. Third, the slower rate of speech and exaggerated articulation could modify plosives so that lengthening occurs. This may provide an explanation for the confusion of plosives for continuant sounds. For example, there were 51 (or, 7.1%) WI nasal-for-plosive confusions when EL speakers used HS and 56 (or, 7.8%) WI nasal-for-plosive confusions when EL speakers used CS. This trend continued in WF position; that is, there were 76 (or, 10.6%) nasal-for-plosive confusions when EL speakers used HS compared to 78 (or, 11.3%) plosive-for-nasal confusions when EL speakers used CS.

Omissions

Plosives. Analysis of omissions indicated that changes in EL speakers' articulation while using CS could have led to listeners omitting fewer plosives. Although no statistically significant effect of speaking condition on omissions was observed conditions, the greatest benefit with CS was observed in WI position; more specifically, 16 fewer plosives were omitted (18% of plosive omissions) across HS and CS. Meanwhile, a relatively similar number of plosives (e.g., 25 and 28) were omitted across HS and CS in WF position.

Fricatives. Listeners correctly identified more fricatives while using CS when compared HS in both WI and WF positions. With 11 and 18 fewer fricatives omitted in WI and WF position respectively, EL speakers may benefit (even to a small degree) while using CS during the production of fricatives.

With few exceptions, fricatives were perceived similarly across HS and CS. Scores for WI voiced fricatives are similar to previously reported data of 12-16% for /v/ and 19-32% for /z/ (Weiss & Basili, 1985). Lower SI scores were observed for WI voiced

fricatives than those in WF position across HS and CS, which has been reported previously in EL speakers (Weiss & Basili, 1985; Weiss et al., 1979).

The potential reason for better SI scores for WF voiced fricatives is twofold. Based on the work of Raphael (1972), Weiss et al. (1979) indicated that fricatives might benefit from durational characteristics of the preceding vowel in order to maintain voicing characteristics. Research has indicated that these vowel cues are well-preserved in EL speech (Weiss et al., 1979). In addition, CS can increase the durations of vowels as speakers attempt to make their speech clearer to the listener (Picheny et al., 1986). In the present study, vowels were highly intelligible in both HS and CS conditions (85.4% and 82.7%). This may explain how correct listener identifications of /v/ increased by 45.8% in HS and 58.3% in CS in WF position. In addition, correct identification of /z/ in WF position increased by 70% in HS and 55.8% in CS. Taken together, research supports the notion that CS can improve the SI of WI fricatives. Unfortunately, no perceptible differences were observed for WF fricatives across HS and CS conditions in the present study.

Conclusions

This is the first study to examine the potential influence of CS on EL speakers' word and consonant SI by word position. The present findings provide initial evidence suggesting that volitional attempts to improve EL speech using CS do not result in large changes relative to listener perceptions of words. However, the potential exists for future research to demonstrate 'clinically meaningful' improvement in the SI of sentence-level and connected speech when EL speakers use CS. Previous reports of CS leading to improvements in SI of ~8% have been deemed of value in challenging contexts (Tjaden

et al., 2014; Van Nuffelen, De Bodt, Vanderwegen, Van de Heyning, & Wuyts, 2010). Although the present study did not use a perceptually challenging perceptual context (e.g., multi-talker babble), naïve listeners were required to transcribe single words spoken by EL speakers in HS and CS. Given the unique and unnatural acoustic and perceptual qualities of EL speech when compared to laryngeal speech, the transcription task in Experiment 1 could be considered challenging for naïve participant listeners. Therefore, this scenario provides a potential means to discuss any possible ‘clinical meaningfulness’ of SI improvements with future data.

The present study provides valuable information into the potential utility of CS on the SI of words and consonants produced by EL speakers using currently available devices. General comparisons made between Servox Digital and TruTone devices resulted in no significant effect on the SI of words or consonants. Moreover, it is important to note that none of the EL speakers included were judged to be highly proficient in the use of the intonation controls afforded by the TruTone.

For the significantly different consonant scores in WI position, the increase was only 1.6% in CS compared to HS. It should be noted that voiced consonants in WF were perceived correctly more than voiceless consonants given, among other potential reasons, the voiced nature of EL speech. Given some of the limited changes observed in the present study, future research should consider controlling articulatory rates during CS in order to further assess whether improved word and consonant SI will occur. For example, will a slower rate of speech (e.g., monitored in syllables per second) lengthen vowels to a degree that permits notable differences to be achieved in listener perception of WF consonants between HS and CS? Overall, the current findings may provide an initial step

toward improving the SI of EL speakers through modifications employing the concept of CS. It is believed that the present study's findings highlight the difficulty in improving SI for a speech signal that is based on an external, electronic voicing source. While EL speakers might have only gained a very small improvement in SI for the present study, future research should consider investigating the complex acoustic changes that occur during the application of CS in this unique population.

References

- American Speech-Language-Hearing Association (2015). Retrieved from:
http://ncepmaps.org/_gl/_104/
- Barney, H. L., Haworth, F. E., & Dunn, H. K. (1959). An experimental transistorized artificial larynx. *The Bell System Technical Journal*, 38(6), 1337-1356.
- Beaudin, P.G. (2002). *A paired-comparison evaluation of electronic artificial laryngeal device preference for vowel stimuli*. Unpublished master's thesis, Western University, Ontario, Canada.
- Bennett, S., & Weinberg, B. (1973). Acceptability ratings of normal, esophageal, and artificial larynx speech. *Journal of Speech and Hearing Research*, 16, 608-615.
- Beukelman, D. R. Fager, S. Ullman, C. Hanson, E., & Logemann, J. (2002). The impact of speech supplementation and clear speech on the intelligibility and speaking rate of people with traumatic brain injury. *Journal of Medical Speech-Language Pathology*, 10, 237-242.
- Black, J. W. (1949). Natural frequency, duration, and intensity of vowels in reading. *Journal of Speech and Hearing Disorders*, 14(3), 216-221.
- Doyle, P.C. (1994). *Foundations of voice and speech rehabilitation following laryngeal cancer*. San Diego: Singular.
- Doyle, P.C., & Keith, R. L. (Eds) (2005) *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing*. Austin, TX: Pro-Ed.

- Eadie, T. L., & Doyle, P. C. (2004). Auditory-perceptual scaling and quality of life in tracheoesophageal speakers. *The Laryngoscope*, *114*(4), 753-759.
- Eadie, T. L., Otero, D., Cox, S. R., Johnson, J., Baylor, C., Yorkston, K., & Doyle, P. C. (in press). The relationship between communicative participation and post-laryngectomy speech outcomes. *Head and Neck*.
- Egan, J. P. (1948). Articulation testing methods. *Laryngoscope*, *58*, 955-991.
- Espy-Wilson, C. Y., Chari, V. R., MacAuslan, J. M., Huang, C. B., & Walsh, M. J. (1998). Enhancement of electrolaryngeal speech by adaptive filtering. *Journal of Speech, Language, and Hearing Research*, *41*, 1253-1264.
- Hillman, R. E., Walsh, M. J., Wolf, G. T., Fisher, S. G., & Hong, W. K. (1998). Functional outcomes following treatment for advanced laryngeal cancer. Part I - Voice preservation in advanced laryngeal cancer. Part II - Laryngectomy rehabilitation: The state of the art in the VA system. Research Speech-Language Pathologists. Department of Veterans Affairs Laryngeal Cancer Study Group. *Annals of Otology, Rhinology and Laryngology Supplement*, *172*, 1-27.
- Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility in dysarthria. *Journal of Speech and Hearing Disorders*, *54*, 482-499.
- Krause, J.C., & Braida, L.D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *Journal of the Acoustical Society of America*, *112* (5), 2165-2172.

- Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *Journal of the Acoustical Society of America*, *115*(1), 362-378. Naïve
- Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. *Journal of Speech, Language, and Hearing Research*, *36*(1), 21-40.
- Lam, J., & Tjaden, K. (2013). Intelligibility of clear speech: Effect of instruction. *Journal of Speech, Language, and Hearing Research*, *56*(5), 1429-1440.
- McCroskey, R. L., & Mulligan, M. (1963). The relative intelligibility of esophageal speech and artificial-larynx speech. *Journal of Speech and Hearing Disorders*, *28*(1), 37-41.
- Meltzner, G. S., & Hillman, R. E. (2005). Impact of aberrant acoustic properties on the perception of sound quality in electrolarynx speech. *Journal of Speech, Language, and Hearing Research*, *48*(4), 766-769.
- Meltzner, G. S., Kobler, J. B., & Hillman, R. E. (2003). Measuring the neck frequency response function of laryngectomy patients: Implications for the design of electrolarynx devices. *The Journal of the Acoustical Society of America*, *114*(2), 1035-1047.
- Mendenhall, W. M., Morris, C. G., Stringer, S. P., Amdur, R. J., Hinerman, R. W., Villaret, D.B., & Robbins, K. T. (2002). Voice rehabilitation after total laryngectomy and postoperative radiation therapy. *Journal of Clinical Oncology*, *20*, 2500-2505.

- Payton, K.L., Uchanski, R.M., & Braida, L.D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America*, 95, 1581–1592.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing. I. Intelligibility differences between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 28, 96–103.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 29, 434–446.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1989). Speaking clearly for the hard of hearing III: An attempt to determine the contribution of speaking rate to differences in intelligibility between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 32(3), 600-603.
- Qi, Y. Y., & Weinberg, B. (1991) Low-frequency energy deficit in electrolaryngeal speech. *Journal of Speech, Language, and Hearing Research*, 34(6), 1250-1256.
- Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *The Journal of the Acoustical Society of America*, 51(4B), 1296-1303.
- Royal College of Speech-Language Therapists (2005). Royal College of Speech Language Therapists Clinical Guidelines. S. Taylor-Goh (Ed.). Bicester: Speechmark Publishing Ltd.

- Saikachi, Y., Stevens, K. N., & Hillman, R. E. (2009). Development and perceptual evaluation of amplitude-based F0 control in electrolarynx speech. *Journal of Speech, Language, and Hearing Research, 52*(5), 1360-1369.
- Schiavetti, N. (1992). *Scaling procedures for the management of speech intelligibility*. In R. Kent (Ed.) *Intelligibility in speech disorders: Theory, measurement, and management*. Philadelphia, PA: John Benjamins Publishing Co.
- Shames, G. H., Font, J., & Matthews, J. (1963) Factors related to speech proficiency of the laryngectomized. *Journal of Speech and Hearing Disorders, 28*, 273-287.
- Smiljanić, R., & Bradlow, R. A. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Lang Linguist Compass, 3*(1), 236–264.
- Tardy-Mitzell, S., Andrews, M. L., & Bowman, S. A. (1985). Acceptability and intelligibility of tracheoesophageal speech. *Archives of Otolaryngology, 111*(4), 213-215.
- Tjaden, K., Sussman, J.E., & Wilding, G. E. (2014). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research, 57*, 779-792.
- Uchanski, R. M., Choi, S. S., Braida, L. D., Reed, C. M., & Durlach, N. I. (1996). Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate. *Journal of Speech, Language, and Hearing Research, 39*(3), 494-509.
- Van Nuffelen, G., De Bodt, M., Vanderwegen, J., Van de Heyning, P., & Wuyts, F. (2010). Effect of rate control on speech production and intelligibility in dysarthria. *Folia Phoniatica et Logopaedica, 62*(3), 110-119.

- Verdolini, K., Skinner, M. W., Patton, T., & Walker, P. A. (1985). Effect of amplification on the intelligibility of speech produced with an electrolarynx. *Laryngoscopy*, 95: 720-726.
- Ward, E. C., Koh, S. K., Frisby, J., & Hodge, R. (2003). Differential modes of alaryngeal communications and long-term voice outcomes following pharyngolaryngectomy and laryngectomy. *Folia Phoniatica et Logopaedica*, 55 (1), 39-49.
- Watson, P. J., & Schlauch, R. S. (2009). Fundamental frequency variation with an electrolarynx improves speech understanding: A case study. *American Journal of Speech-Language Pathology*, 18(2), 162–167.
- Watson, P. J., & Williams, S. E. (1987). Laryngectomees' and nonlaryngectomees' perceptions of three methods of alaryngeal voicing. *J Comm Disord.*, 20, 295-304.
- Weiss, M S., & Basili, A. M. (1985). Electrolaryngeal speech produced by laryngectomized subjects: Perceptual characteristics. *Journal of Speech and Hearing Research*, 28, 294-300.
- Weiss, M. S., Yeni-Komshian, G. H., & Heinz, J. M. (1979). Acoustical and perceptual characteristics of speech produced with an electronic artificial larynx. *The Journal of the Acoustical Society of America*, 65(5), 1298-1308.
- Yorkston, K. M., Hakel, M., Beukelman, D. R., & Fager, S. (2007). Evidence for effectiveness of treatment of loudness, rate, or prosody in dysarthria: A systematic review. *Journal of Medical Speech-Language Pathology*, 15(2), XI-XXXVI.

Chapter 3

The Influence of Clear Speech on Acoustic Characteristics of Electrolaryngeal Speakers

Acoustic characteristics have long been the focus of research on both normal and disordered speech production. This includes explorations of frequency, intensity, and temporal characteristics of the speech signal. Research findings on the temporal aspects of normal speech frequently highlight the importance of the contexts in which phonetic stimuli occur (Öhman, 1967; Raphael, 1972; Theodore, Miller, & DeSteno, 2009; Umeda, 1975; 1977). These contexts range from phonetic-level analyses, involving individual segmental durations within words (e.g., consonant-vowel-consonant) to more global, sentence-level analyses. Thus, temporal alterations at multiple levels of speech production have been considered in both populations of normal speaker and those with speech disorders. One specific temporal measure is that of speech rate. Speech rate is often measured in the number of syllables or words produced in a given time period (i.e., syllables or words divided by time), a measure that has been shown to vary considerably across individuals who speak with and without a larynx (Goldman-Eisler, 1954; 1956; Robbins, Fisher, Blom, & Singer, 1984). This variability may be attributed to the syllable or word length of an utterance, the number and duration of pauses, the speaker's rate of breathing, and articulation rate during speech production (Goldman-Eisler, 1954; 1956). In addition, speech rate has been shown to be similar to articulation rate when minimal pauses are present during speech (Goldman-Eisler, 1956). Crystal and House (1990) have measured articulation rate by calculating "...the average syllable duration for interpause intervals..." (p. 101). They found that articulation rate can naturally vary due to the

number of phones within syllables (Crystal & House, 1990). This variability in speech and articulation rate is also believed to be of particular importance when attempting to improve speech intelligibility (SI) through rate modification. Of all phonetic-level units, vowels appear to be the most sensitive speech sounds to changes in speech and articulatory rate.

A review of historic literature highlights the importance of vowels within words. Öhman (1967) suggested that consonants are merely “superimposed on a context dependent vowel substrate that is present during all of the consonantal gesture.” (p. 165). Depending on tongue height, oral cavity size, and area of oral or pharyngeal constriction, vocal tract configuration can change both the formant frequency and duration of vowels. Vowels also are influenced by the context in which they occur. In particular, vowel duration is the most sensitive acoustic feature relative to that of neighboring phonemes. It has been suggested that vowel duration can be impacted by the voicing of surrounding consonants, while manner and place features have relatively less influence on duration (House & Fairbanks, 1953; Raphael, Dorman, Freeman, & Tobin, 1975; Umeda, 1975). For example, Raphael (1972) found that vowels preceding voiceless consonants are approximately two-thirds to one-half of the duration when compared to vowels that precede voiced consonants.

In addition to durational data for vowels, their formant frequencies (or, the resonant energies generated in the vocal tract during speech) have also been thoroughly investigated. Most prominently, Peterson and Barney (1952) examined the formant structure of 10 English vowels produced by 33 men, 28 women, and 15 children. Following measurement of the formants generated and the calculation of the acoustic

relationships between the first formant (F1) and second formant (F2), vowel data were illustrated using F1/F2 plots to show the ‘vowel space’ for each vowel. Briefly, a vowel space provides a two-dimensional representation of individuals’ acoustic and articulatory space plotted according to inherent F1 and F2 formant frequencies. Peterson and Barney’s (1952) F1/F2 plots indicated that vowel categories are not defined by a specific formant frequency, but by the proportional relationship between formants. While absolute vowel formant frequencies were greatest for children, followed by women and then men, the proportional relationships between F1 and F2 were maintained. The general trend indicated that F1 frequencies were higher and F2 frequencies were lower as vowel height and tongue advancement were reduced. Thus, previous investigations suggest that reductions of speech rate, in combination with increased mouth opening, which correlates with tongue height, can potentially influence vowel formant frequencies and expand the vowel space (Ferguson & Kewley-Port, 2002, 2007; Picheny, Durlach, & Braida, 1986).

Interestingly, larger vowel spaces have been observed with higher levels of SI in normal talkers (Bradlow, Torretta, & Pisoni, 1996) and those with neurological conditions (e.g., amyotrophic lateral sclerosis) (Turner, Tjaden, & Weismer, 1995). Further, some research has indicated that reducing speech rate and over-articulating in an effort to make oneself clearer can increase the vowel space (i.e., expansion of vowel spaces leading to modification of formant frequency characteristics) (Chen, 1980; Ferguson & Kewley-Port, 2002, 2007; 1986; Moon & Lindblom, 1994; Picheny et al., 1986). For this reason, clear speech (CS) has been suggested as a prescribed style of speaking that encourages individuals to slow their rate of speech and over-articulate in an effort to make it clearer and more understandable to the listener (Picheny, Durlach &

Braida, 1985; Picheny et al., 1986). Picheny et al. (1986) found that CS produced by normal speakers was significantly longer in duration than their typical conversational speech. In fact, sentences produced using CS were twice the duration of the same sentences spoken using normal (or, ‘conversational’) speech. These differences in speaking rate were attributed to both the CS users’ ability to increase the duration of individual speech sounds and the addition or lengthening of pauses (Picheny et al., 1986). CS also produced numerous phonetic changes, including a decrease in the frequency of vowel reduction and increases in vowel duration (Ferguson & Kewley-Port, 2002; Picheny et al., 1986). Ferguson and Kewley-Port (2002) reported that vowels were approximately twice as long during CS when compared to normal, conversational speech when spoken by a healthy male talker. In addition, CS also has been shown to result in formant frequency changes for vowels, a change that is characterized by vowel space expansion (Ferguson & Kewley-Port, 2002; Moon & Lindblom, 1994). However, research has suggested that speaking rate alone is not the only important aspect of CS (Picheny, Durlach, & Braida, 1989; Uchanski, Choi, Braida, Reed, & Durlach, 1996). Lam, Tjaden, and Wilding (2012) indicated that, when comparing three different instruction sets (e.g., ‘speak clearly’, ‘talk to someone with a hearing impairment’, and ‘over-enunciate’), the ‘over-enunciate’ instructions appeared to produce the greatest change across several acoustic measures.

Until now CS has only been applied to individuals with disorders of speech production or speech reception difficulties (Beukelman, Fager, Ullman, Hanson, & Logemann, 2002; Picheny et al., 1985; Tjaden, Sussman, & Wilding, 2014). Therefore, the present study is concerned with the impact of CS on the acoustic characteristics of

speech produced by individuals who have undergone total laryngectomy (TL) and use of the artificial electronic larynx.

TL is a procedure that removes the larynx and necessitates the use of a postsurgical ‘alaryngeal’ method of verbal communication. Alaryngeal speakers typically produce speech at a slower rate than normal speakers and this varies according to the method of post-laryngectomy speech used, as well as the speaker (Doyle & Eadie, 2005). For example, Robbins et al. (1984) found that normal speakers had a speech rate of approximately 173 words per minute (WPM). In comparison, speakers who undergo surgical-prosthetic voice restoration and use tracheoesophageal (TE) speech (Singer & Blom, 1980) to generate a pulmonary powered ‘esophageal’ speech signal, may approximate relatively normal speaking rates of ~127 to 138 WPM (Robbins et al., 1984; Trudeau & Qi, 1990). In contrast, the speech of individuals who use traditional esophageal speech (ES) that relies on the use of air that injected or insufflated into the esophagus (Van den Berg & Moolenaar-Bijl, 1959) is substantially reduced in rate from that of normal speakers. ES speakers may demonstrate speaking rates that range from 1.79 to 2.24 ($M = 2.01$) syllables per second or 99.1 to 114.3 WPM (Gandour, Weinberg, & Rutkowski, 1980; Hoops & Noll, 1969; Robbins et al., 1984; Snidecor & Curry, 1959). Finally, individuals who use an electrolarynx (EL), which involves use of an external, electronic voicing source that is placed against the neck, have demonstrated speech rates of approximately 130 WPM, one that is within the normal range (Hillman, Walsh, Wolf, Fisher, & Hong, 1998). Given the variability in speech rate among alaryngeal speakers, it is important to understand how the modification of speech rate through the use of CS can potentially influence acoustic characteristics at the phonetic and word-level (Lindblom,

1963; 1964, Miller et al., 1986; Miller & Volaitis, 1989; Kessinger & Blumstein, 1997; Picheny et al., 1986; Theodore et al., 2009).

Numerous authors have acknowledged wide variability in the acoustic characteristics both among and between consonants and vowels produced by laryngeal and alaryngeal speakers (Allen, Miller, & DeSteno, 2003; Christensen & Weinberg, 1984; Doyle, Danhauer, & Reed, 1988; Gandour et al., 1980; Hillebrand et al., 1995; Meltzner & Hillman, 2005; Peterson & Barney, 1952; Robbins et al., 1984; Sacco, Mann, & Schultz, 1967; Sisty & Weinberg, 1972; Weiss & Basili, 1985; Weiss, Yeni-Komshian, & Heinz, 1979). Further, research has indicated that various acoustic features (i.e., temporal features, frequency) change when speaking rate is modified (Picheny et al., 1986; Theodore et al., 2009). A reduction in speaking rate, for example, has been shown to increase phoneme and syllable durations for normal speakers (Kessinger & Blumstein, 1997; Miller et al., 1986; Theodore et al., 2009). These increases in phoneme and syllable durations consequently contribute toward longer word and utterance durations, which have implications for the perception of specific phonemes (Miller & Volaitis, 1989). Since allophones of phonemes can have their own unique set of acoustic characteristics, violation of these features and their distinctions are seen frequently in alaryngeal speech (e.g., voicing errors). Therefore, it is important to consider the potential influence of both a vocal tract that is altered following TL, in addition to its interaction with an alaryngeal voice source alter on the acoustic characteristics of speech postlaryngectomy.

The acoustic differences in alaryngeal speech have been shown to occur as a result of the interplay between the new alaryngeal voice source and vocal tract characteristics following TL. Voicing errors and alterations in vowel durations have been

reported for distinct groups of alaryngeal speakers (Christensen et al., 1978; Christensen & Weinberg, 1976; Doyle et al., 1988; Gandour et al., 1980; Jongmans, Hilgers, Pols, & van As-Brooks, 2006). While ES speakers often produce more voicing and durational errors than TE speakers, the linguistic rules governing vowel duration are relatively maintained. This is partially due to the fact that ES and TE are ‘intrinsic’ methods of alaryngeal speech. In contrast, EL speech alterations may be the result of the speaker’s use of an externally-based, electronic and continuously voiced alaryngeal source (Weiss et al., 1979). Additionally, it is important to acknowledge that TL results in a reduced effective length a vocal tract (Diedrich & Youngstrom, 1966). This reduction in the effective length of the vocal tract has been shown to increase formant frequencies for ES speakers (Sisty & Weinberg, 1972). Although no direct evidence exists to indicate that a reduction in vocal tract length results in similar frequency formant changes for EL speech, similar changes would be expected regardless of the type of alaryngeal voicing source due to TL. Research has documented, however, the difficulty of transferring the EL signal across neck tissue and the resulting neck-transfer function in EL speakers (Meltzner, Kobler, & Hillman, 2003). As a result, such changes following TL also may impact EL speech. Therefore, individuals who use an EL may potentially face challenges in the acoustic structure (e.g., frequency) of their speech due to the unique use of an external voice source that must interact with modified neck tissue.

Overall, research has clearly identified that EL voice and speech are acoustically and perceptually different than ES and TE speech (Doyle & Eadie, 2005; Meltzner & Hillman, 2005; Weiss et al., 1979; Yeni-Komshian, Weiss, & Heinz, 1975). Unfortunately, minimal research has investigated the durational properties of speech

sounds and words for EL speech. One important consideration moving forward is that, if the linguistic rules governing vowel duration are preserved in alaryngeal speech, then EL speakers could experience a significant increase in word and vowel durations similar to findings previously observed in CS research. In addition, due to the documented relationship between articulatory movement (i.e., increased mouth opening) and formant frequencies (Stevens & House, 1955), the potential impact of CS on formant structure in EL speech must be explored. That is, if EL speakers increase oral cavity size while slowing their rate during CS, then subsequent changes in the vowel space and resultant formant frequency characteristics should occur. Therefore, the purpose of this study was to determine the potential impact of CS on the duration of words and their intrinsic vowel component, in addition to altering the fundamental frequency and formant frequency characteristics of vowels produced by EL speakers.

Method

Speakers

Ten adult males ($M_{age} = 74$ years; $range = 59-87$ years) who underwent TL participated as speakers. All speakers were at least 24 months ($M = 133$ months; $range = 24-300$ months) postlaryngectomy and used an EL device as their primary method of alaryngeal speech since their TL. Seven speakers had a neck dissection as part of their TL. In addition, all speakers received radiation therapy (RT) either before ($n=4$), after ($n=5$), or before and after TL ($n=1$). Two speakers received CCRT before ($n = 1$) and after ($n=1$) TL. Speakers indicated that they were in good general health with no known neurological, medical or psychological conditions. Although no formal hearing screening was performed, every speaker reported no known hearing difficulties. However, given the

age and previous medical history, some level of hearing loss cannot be ruled out. All speakers indicated that English was their native language. Lastly, all ten speakers were the same participants for Experiment 1.

Every speaker used an EL device as their primary method of alaryngeal communication method. As part of their participation in the current study, speakers were asked to bring their own EL device to each recording session. In total, there was an equal representation of two commercially available EL devices across the speakers with five using a Servox Digital (Servona GmbH, Troisdorf, Germany) and five using a TruTone device (Griffin Laboratories, Temecula, CA) device.

Speech Stimuli

A list of 18 monosyllabic English words, 17 with a consonant-vowel-consonant (CVC) structure and one with a CV structure, served as speech stimuli. Words containing consonants in WI and word-final (WF) position were selected from a larger, 66-word list created by Weiss and Basili (1985) (see Appendix A). This larger list of words was modified to ensure an equal representation of each consonant in WI and WF position. In total, six plosives (/p/, /b/, /t/, /d/, /k/ and /g/), seven fricatives (/f/, /v/, /s/, /z/, /ʃ/, /θ/ and /ð/), two affricates (/tʃ/ and /dʒ/), and two nasals (/m/ and /n/) were represented in the 18 stimulus items. Sixteen of the 18 stimuli contained target phonemes in word-initial *and* word-final position, and two additional words (i.e., ‘know’ and ‘loathe’) were included to represent the word-initial nasal (e.g., /n/) and the word-final (e.g., ð), voiced dental fricative. Finally, a total of six vowels (/i/, /ɪ/, /ɛ/, /æ/, /eɪ/, and /oʊ/) were represented in the word list, although unevenly distributed due to the use of real word stimuli.

Data Acquisition

Recording of speech stimuli. Speech stimuli were recorded in a quiet room free of background noise. Recordings were obtained immediately after providing informed consent (Western University Ethics Research Board Approval #105382) (see Appendices B and C) and the collection of demographic information (see Appendix D) from each speaker. The recordings began with placement of a unidirectional microphone (Shure PG-81, Niles, IL) that was placed approximately 15 cm above each speaker's mouth at a 45-degree angle. The microphone was attached to a pre-amplifier (M-Audio, Avid Technology, Burlington, MA) and laptop computer (Dell Inspiron, Round Rock, TX) with SonaSpeech II software (KayPentax, Lincoln Park, NJ). A sampling rate of 44.1 kHz was used for all recordings. Volume levels were adjusted manually before each recording session and also were monitored using the sound meters in SonaSpeech II during recording to prevent over- or under-driving the input signal.

The same ten speakers from Experiment I were provided with a print list of the 18 words and provided with the following instructions: "Please take a moment to look over the words. Once you are ready, please read each word. If you make a mistake, I will ask you to repeat the word(s) once you finish reading". This was referred to as the habitual speech (HS) condition. Once the word list was recorded in HS, each speaker was next provided with instructions to read the same word list using clear speech (CS). In order to produce CS participants were asked, "Now I would like you to re-read the words by speaking as clearly as possible. This will involve slowing down while speaking and over-articulating" (Picheny et al., 1985). Every participant speaker rehearsed reading words using this style of speaking prior to recording. Therefore, each speaker was required to first

read the word list in HS and then re-read the word list a second time list using CS. This method of not counter-balancing sessions was deliberately used to control for potential carryover effects from the experimental (i.e., CS) speaking condition had that been recorded first. All recording sessions lasted approximately 20 minutes.

Editing word stimuli. After all 10 speakers provided their recording in HS and CS, two separate audio files containing 36 words each (i.e., 18 words in HS and 18 words in CS) were edited using Audacity 2.0.5 (Mazzoni & Dannenberg, 2013). Recording noise on each audio file was removed using the ‘Noise Removal’ tool within Audacity. This was completed by highlighting a small window of silence (i.e., non-speech recording), obtaining the noise profile in Audacity, and then allowing the software to remove any audible track noise. Speech stimuli were not altered as a result of this process. Individual words on each sound file were then highlighted, copied and then pasted into new audio tracks and saved as individual audio files in .wav format. After editing, there was a total of 360 audio files composed of single words [18 words x 10 speakers x 2 speaking conditions].

Acoustic Analysis

Acoustic analysis centered on objectively measuring several acoustic characteristics of words (i.e., durations) and vowels (i.e., durational measures, fundamental frequency, and formant frequencies) for stimuli produced in both HS and CS conditions. All analyses were conducted using Version 5.4.17 of Praat (Boersma & Weenink, 2015). In order to ensure a reliable and accurate measurement method, a combination of careful visual inspection of spectrograms (e.g., voicing, intensity, and

formant patterning) and waveforms, in addition to auditory playback were maintained for all stimuli.

Duration Measurement

Overall word duration. Overall word durations for HS and CS stimuli were computed by measuring the entire word duration from the beginning to the end of each recorded word. Each edited audio file containing a single stimulus word was opened in Praat (Boersma & Weenink, 2015). The beginning of the word was selected where EL speakers turned on their device. This window was lengthened until the end of the word, indicated by a termination of EL device and visual confirmation of no further speech sound production (e.g., release burst or frication). The time of the highlighted window was recorded in milliseconds (ms).

Overall vowel duration. Four monophthongs (/i/, /ɪ/, /ɛ/, and /æ/) and two diphthongs (/eɪ/ and /oʊ/) were represented in the list of 18 stimulus words. Measurement of vowel duration for the monophthongs and diphthongs began at the first zero crossing after the WI stop release involving steady-state vowel formant patterning. The entire steady vowel was highlighted and ended at a zero crossing where there was a lack of steady state vowel formant pattern. After the highlighted area was selected, the duration provided by Praat was recorded in ms.

Fundamental frequency. Fundamental Frequency (F_0) data were collected for non-speech and EL speech data using Praat. The mean F_0 was obtained for coupled (i.e., device on neck with mouth open), non-speech ELs signals for every participant speaker. To do this, frequency measurements were taken using the same, randomly chosen word for each speaker (i.e., ‘catch’). Frequency measurements were obtained during a time

interval within the stop gap of the WF affricate. After listening to the selected area to confirm only EL noise was present, the investigator clicked in the center of the time interval and selected 'Pitch' and 'Get pitch' from the Praat toolbar. The F_0 values were confirmed by a blue, pitch contour shown in the Praat object window.

To obtain F_0 -related measurements for vowels, the investigator began by displaying the blue pitch line on the spectrogram window within Praat by selecting 'Pitch' > 'Show pitch' in the menu. Next, selecting the middle of on the blue line (which is located in the middle of the phoneme) produces an estimate of the F_0 . Additionally, the blue line is time-linked to the spectrogram, further permitting identification of the temporal mid-point of the vowel. To confirm this, the investigator selected 'Pitch' > 'Get pitch' from the Praat menu to obtain the F_0 for the selected data point, and the F_0 in Hertz (Hz) was recorded. Similarly, formant data were obtained by identifying and selecting the middle of the vowel and selecting 'Formant' > 'Get first formant'. This procedure was repeated for the second and third formants (i.e., 'Get second formant' and 'Get third formant'); data were then extracted and entered into a database for later statistical analysis.

Data Analyses

Word duration. A repeated measures analysis of variance (ANOVA) was used to assess the effect of speaking condition and EL device on word duration. Specifically, statistical comparisons were conducted on overall word durations between HS and CS, followed by overall word durations within device groups comparisons (e.g., Servox Digital HS vs. CS, TruTone HS vs. CS), and then overall word durations between device groups across speaking conditions (e.g., Servox Digital HS vs. TruTone HS). In addition,

the magnitude of effect for speaking condition was determined by calculating partial eta squared. Interpretation of effect size followed guidelines by Cohen (1988), which includes 0.01 (small effect), 0.06 (medium effect), and 0.14 (large effect). An a priori significance level was set at $p < .05$ for all statistical analyses. A Bonferroni correction was used for post-hoc testing.

Vowel duration. Overall vowel durations were analyzed using a repeated measures ANOVA. Specifically, analyses were conducted between speaking conditions, followed by comparisons of overall vowel durations within device groups, and overall vowel durations between device groups. This was followed by comparisons of overall frequency characteristics of vowels between speaking conditions, within device group and between device group comparisons. Once again, the magnitude of effect for each analysis was determined by calculating partial eta squared, and the interpretation of effect size followed guidelines by Cohen (1988) (e.g., 0.01 for a small effect, 0.06 for a medium effect, and 0.14 for a large effect). A Bonferroni correction was used for post-hoc testing. An a priori significance level was set at $p < .05$ for all statistical analyses.

Results

Whole-Word Stimuli

Overall word duration. Mean overall word durations for EL speakers in HS and CS are shown in Table 3.1 and represented graphically in Figure 3.1. The mean overall durations for the 17 CVC-words spoken by EL speakers were 596 ms ($SD = 112$ ms; $range = 462-736$ ms) in HS and 653 ms ($SD = 133$ ms; $range = 497-817$ ms) in CS. Overall, the mean durations for 17 CVC-words were found to be longer in CS compared to HS. Results from a repeated measures ANOVA indicated that there was a significant

Table 3.1

Overall Mean Word Durations for Electrolaryngeal (EL) Speakers During Habitual Speech (HS) and Clear Speech (CS)

Words	EL Speakers							
	HS				CS			
	M	SD	Low	High	M	SD	Low	High
Catch	579	145	440	877	641	159	466	948
Mass	685	118	490	876	765	147	578	1000
Pad	605	201	370	986	629	185	374	930
Sack	628	126	398	794	672	152	446	926
Dab	490	126	339	747	565	172	387	877
Teeth	504	112	386	721	543	122	379	774
Jeep	467	103	328	632	538	173	350	958
Shave	728	186	447	1129	817	196	618	1169
Zag	648	192	437	1102	748	247	474	1233
Badge	728	187	494	1052	761	213	576	1259
Gain	534	122	336	702	587	86	468	692
Vet	464	133	330	767	498	151	273	739
Chief	522	128	322	757	578	123	388	834
These	702	118	521	943	771	130	547	960
Fish	538	97	342	673	557	121	359	750
Theme	575	125	339	765	630	158	326	852
Know	551	160	356	904	629	176	482	1045
Loathe	736	203	486	1163	799	208	543	1210

Note. Duration data are in milliseconds (ms).

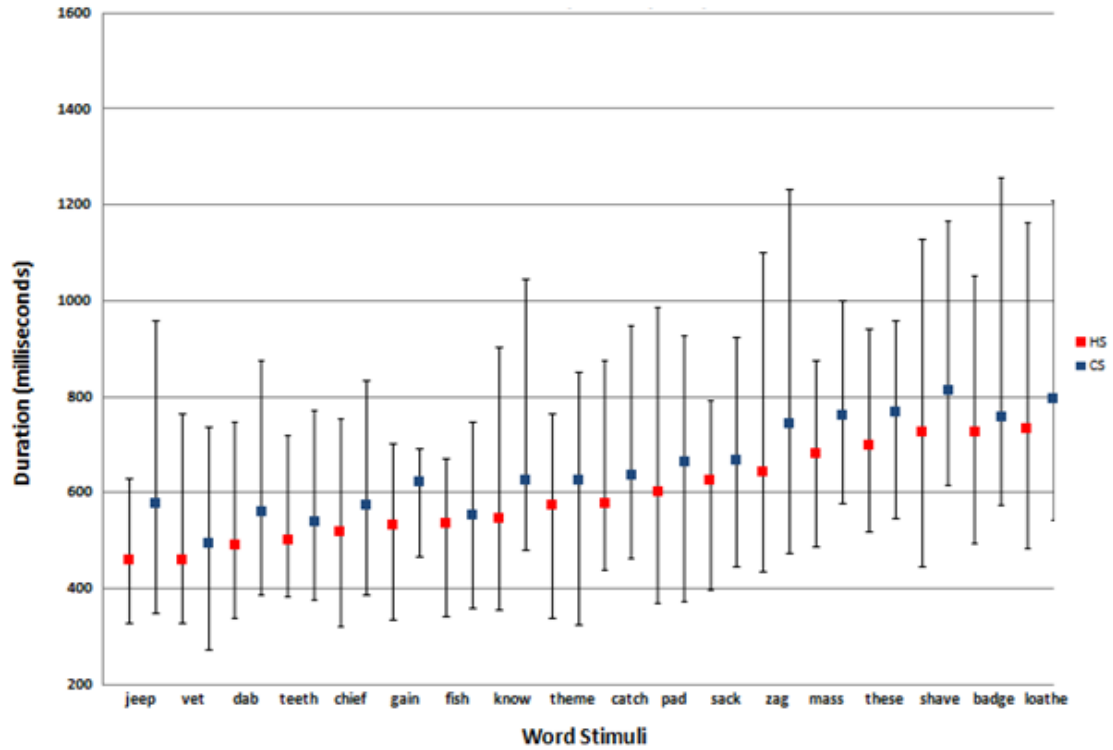


Figure 3.1. Overall word durations and ranges by electrolaryngeal speakers. Words are arranged from shortest to longest duration. Duration data are in milliseconds (ms). *HS* = habitual speech; *CS* = clear speech.

effect of speaking condition on overall word duration, $F(1, 8) = 17.310$, $p < .01$, partial $\eta^2 = .684$. Speaking condition was deemed to have a large effect on word durations. Post-hoc testing indicated that overall word durations were significantly longer in CS compared to HS ($p < .05$).

The mean overall durations for the single CV-word (e.g., 'know') were 551 ms ($range = 356- 904$ ms) in HS and 629 ms ($range = 482$ ms-1.05 s) in CS. This difference was found to be statistically significant using a repeated measures ANOVA, $F(1,8) = 13.965$, $p < .01$, partial $\eta^2 = .636$. The magnitude of the effect indicates that speaking condition demonstrated a large effect on all word durations. Post-hoc testing indicated that durations of 'know' were significantly longer in CS compared to HS ($p < .05$).

In order to assess the potential influence of EL device on overall word durations, data for Servox Digital and TruTone speakers are shown in Tables 3.2a and 3.2b, respectively. Further, data are also presented according to Servox and Trutone in Figures 3.2a and 3.2b, respectively. For Servox speakers, all 17 CVC-words were spoken 62 ms slower during CS ($M = 682$ ms; $range = 513- 880$ ms) compared to HS ($M = 620$ ms; $range = 423-811$ ms). The single CV-word (e.g., 'know') was 109 ms longer in duration when Servox speakers used CS ($M = 642$ ms; $range = 482$ ms-1.05 s) compared to HS ($M = 533$ ms; $range = 356-904$ ms). Repeated measures ANOVA revealed no significant effect of speaking condition on the duration of words spoken by Servox speakers.

On average, TruTone users produced the 17 CVC-words 52 ms slower when they used CS ($M = 624$; $range = 483-776$ ms) compared to HS ($M = 572$ ms; $range = 440-702$

Table 3.2a

Overall Mean Word Durations for Servox Speakers During Habitual Speech (HS) and Clear Speech (CS)

Words	Servox							
	HS				CS			
	<i>M</i>	<i>SD</i>	<i>Low</i>	<i>High</i>	<i>M</i>	<i>SD</i>	<i>Low</i>	<i>High</i>
Catch	627	196	440	877	680	199	466	948
Mass	668	147	490	876	794	203	578	1000
Pad	668	267	370	986	681	235	443	930
Sack	626	153	398	770	729	202	446	926
Dab	540	162	339	747	589	178	434	877
Teeth	523	140	386	721	533	121	379	716
Jeep	453	108	328	625	583	216	454	958
Shave	770	257	447	1129	880	225	632	1169
Zag	675	269	437	1102	803	345	474	1233
Badge	798	227	591	1052	852	260	634	1259
Gain	555	115	442	702	583	82	507	689
Vet	423	67	359	499	513	175	273	739
Chief	566	173	322	757	588	164	388	834
These	712	159	521	943	765	158	547	912
Fish	541	126	342	673	552	145	359	750
Theme	576	154	339	765	639	218	326	852
Know	533	215	356	904	642	233	482	1045
Loathe	811	238	555	1163	825	251	543	1210

Note. Duration data are in milliseconds (ms).

Table 3.2b

Overall Mean Word Durations for TruTone Speakers During Habitual Speech (HS) and Clear Speech (CS)

Words	TruTone							
	HS				CS			
	<i>M</i>	<i>SD</i>	<i>Low</i>	<i>High</i>	<i>M</i>	<i>SD</i>	<i>Low</i>	<i>High</i>
Catch	530	54	476	601	601	116	496	780
Mass	702	96	555	797	736	73	620	806
Pad	542	99	395	618	576	120	374	682
Sack	631	111	492	794	614	54	548	672
Dab	440	55	383	522	541	183	387	843
Teeth	485	87	421	638	552	135	404	774
Jeep	480	109	375	632	493	125	350	618
Shave	685	83	572	778	753	161	618	1024
Zag	621	93	491	712	693	106	603	870
Badge	658	121	494	796	671	118	576	874
Gain	514	139	336	699	591	100	468	692
Vet	504	178	330	767	483	141	307	691
Chief	478	48	427	554	568	81	463	655
These	692	74	611	785	776	113	676	960
Fish	535	73	438	612	562	109	442	687
Theme	575	108	428	721	620	94	503	762
Know	569	103	448	733	615	122	498	791
Loathe	661	149	486	826	773	181	586	1040

Note. Duration data are in milliseconds (ms).

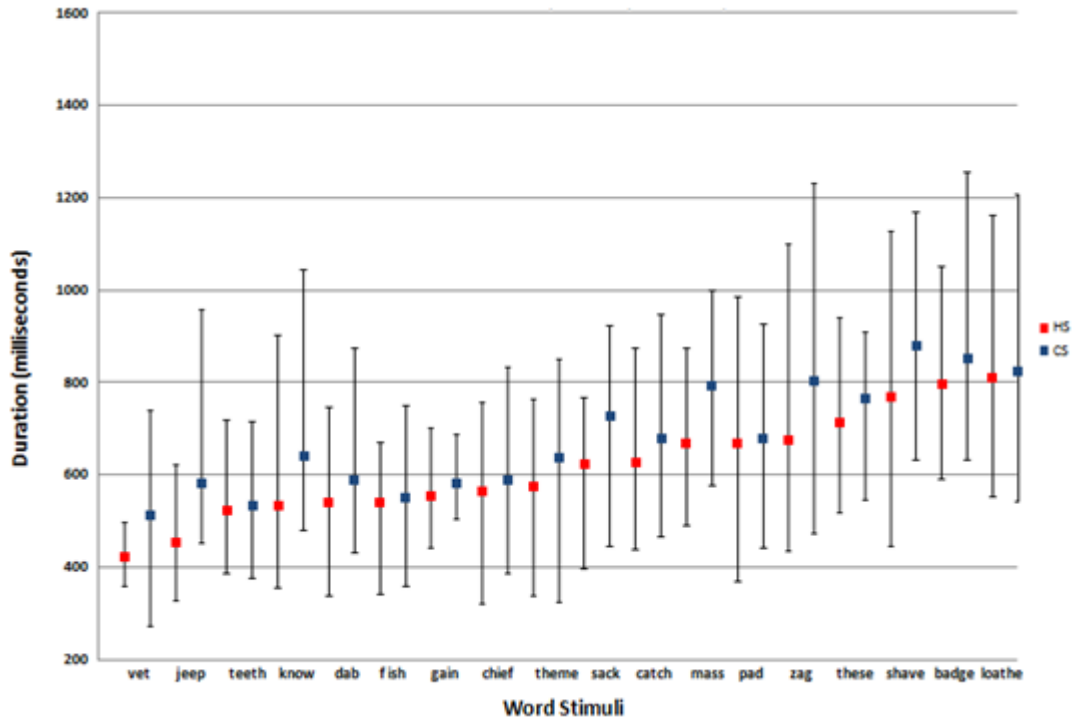


Figure 3.2a. Mean word durations by Servox speakers. Words arranged from shortest to longest duration. Duration data are in milliseconds (ms). *HS* = habitual speech; *CS* = clear speech.

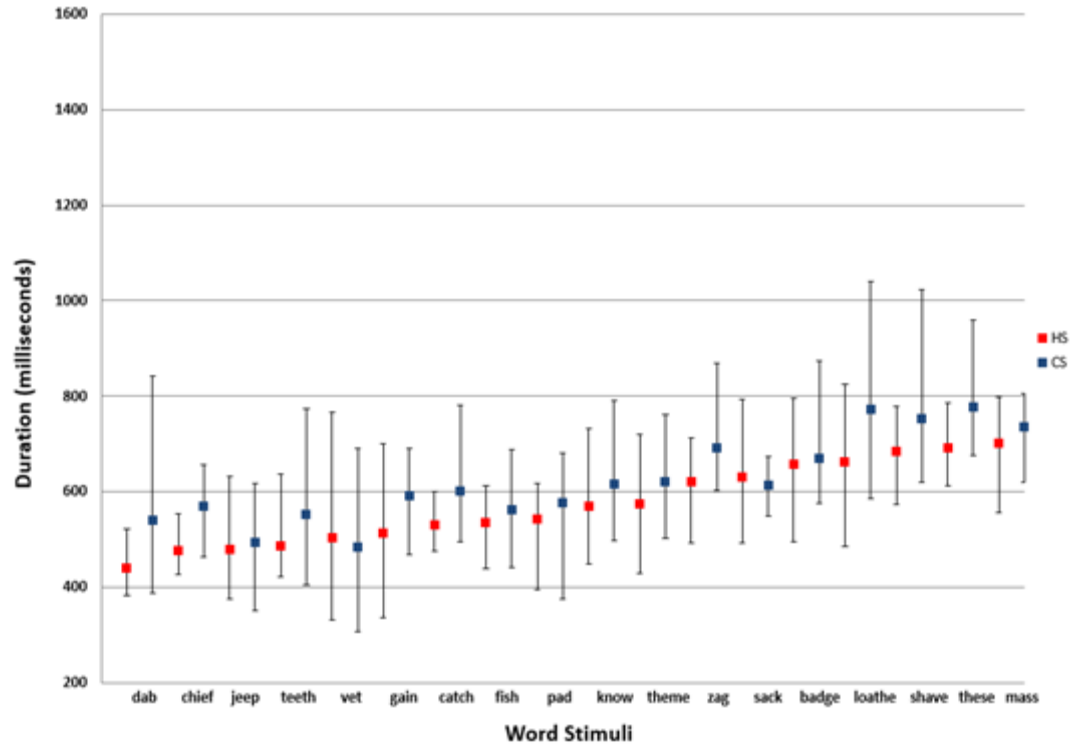


Figure 3.2b. Mean word durations by TruTone speakers. Words arranged from shortest to longest duration. Duration data are in milliseconds (ms). *HS* = habitual speech; *CS* = clear speech.

ms). The single CV word (e.g., ‘know’) was spoken 46 ms slower when TruTone speakers used CS ($M = 615$ ms; $range = 498-791$ ms) compared to HS ($M = 569$ ms; $range = 448-733$ ms). Repeated measures ANOVA revealed no significant effect of speaking condition on the duration of words spoken by Servox speakers.

Finally, analyses between Servox Digital and TruTone users indicated that there was no influence of device on word duration in HS and CS.

Vowel Stimuli

Overall vowel duration. Mean overall vowel durations for EL speakers in HS and CS are shown in Table 3.3. The mean overall durations for vowels within the 17 CVC-words spoken by EL speakers were 333 ms ($SD = 76$ ms) in HS and 354 ms ($SD = 71$ ms) when using HS and CS, respectively. Results from the repeated measures ANOVA indicated that there was a significant effect of speaking condition on overall vowel duration, $F(1,8) = 12.149$, $p < .01$, partial $\eta^2 = .603$. The magnitude of the effect indicated that speaking condition demonstrated a large effect on overall vowel durations. Post-hoc testing indicated that overall vowel durations were significantly longer in CS compared to HS ($p < .05$).

The mean overall durations for the vowel in single CV-word (e.g., ‘know’) were 551 ms ($range = 356-904$ ms) in HS and 629 ms ($range = 482$ ms-1.05 s) in CS. Repeated measures ANOVA indicated a significant effect of speaking condition on the single CV-vowel duration, $F(1,8) = 9.127$, $p < .05$, partial $\eta^2 = .533$. The magnitude of the effect indicated that speaking condition demonstrated a large effect on vowel durations. Post-hoc testing indicated that the single CV-vowel duration was significantly longer in CS compared to HS ($p < .05$).

Table 3.3

Mean Vowel Durations for Electrolaryngeal (EL) Speakers During Habitual Speech (HS) and Clear Speech (CS)

Vowels	Servox							
	HS				CS			
	<i>M</i>	<i>SD</i>	<i>Low</i>	<i>High</i>	<i>M</i>	<i>SD</i>	<i>Low</i>	<i>High</i>
/i/	304	133	129	633	307	115	147	531
/ɪ/	205	35	150	247	231	73	108	300
/ɛ/	175	44	137	233	277	205	152	780
/æ/	370	139	189	723	407	145	212	769
/eɪ/	366	129	231	693	407	112	270	642
/oo/	429	145	226	651	467	152	249	780
Vowels	TruTone							
	HS				CS			
	<i>M</i>	<i>SD</i>	<i>Low</i>	<i>High</i>	<i>M</i>	<i>SD</i>	<i>Low</i>	<i>High</i>
/i/	315	82	153	461	329	114	175	643
/ɪ/	264	45	210	330	257	59	190	336
/ɛ/	184	47	120	240	193	42	142	248
/æ/	347	74	199	473	347	83	163	486
/eɪ/	411	69	276	495	430	89	278	548
/oo/	414	97	302	586	494	120	380	693

Note. Duration data are in milliseconds (ms).

On average, Servox users exhibited monophthong vowel durations that were 21 ms slower in CS compared to HS (*range* = 129-723 and 108-769, respectively). When the duration of diphthongs were assessed, they were produced 39 ms slower in CS (*range* = 249-780 ms in CS; *range* = 226-693 ms in HS). In addition, TruTone speakers produced monophthongs similarly in HS and CS conditions; there was a difference of 20 ms for monophthongs in CS when compared to HS (*range* = 142-643 ms and 125-633 ms, respectively). Diphthongs were produced approximately 50 ms slower in CS (*range* = 249-780 ms in CS; *range* = 226-693 ms in HS). Overall, there was not significant effect of EL device used on vowel durations spoken in HS or CS.

Non-speech fundamental frequency. Non-speech F_0 measurements were obtained for each device and these data are shown in Table 3.4. On average, Servox Digital users produced an average F_0 of 77.5 Hz (*range* = 46.7-88.4 Hz) while TruTone users produced an average F_0 of 87.8 Hz (*range* = 78.1 to 93.3 Hz). There was no significant effect of EL device used on non-speech device F_0 .

Vowel fundamental and formant frequencies. Overall F_0 and formant frequency data for vowels are shown in Table 3.5. Vowels were produced by Servox users with a mean F_0 of 77.7 Hz (*range* = 46.7-84.9 Hz) in HS and 77.6 Hz (*range* = 46.6 to 84.9 Hz) in CS; TruTone users exhibited a mean F_0 of 83.7 Hz (*range* = 63.3-97.1 Hz) in HS and 85.3 Hz (*range* = 80.0-104.7 Hz) in CS. Results from the repeated measures ANOVA indicated no significant influence of speaking condition or device on F_0 during the production of vowels.

Table 3.4

Fundamental Frequency of Non-Speech Data by Electrolaryngeal (EL) Speakers

Servox	Speaker 1	Speaker 5	Speaker 6	Speaker 8	Speaker 10	Average	Overall
(F₀ Hz)	83.7	46.7	83.8	88.4	84.8	77.5	
TruTone	Speaker 2	Speaker 3	Speaker 4	Speaker 7	Speaker 9	Average	
(F₀ Hz)	83.6	90.4	93.7	78.1	93.3	87.9	82.7

Note. All data are provided in hertz (Hz). F₀ = fundamental frequency.

Table 3.5

Average Fundamental and Formant Frequencies of Vowels Produced by Servox Digital and TruTone Speakers in Habitual Speech (HS) and Clear Speech (CS)

		Servox					
		/i/	/ɪ/	/ɛ/	/æ/	/eɪ/	/oʊ/
F0	HS	74.8	74.8	79.4	74.3	83.3	79.4
	CS	74.5	74.8	79.2	74.5	83.4	79.3
F1	HS	726.3	612.3	677.5	762.4	588.4	652.2
	CS	723.3	618.6	662.7	762.0	560.8	640.7
F2	HS	2274.1	2090.7	2051.9	1881.5	2073.9	1328.5
	CS	2148.1	2104.8	1959.1	1868.0	2032.5	1372.1
F3	HS	2798.4	2798.4	2671.2	2588.6	2748.7	2835.3
	CS	2851.0	2851.0	2534.0	2489.0	2777.0	2685.7
		Trutone					
		/i/	/ɪ/	/ɛ/	/æ/	/eɪ/	/oʊ/
F0	HS	83.0	82.8	84.6	81.0	85.7	85.4
	CS	85.4	84.4	85.0	83.1	87.2	86.8
F1	HS	616.3	548.0	674.4	728.1	591.6	626.8
	CS	619.4	583.3	718.4	753.6	571.7	628.1
F2	HS	2090.5	2023.0	1753.6	1852.8	2045.0	1132.0
	CS	2083.0	2020.2	1857.7	1882.7	2051.9	1203.2
F3	HS	2792.6	2691.0	2548.7	2474.7	2552.7	2633.1
	CS	2608.7	2699.7	2541.9	2516.8	2618.0	2626.4

Note. Frequency measurements are in Hertz (Hz). F₀ = fundamental frequency; F₁ = first formant; F₂ = second formants; F₃ = third formant.

F1/F2 relationship plots for monophthongs (see Figure 3.3a) and diphthongs (see Figure 3.3b) illustrate the relationship of and variation in formant frequencies across speaking condition for Servox Speakers. Formant data for monophthongs and diphthongs produced by TruTone users are shown in Figures 3.4a and 3.4b, respectively. In each figure, individual speaker productions were arbitrarily enclosed in a loop in an approach used previously (Peterson & Barney, 1952) with each loop containing more than 90% of the productions for a given vowel. Data in Figures 3.3a and 3.4a indicate considerable overlap between vowel formants. In addition, the tightly clustered data points in Figures 3.3a and 3.4a suggest some neutralization of vowels, especially as EL speakers move across speaking condition. This is supported by the acoustic data provided in Table 3.5 showing relatively similar frequency values for all vowels for F_0 through F_3 . Further, individual formant plots for each EL speakers' monophthongs and diphthongs are shown in Figures 3.5a and b through 3.14a and b.

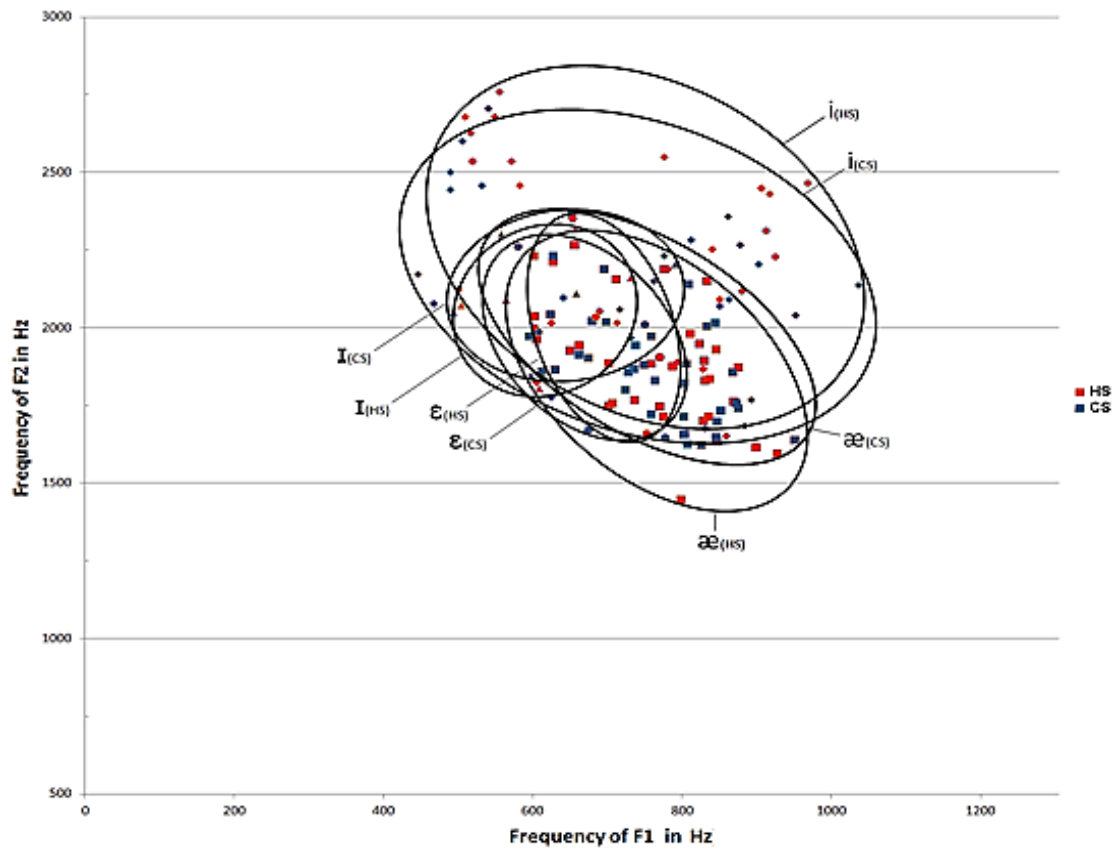


Figure 3.3a. F1/F2 plots of monophthongs produced by Servox speakers. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

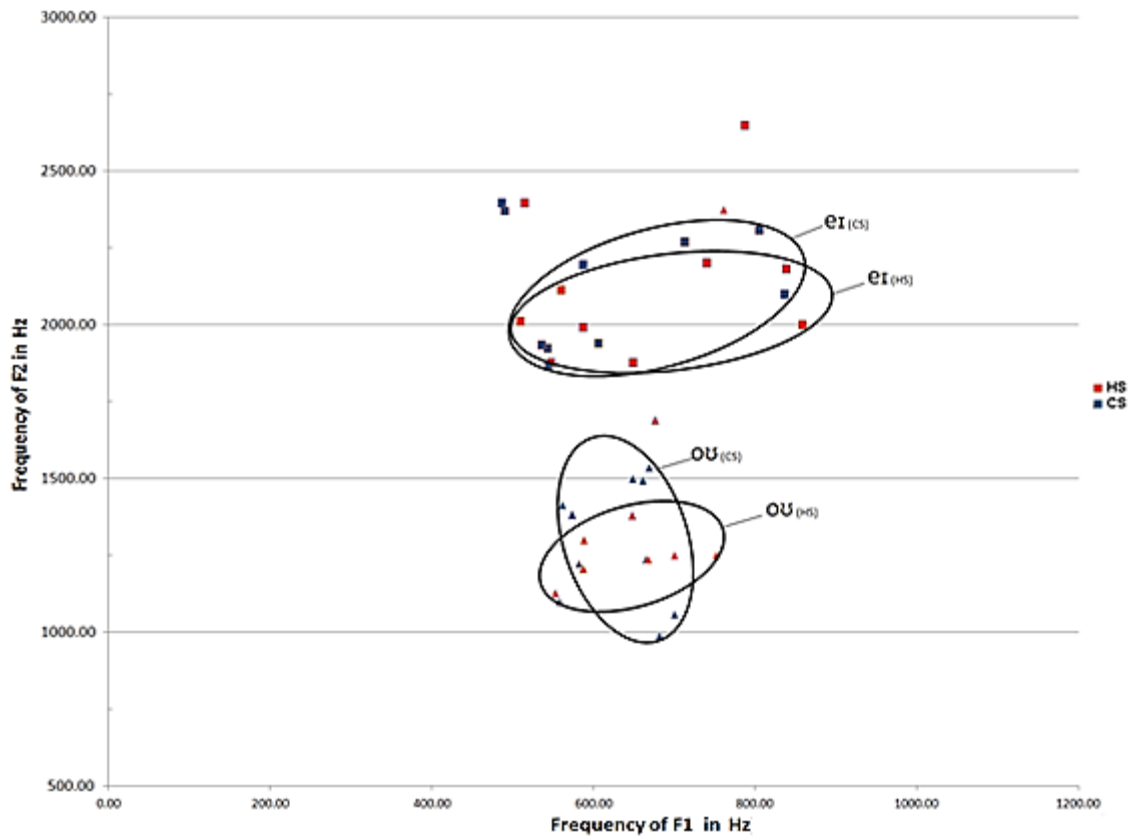


Figure 3.3b. F1/F2 plots of diphthongs produced by Servox speakers. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

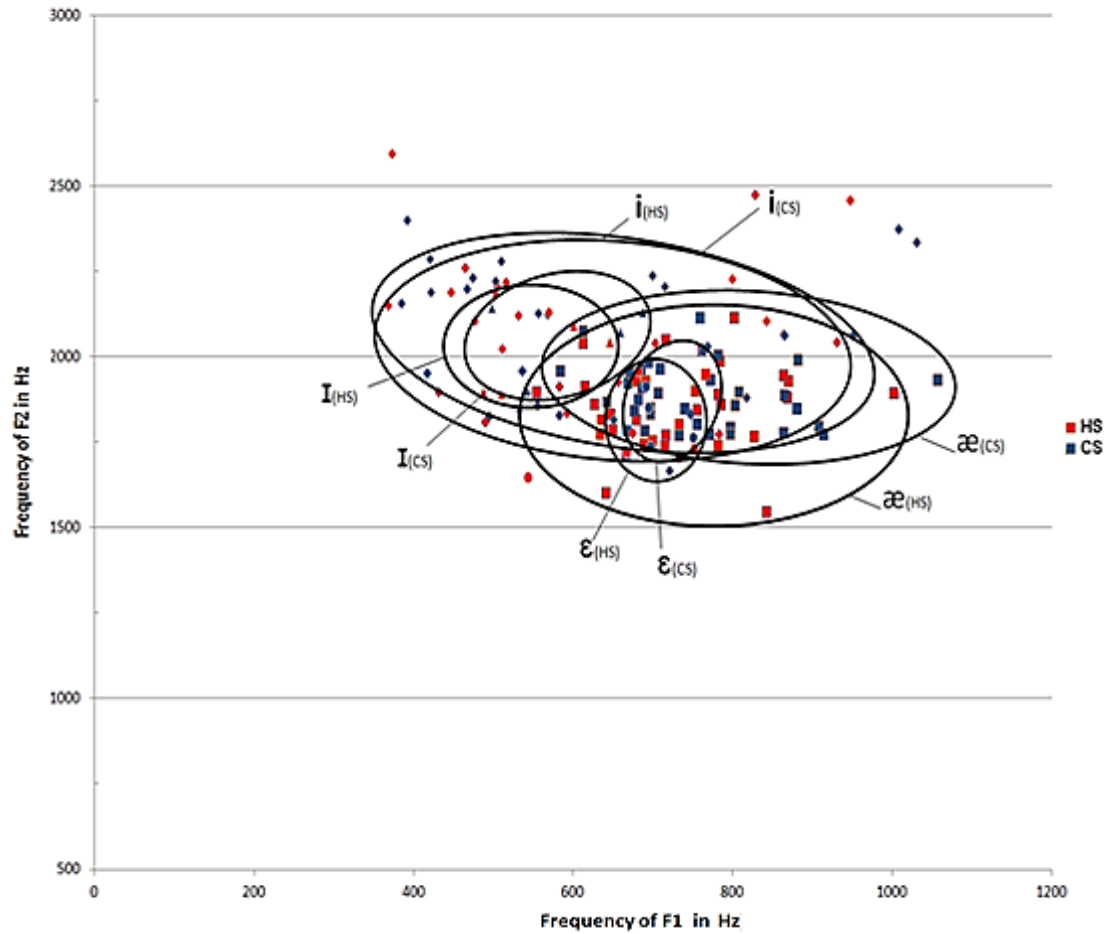


Figure 3.4a. F1/F2 plots of monophthongs produced by TruTone speakers. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

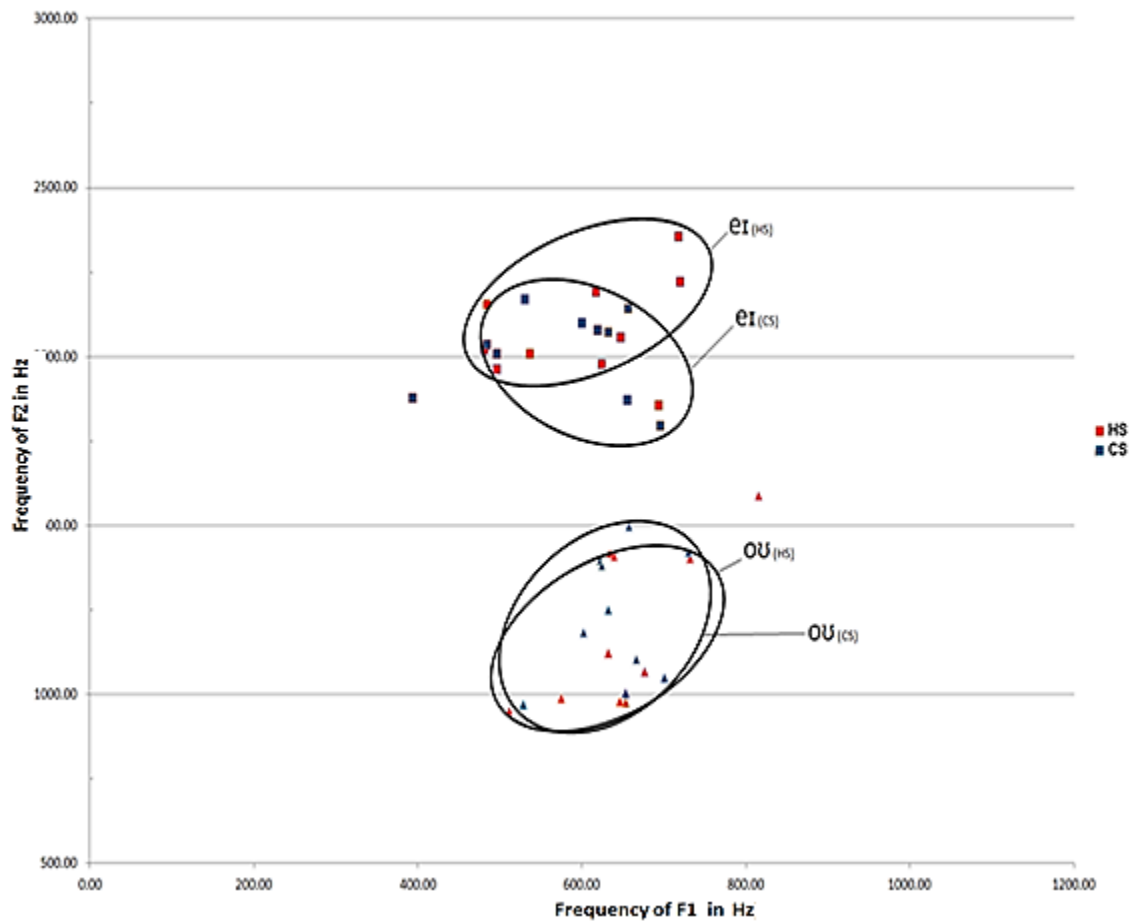


Figure 3.4b. F1/F2 plots of diphthongs produced by TruTone speakers. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

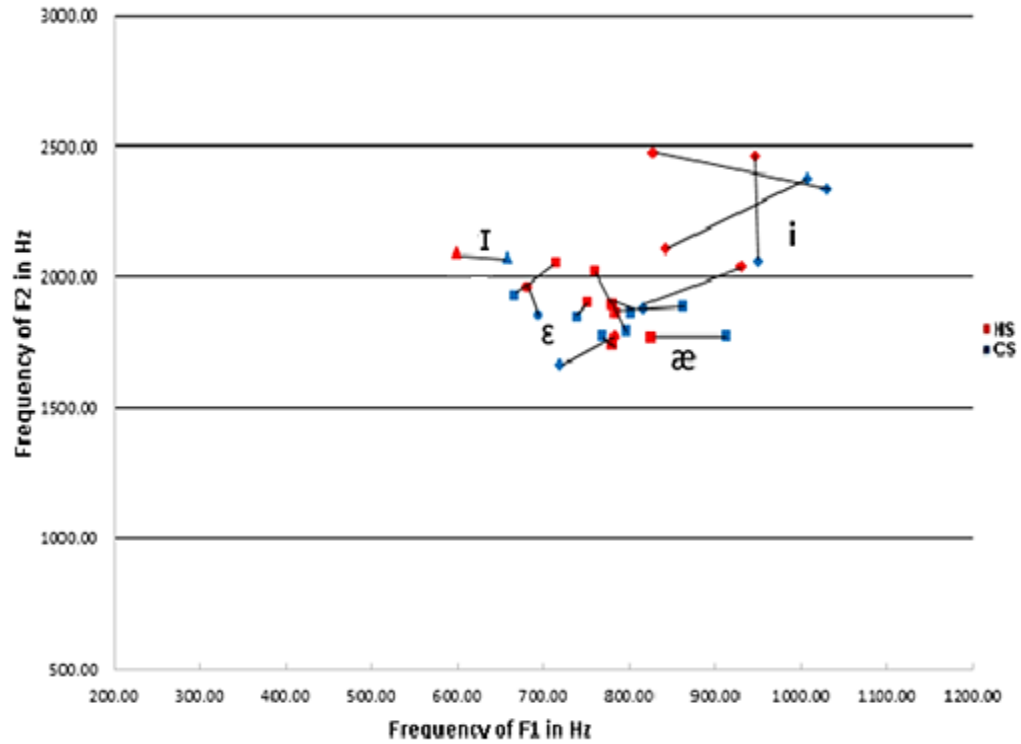


Figure 3.5a. F1/F2 plots of monophthongs produced by Speaker 1. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

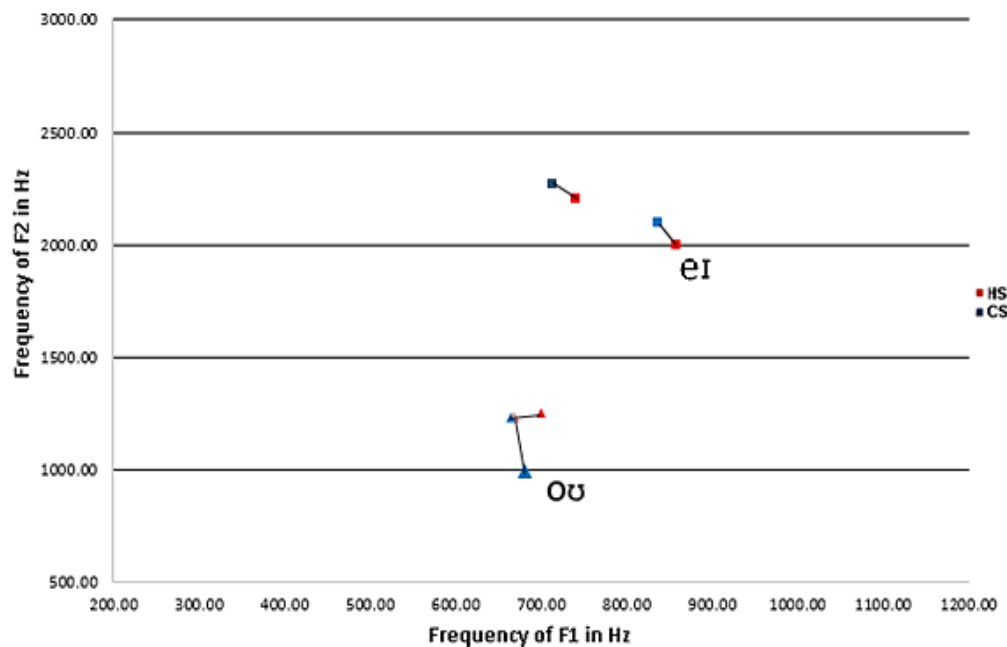


Figure 3.5b. F1/F2 plots of diphthongs produced by Speaker 1. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

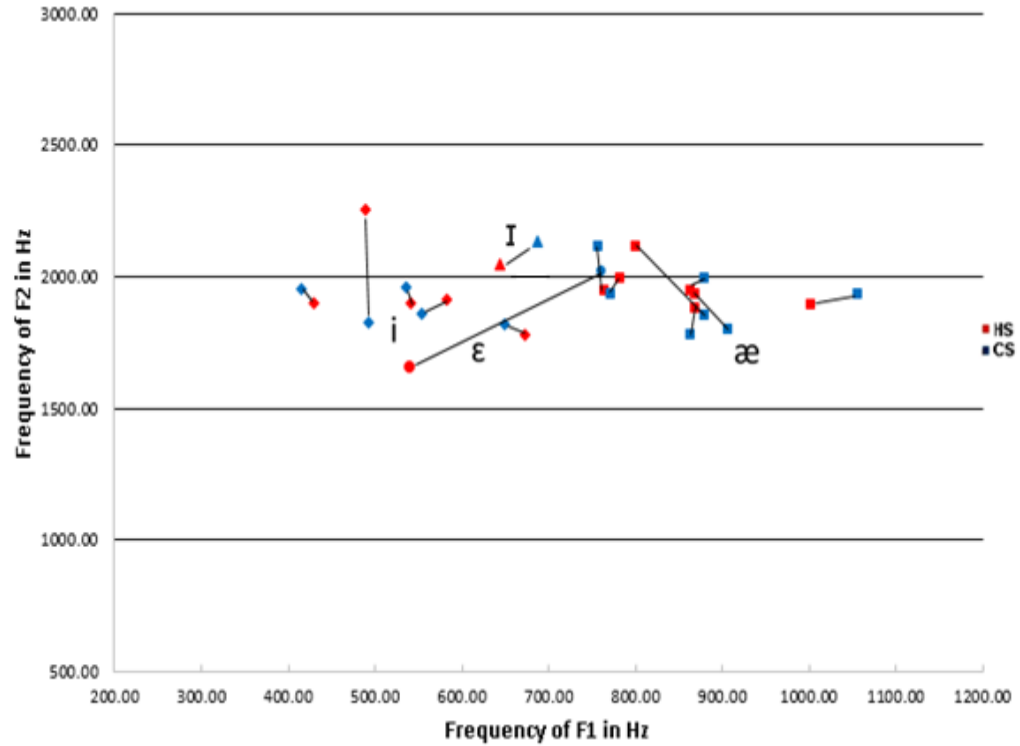


Figure 3.6a. F1/F2 plots of monophthongs produced by Speaker 2. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

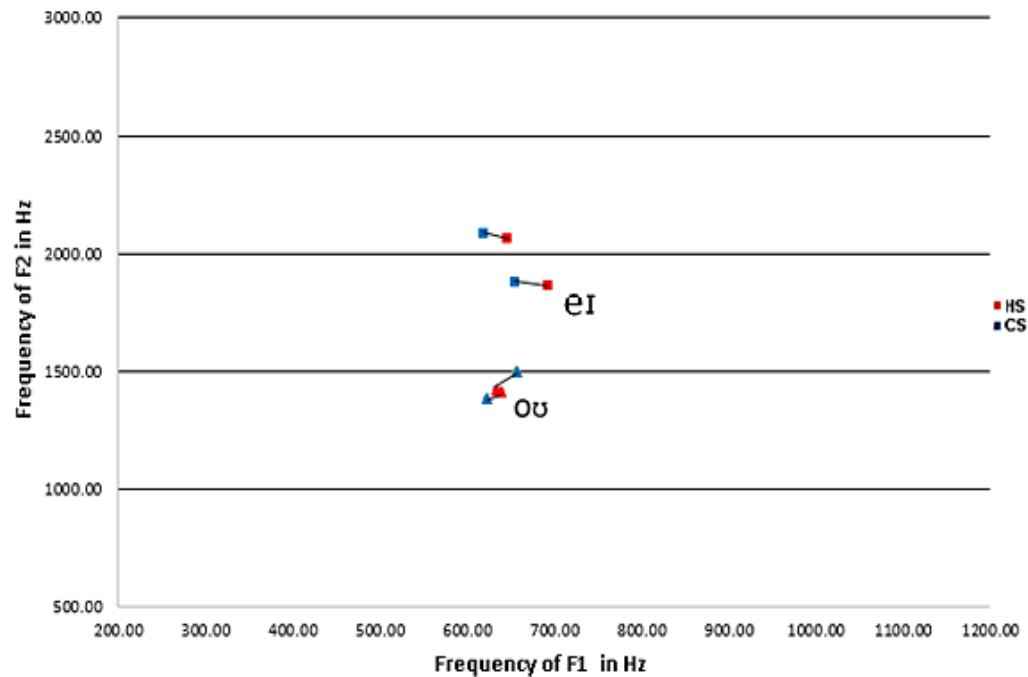


Figure 3.6b. F1/F2 plots of diphthongs produced by Speaker 2. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

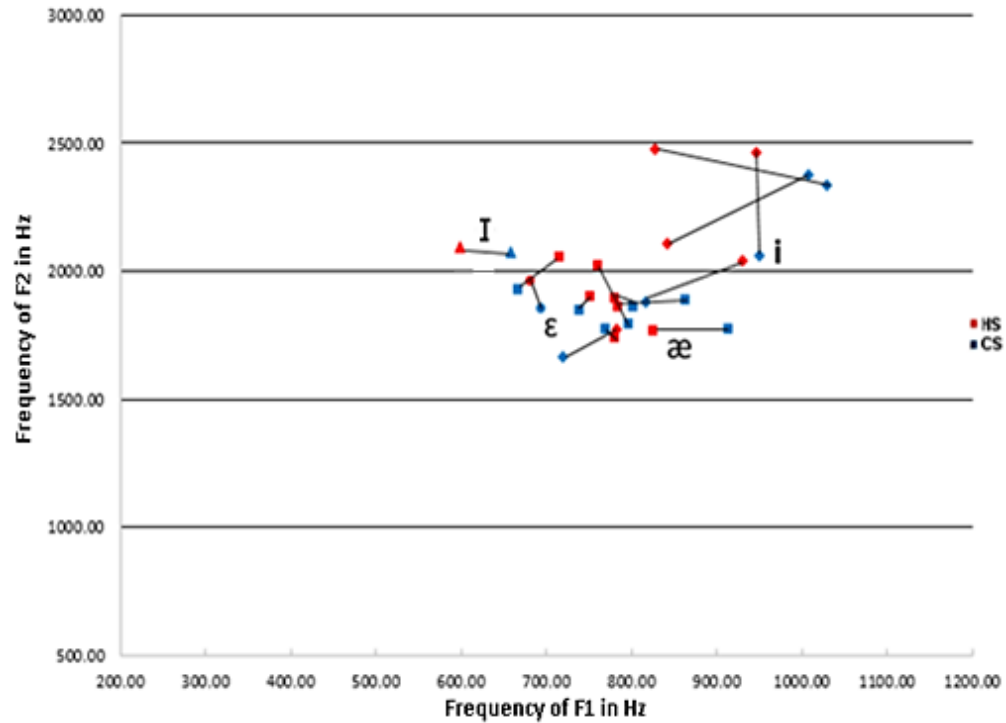


Figure 3.7a. F1/F2 plots of monophthongs produced by Speaker 3. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

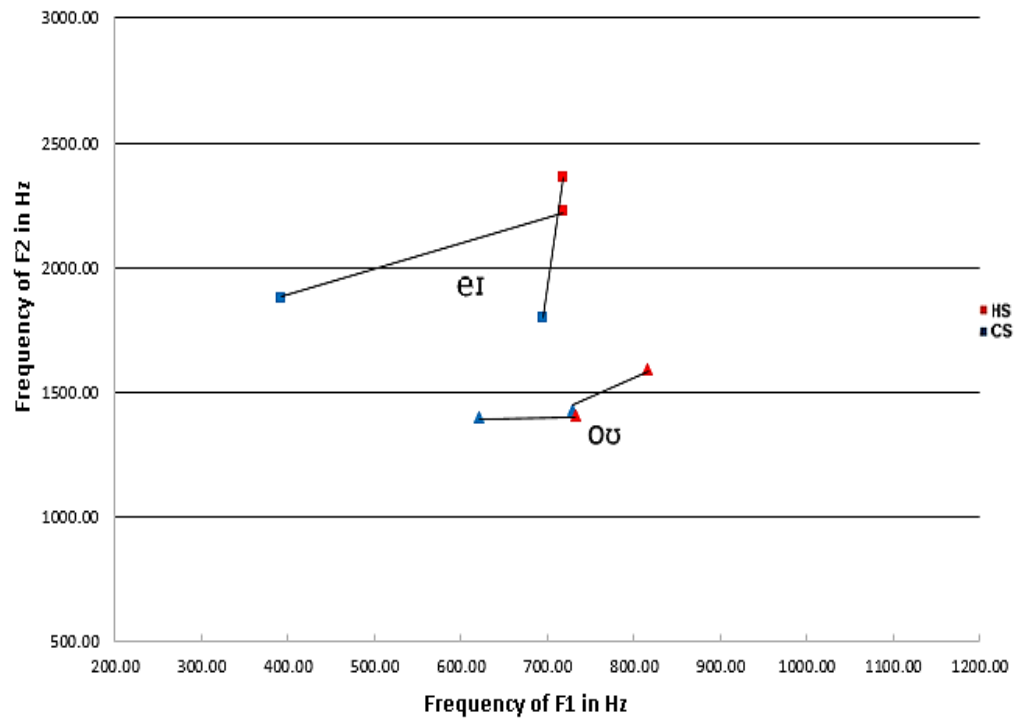


Figure 3.7b. F1/F2 plots of diphthongs produced by Speaker 3. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

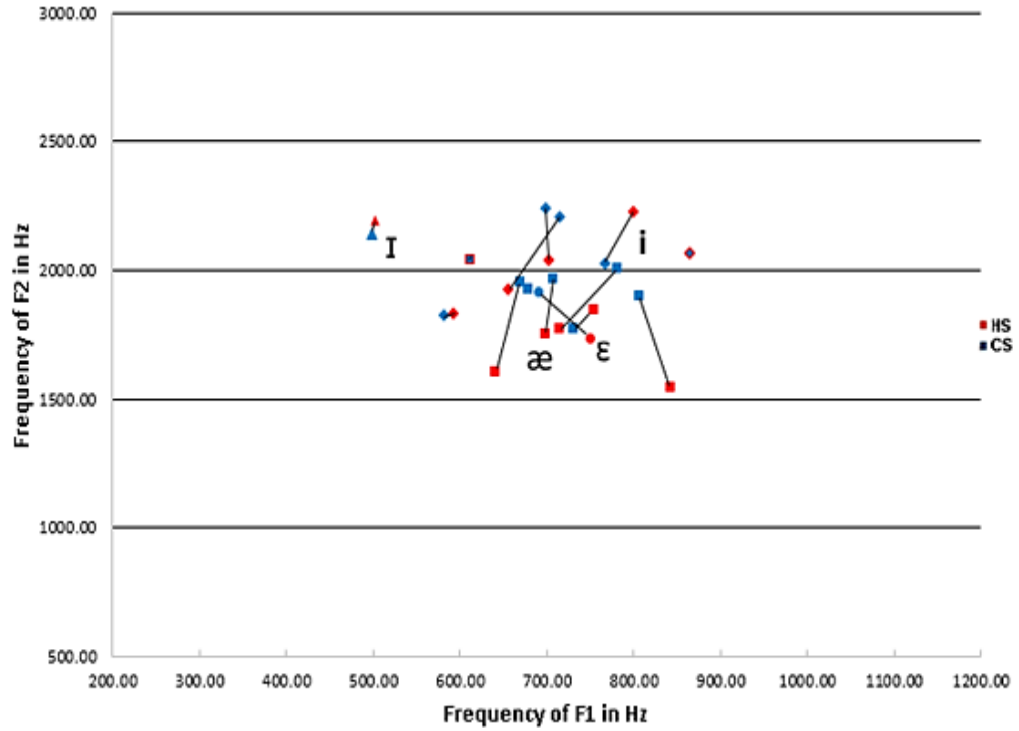


Figure 3.8a. F1/F2 plots of monophthongs produced by Speaker 4. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

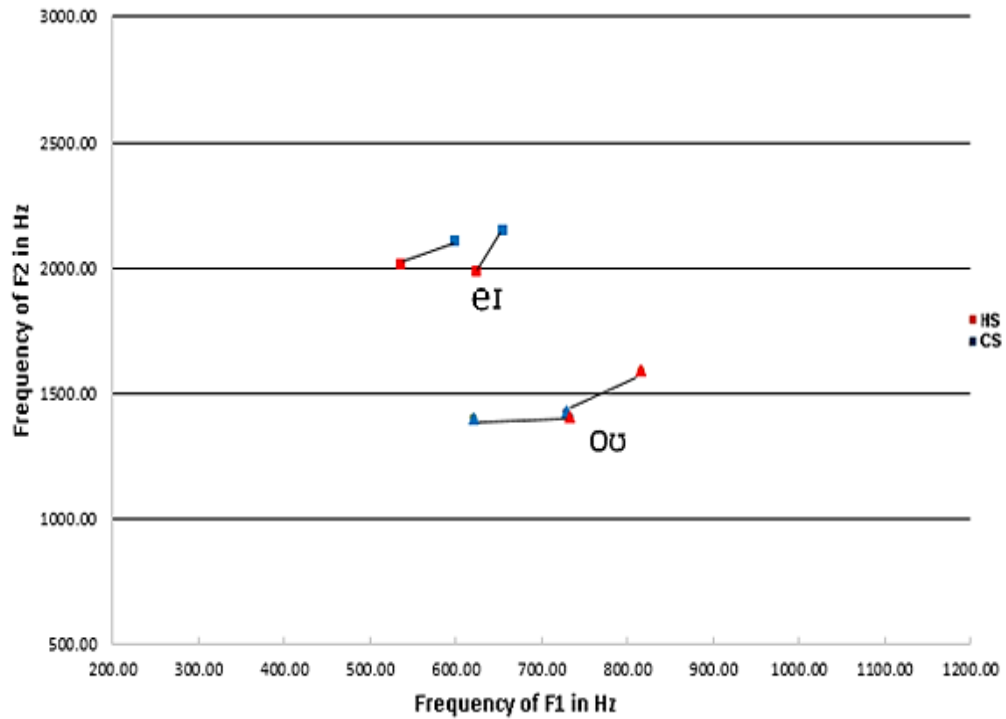


Figure 3.8b. F1/F2 plots of diphthongs produced by Speaker 4. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

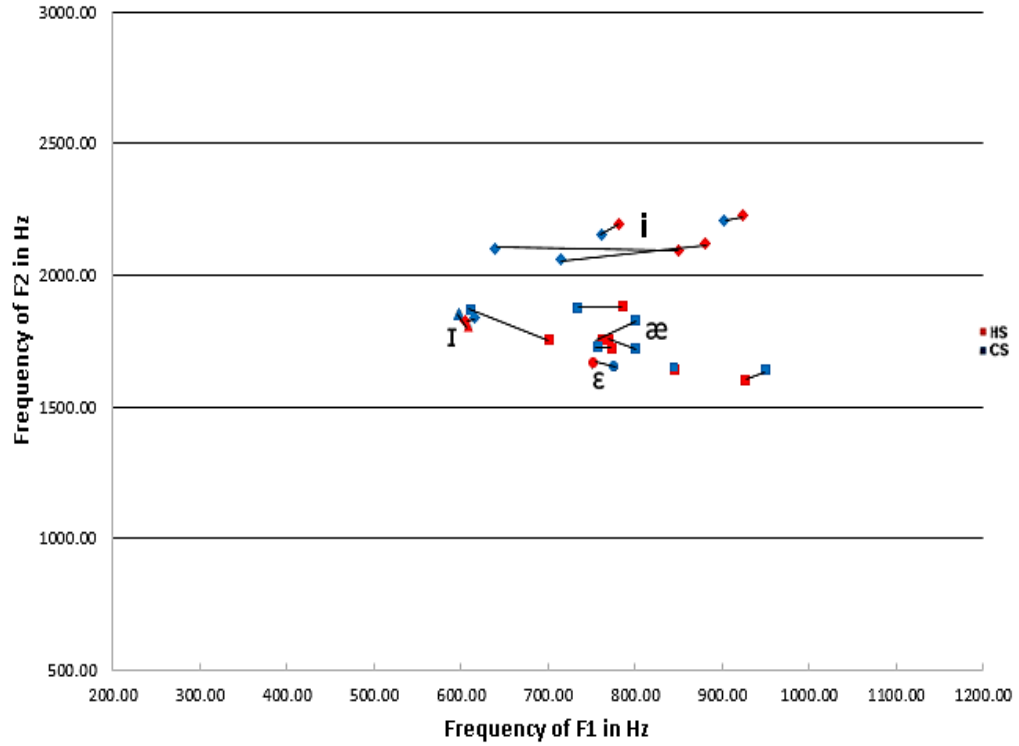


Figure 3.9a. F1/F2 plots of monophthongs produced by Speaker 5. *F1* = first formant; *F2* = second formant; *Hz*; hertz; *HS* = habitual speech; *CS* = clear speech.

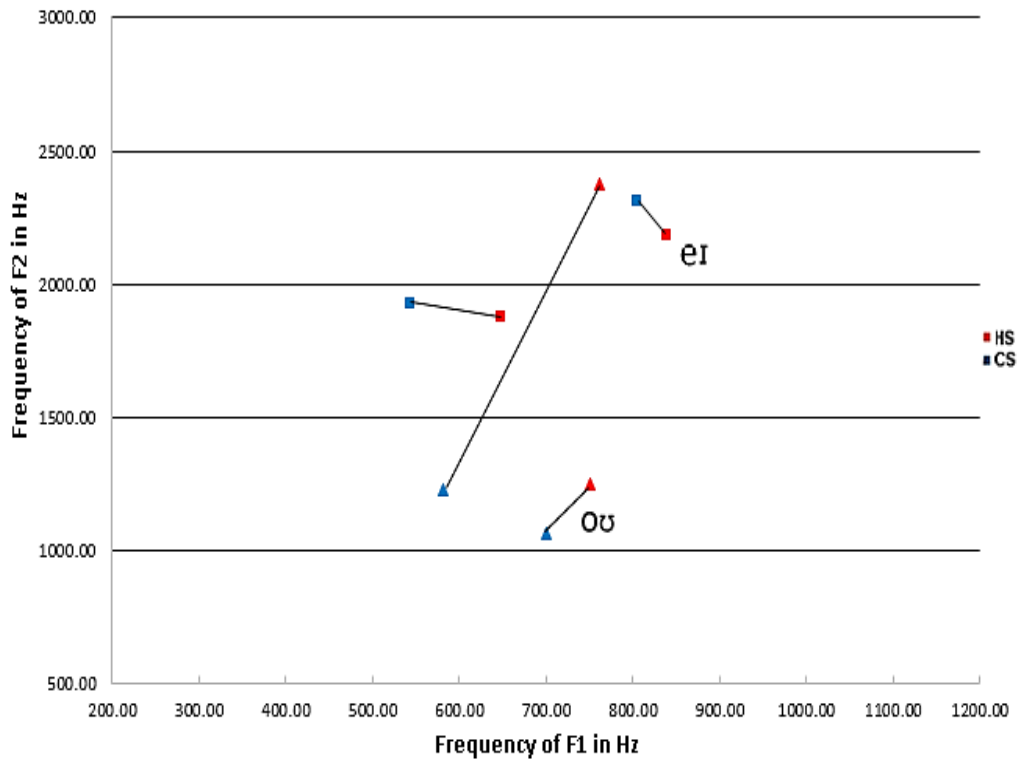


Figure 3.9b. F1/F2 plots of diphthongs produced by Speaker 5. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

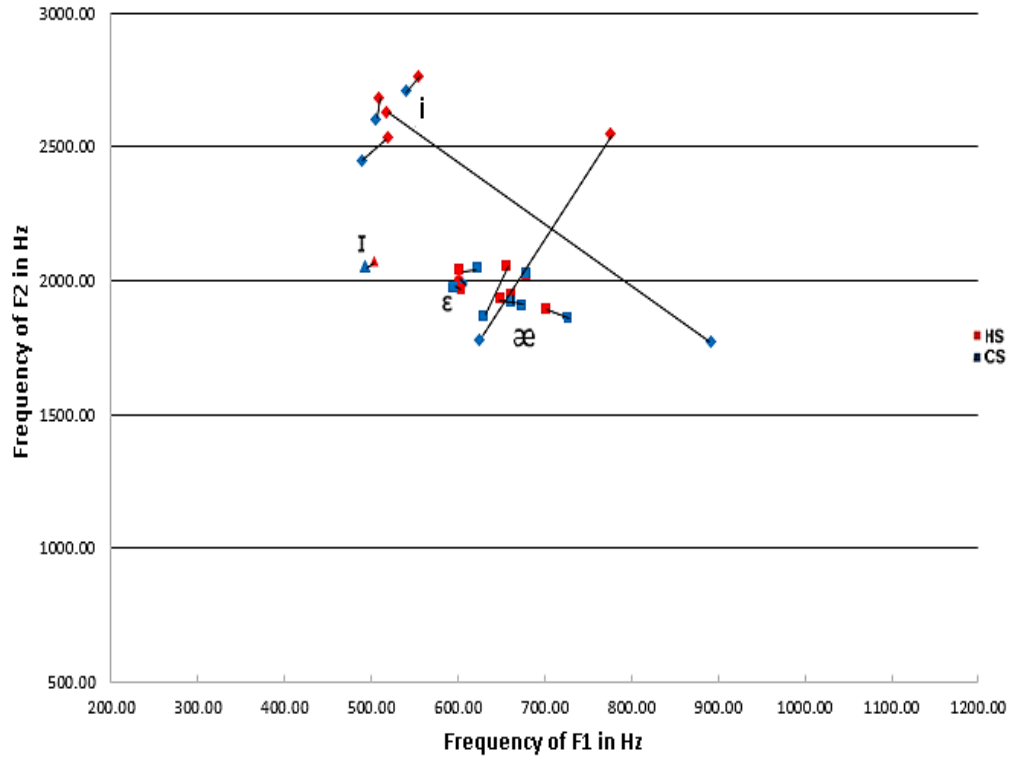


Figure 3.10a. F1/F2 plots of monophthongs produced by Speaker 6. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

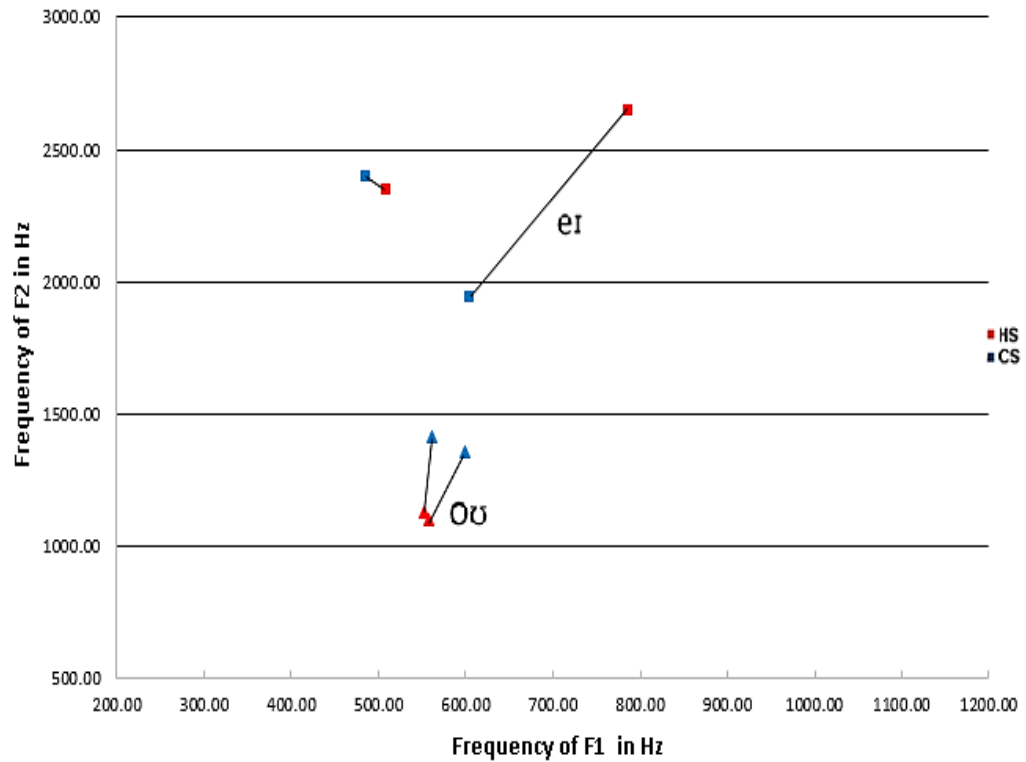


Figure 3.10b. F1/F2 plots of diphthongs produced by Speaker 6. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

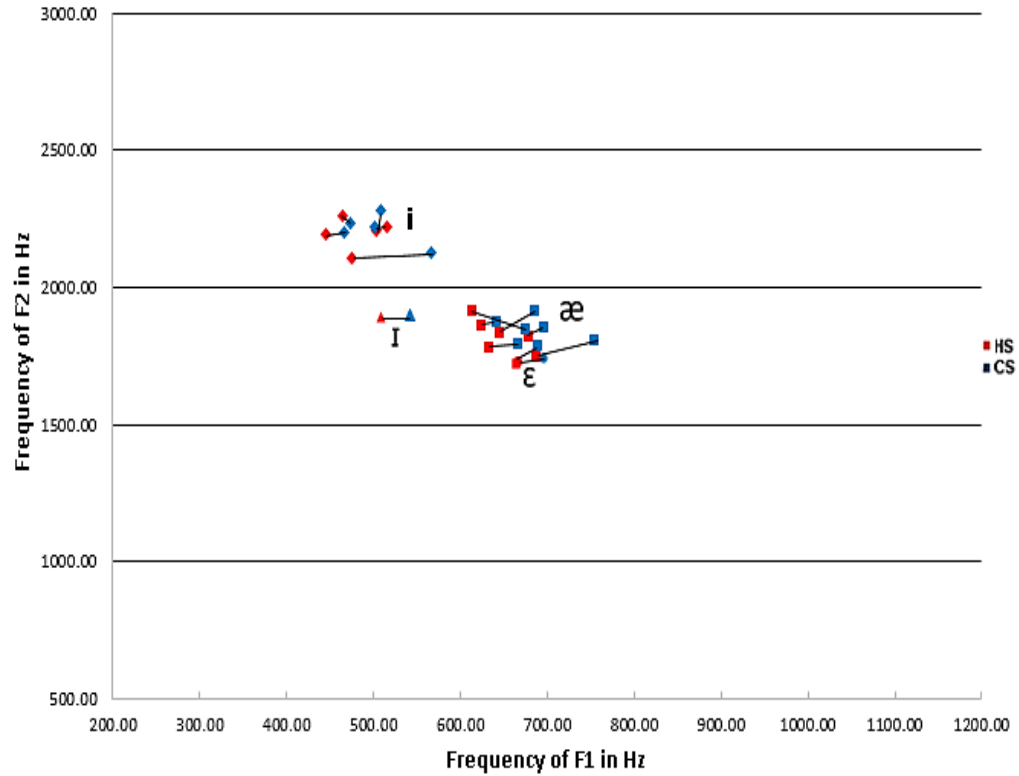


Figure 3.11a. F1/F2 plots of monophthongs produced by Speaker 7. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

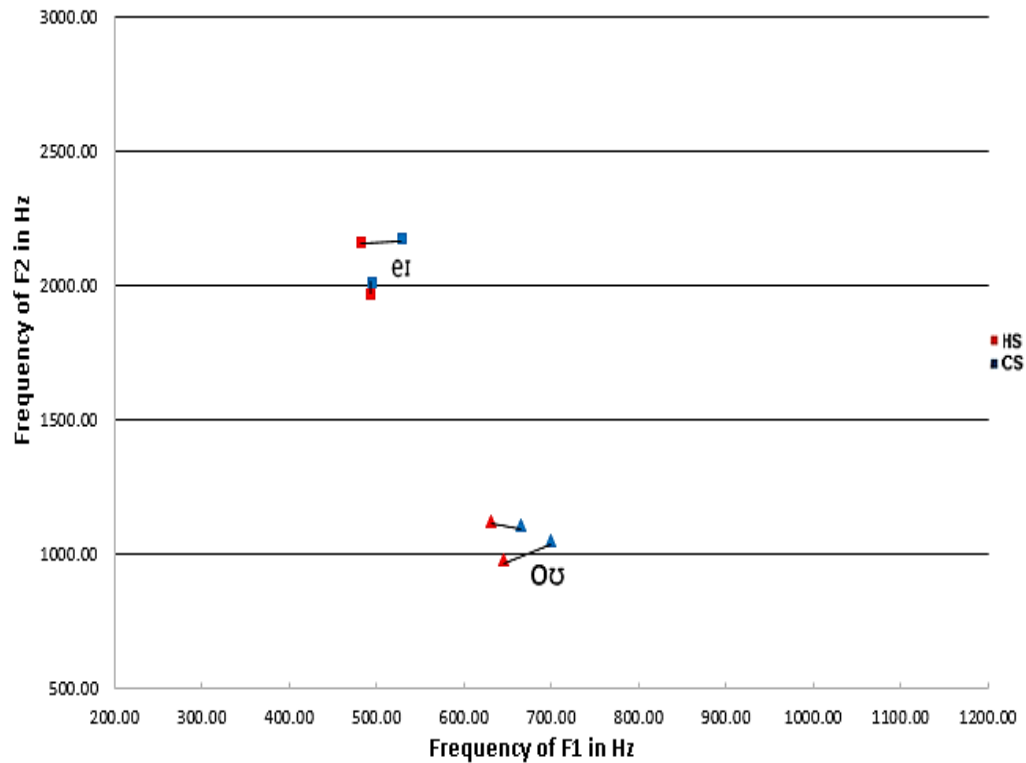


Figure 3.11b. F1/F2 plots of diphthongs produced by Speaker 7. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

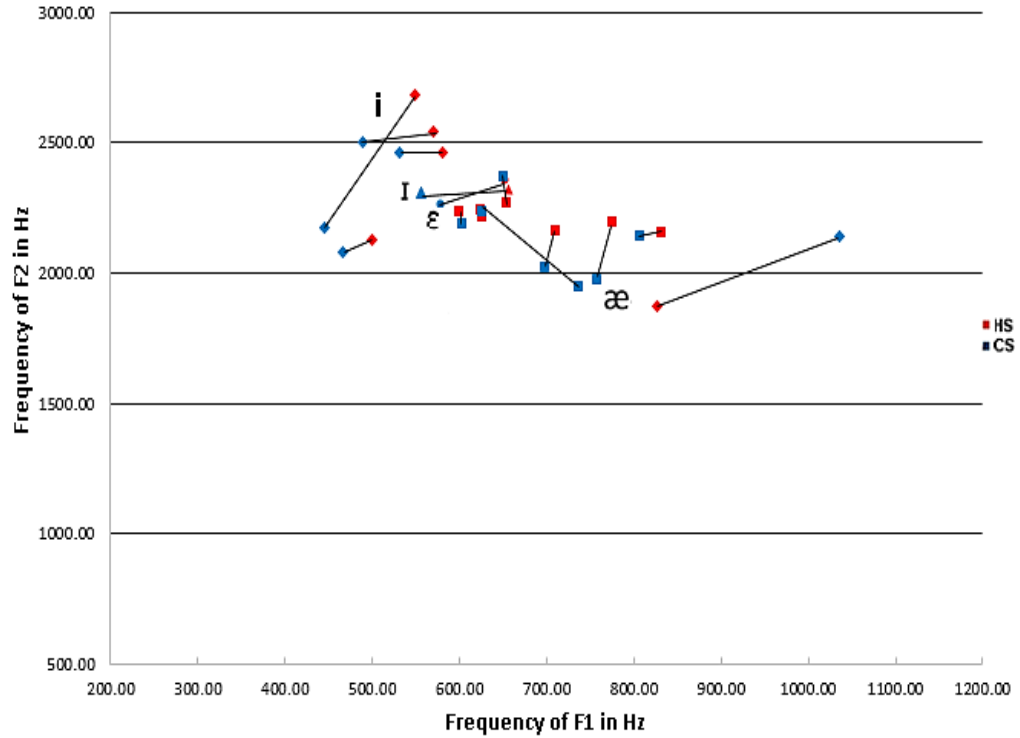


Figure 3.12a. F1/F2 plots of monophthongs produced by Speaker 8. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

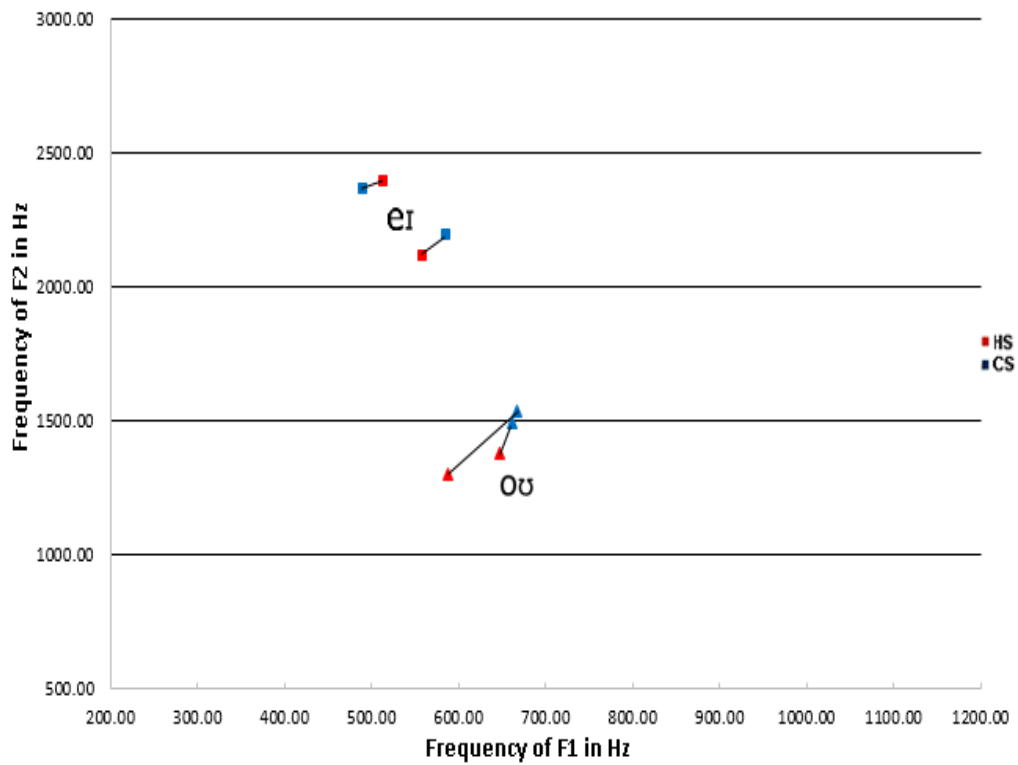


Figure 3.12b. F1/F2 plots of diphthongs produced by Speaker 8. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

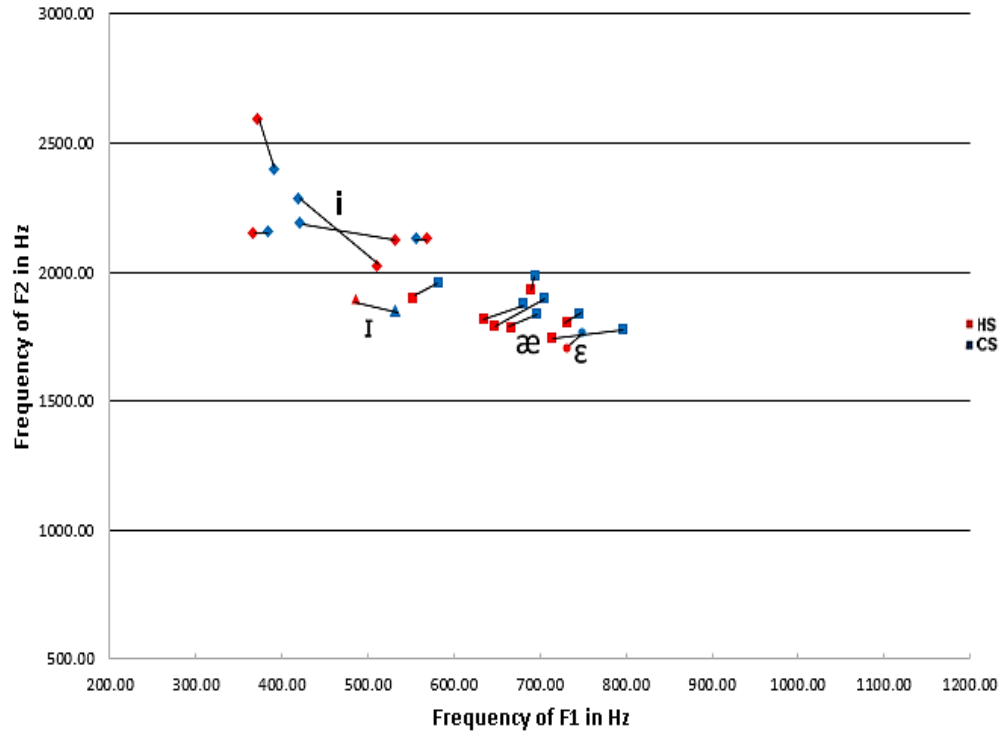


Figure 3.13a. F1/F2 plots of monophthongs produced by Speaker 9. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

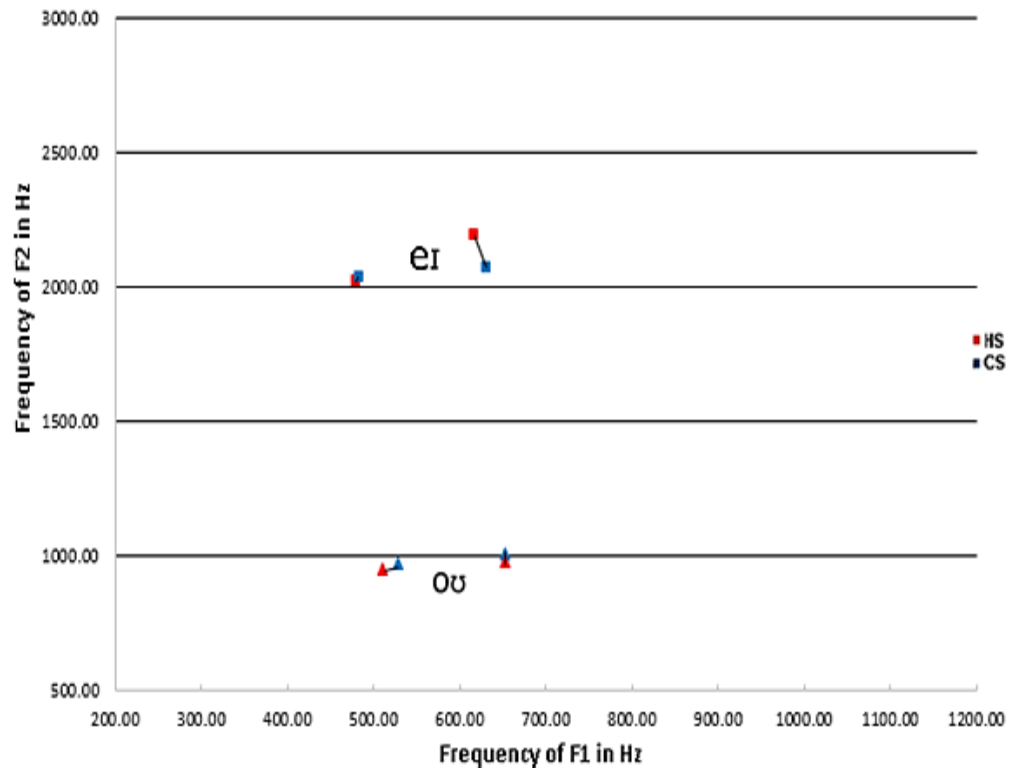


Figure 3.13b. F1/F2 plots of diphthongs produced by Speaker 9. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

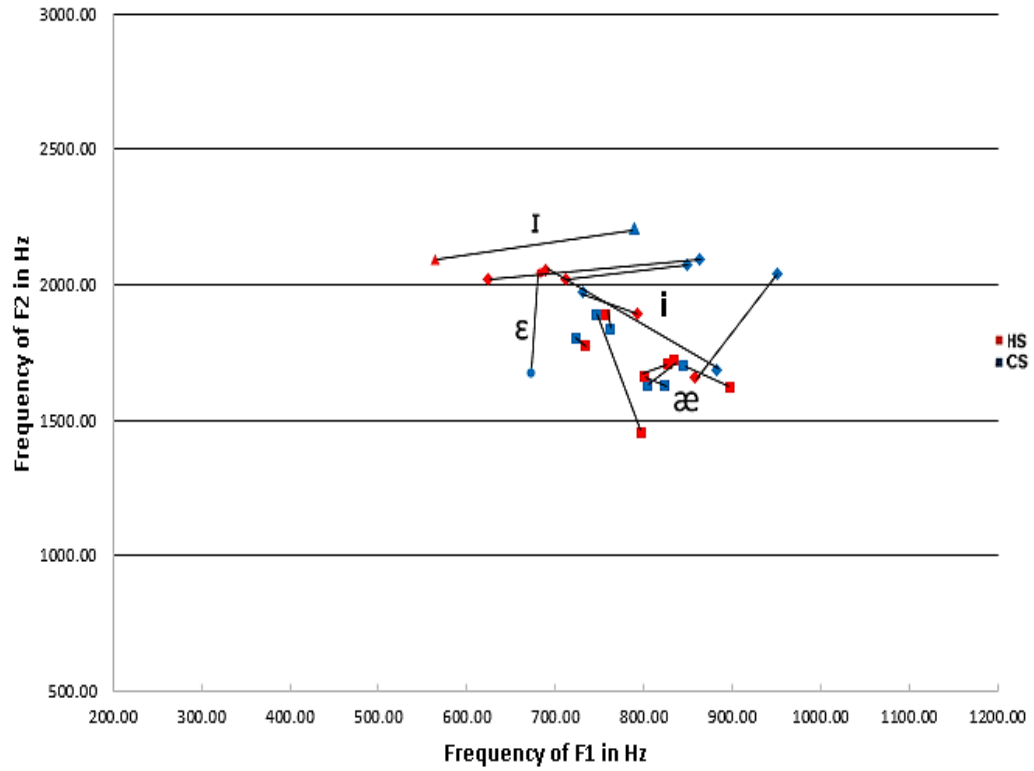


Figure 3.14a. F1/F2 plots of monophthongs produced by Speaker 10. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

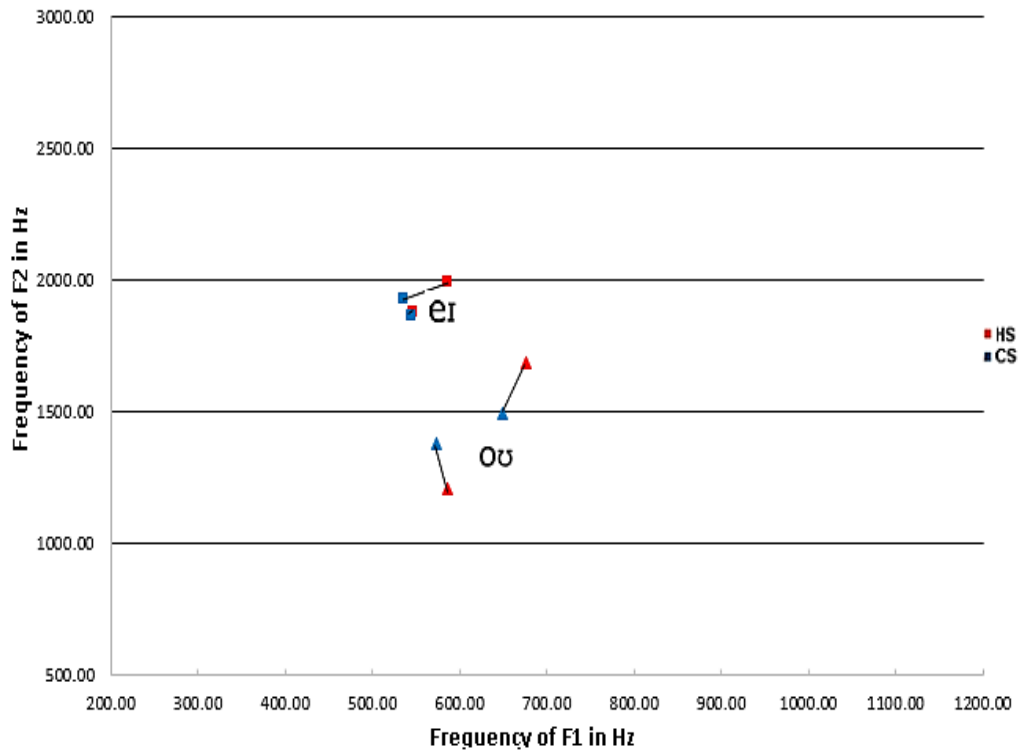


Figure 3.14b. F1/F2 plots of diphthongs produced by Speaker 10. *F1* = first formant; *F2* = second formant; *Hz* = hertz; *HS* = habitual speech; *CS* = clear speech.

Discussion

The purpose of this study was to determine the influence of CS on the acoustic characteristics of EL. Specifically, frequency and temporal data were obtained for comparison. CS was originally used to assist individuals with hearing impairment (Picheny et al., 1985; 1986) and more recently, it has been used to facilitate improved communication for individuals with various speech impairments (Beukelman et al., 2002; Tjaden et al., 2014). Results from previous work indicate that CS improves the understandability of speech for individuals with hearing impairment and for individuals listening to those individuals with speech impairments (Beukelman et al., 2002; Ferguson & Kewley-Port, 2002, 2007; Picheny et al., 1986; Tjaden et al., 2014). Therefore, the potential utility of CS was pursued in the present study because EL speech has been shown to demonstrate reduced intelligibility (Barney et al., 1959; Weiss & Basili, 1985; Weiss et al., 1979). Moreover, CS as a therapeutic strategy appears to be a natural fit for EL speakers since speech rehabilitation for this population typically involves a slowed speech rate and over-articulation of speech sounds (Doyle, 1994; 2005).

Research has indicated that phoneme durations increase when speaking rates decrease for normal speakers (Kessinger & Blumstein, 1997; Miller et al., 1986; Theodore et al., 2009). This also has been observed in previous research on CS; in addition, when speech rate was voluntarily reduced in combination with over-articulation of speech sounds, there was an improvement in SI (Lam & Tjaden, 2013; Picheny et al., 1986; Tjaden et al., 2014). Picheny et al. (1986) provided evidence to suggest that it was the lengthening of speech sounds in CS that played a role in such improvements, but follow-up research indicated that a reduction in speech rate was not the only reason why SI improved in CS (Picheny et al., 1989). In the present study, the duration of

monophthongs and diphthongs, in addition to overall word durations were analyzed and compared across EL speakers' productions in HS and CS.

Overall, the present study found that CS resulted in several varied acoustic changes in vowels and words in EL speech. First, vowel durations followed a pattern according to vowel placement within the oral cavity (e.g., high/low, anterior/posterior). For example, EL users' vowel durations were longest for the high /i/ and /æ/ vowels and gradually reduced in duration for mid vowels (e.g., /ε/ and /i/). Previous research also has indicated that vowel durations are nearly twice the duration when spoken in CS compared to vowels spoken in conversational speech (Ferguson & Kewley-Port, 2002; Picheny et al., 1986). In the present study, however, vowel productions only differed by 20 to 30 ms. When examining the data according to EL device grouping, Servox speakers' monophthongs and diphthongs were longer in CS compared to HS by ~22 and ~24 ms. TruTone users did not appear to reduce the durations of monophthongs as much using CS (i.e., ~2 ms). However, TruTone users' diphthong durations were increased in CS compared to HS by ~78 ms. These durational findings are in stark contrast from previous research findings indicating that vowel durations are twice as long in CS compared to conversational speech (Ferguson & Kewley-Port, 2002). It is important to note that this prior work involved a normal speaker who practiced as an audiologist, a profession that involves speaking to individuals with hearing impairment. Furthermore, instructing individuals to speak as though they are talking to individuals with hearing impairment is a hallmark feature of CS (Picheny et al., 1985). Therefore, the audiologist may have been more proficient in producing CS than the EL speakers used in the current study. This is evidenced by comparisons of vowel duration data from Ferguson and Kewley-Port

(2002)'s study and the present study. For example, the normal speaker produced drastically different mean durations of /i/ in conversational and CS (e.g., ~146 ms and ~417 ms, respectively). In the present study, overall mean durations for /i/ produced across all EL speakers was 310 ms in HS and 318 in CS. Ultimately, the data from the present study suggest that EL speakers were not as proficient when producing CS, especially when compared to HS.

The current data indicate that word durations for CS were generally longer than those produced during HS. Servox Digital users increased their mean word duration in CS by ~65 ms compared to productions in HS. This was slightly greater than TruTone users' productions, which were 51 ms longer in duration when using CS compared to HS. Compared to previous work, Picheny et al. (1986) found that stimuli (e.g., sentences) spoken in CS were twice the duration when compared to conversational (or, habitual) speech. In addition, Ferguson and Kewley-Port (2002) found that vowel durations doubled in CS compared to conversational (or, 'habitual') speech. Overall, the present word and vowel durations were not doubled when moving from HS to CS. Interestingly, CS is not known to produce a uniform change in rate of speech (Picheny et al., 1986). In fact, the EL speakers using CS in the present study varied greatly in their rate of speech during the production of words. For example, Servox users' productions of the 17 CVC words ranged from 513 to 880 ms in CS and 423 to 811 ms in HS. A similar range existed for TruTone users; that is, 483 to 776 ms for words in CS and 440 to 702 ms in HS. While closer analysis of the word-level duration data suggests some potential benefits of CS in reducing speech durations, several important limitations emerge.

First, our EL speakers were only provided with instructions to make their speech clearer, to reduce their rate and over-articulate without their overall speech rates being directly controlled or manipulated in any other way. This gave the EL speakers the freedom to control or modify their rate based solely on the instructions provided. Acoustic data from the present study suggest that some, if not most EL speakers, produced speech similarly in HS and CS, and therefore, each speaker could have benefitted from further CS instruction. Second, speakers were required to produce CS after instructions were provided to them during the experimental recording session. This meant that speakers had a limited window to think about the instructions being provided to them with no practice sessions prior to recording. Although providing additional time for the speakers to more actively consider the instructions may have been of benefit, the simplicity of the CS task may decrease the possible influence of such a consideration. However, a majority of these individuals also would have received initial training on the use of an EL device which typically involves a slower rate of speech and stresses the importance of over-articulation during use. All speakers were at least 24 months (up to 300 months) postlaryngectomy, so it is difficult to know if speakers maintained this slower rate and over-articulation from initial instruction. This could in part explain the lack of significant differences between HS and CS across the majority of stimulus conditions.

Other comparisons from the present study focused on fundamental frequency and formant frequencies. First, the F_0 for Servox Digital speakers was found to be approximately 74.8 Hz in HS and 74.9 Hz in CS for Servox users and 81.7 Hz in HS and 83.1 Hz in CS for TruTone users. These values are pre-set frequencies that are emitted

from each EL device, which unlike the dynamic nature of the vocal folds, will consistently vibrate at a relatively predictable pitch for each speaker during operation of each device.

For a typical laryngeal speaker, F1 and F2 frequency values are approximately 500 and 1500 Hz for /ə/, respectively. Sisty and Weinberg (1972) demonstrated that a reduced vocal tract length following TL will increase F1 and F2 frequency characteristics in those who had been laryngectomized and used esophageal speech (Sisty & Weinberg, 1972). General formant changes from the present study indicate that all EL speakers' demonstrated increased formant frequencies that are in line with those reported by Sisty and Weinberg (1972). Thus, the current data highlight a similar frequency response subsequent to a reduced vocal tract length for laryngectomized individuals who use the EL; as a result, it would appear that consistent patterns of change occur regardless of one's primary alaryngeal speaking method. Furthermore, even though the EL speakers in this study exhibited different absolute formant frequency values as a function of unique source and filter characteristics, many of the rules governing vowel formant patterning reported for normal, laryngeal speakers were maintained.

Vowel height followed proportional formant changes described by Peterson and Barney (1952); that is, low F1 and high F2 values shifted to higher F1 and lower F2 frequencies as Servox Digital and TruTone users moved from the high front vowel /i/ to the low front vowel /æ/. This pattern shifted slightly for each unique monophthong and diphthong, but as observed on the F1/F2 plots (see Figures 3.3a, 3.3b, 3.4a, and 3.4b), there is considerable overlap. The current formant data run contrary to some of the earlier findings of Peterson and Barney (1952) who observed more dispersion of the formant

data in their F1/F2 plots. The vowel spaces for EL speakers in the current study demonstrate a greater degree of overlap. It was expected that the production of CS would cause individuals to create a wider mouth opening and, therefore, influence formants by serving to increase the vowel spaces when compared to vowels produced during HS. However, formants for each vowel were relatively similar between device groups and speaking conditions with no predictable changes observed. Unfortunately, given that previous research has indicated a relationship between larger vowel spaces and higher SI in normal individuals and individuals with neurological impairment (Bradlow et al., 1996; Turner et al., 1995), it appears that the overlapped vowel spaces of EL speakers HS and CS productions speak to the general reductions in this alaryngeal communication method.

Unlike previous studies of vowels (Peterson & Barney, 1952), however, the present study manipulated the productive aspects of EL speech through CS. The production of CS was expected to increase F1 formant frequencies due to the requirement for increased mouth opening. Overall analyses of vowel data and F1/F2 plots indicate that CS actually resulted in minimal changes in the frequency of F1 formants when compared to HS values. This brings into question whether EL speakers fully demonstrated the limits of the CS production strategy; that is, did all EL speakers actively increase mouth opening while attempting to slow down rate in an effort to make their speech as clear as possible? Alternatively, were EL speakers already speaking with proper over-articulation, and thereby, could not over-articulate any further? Alongside comparisons to the proficient CS talker used by Ferguson and Kewley-Port (2002), the data suggest that EL

speakers' productions during CS resulted in insignificant frequency findings when compared to HS.

Conclusions

This study investigated the impact of CS on the acoustic characteristics of EL speech. Given the electronic and continuously voiced signal that characterized EL speech, minimal differences were observed between HS and CS conditions. Although minimal differences were observed for either temporal durations or frequencies of vowels, EL speakers' word durations appeared to increase to the greatest extent. While the focus of the present study was on overall word and vowel characteristics, further work will consider how each of the component parts of word stimuli (e.g., stop closure, release, VOT, vowel onset, vowel duration) contributes to longer word durations in CS compared to HS. In addition, it is important for future work to study the potential voicing effects of neighboring stimuli on vowels in EL speech, especially when speakers are instructed to reduce their speech rate. Finally, it is important to establish a criterion that differentiates CS from one's normal, conversational speech and other reduced rate conditions (e.g., slow) in order to properly assess speakers' proficiency in producing CS. Although the external validity of the current data are unknown due to this work being the first of its kind with alaryngeal speakers, future investigations of the potential utility of CS in alaryngeal speakers would appear to be warranted.

References

- Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America*, *113*, 544-552.
- Barney, H. L., Haworth, F. E., & Dunn, H. K. (1959). An experimental transistorized artificial larynx. *Bell system technical Journal*, *38*(6), 1337-1356.
- Bennett, S., & Weinberg, B. (1973). Acceptability ratings of normal, esophageal, and artificial larynx speech. *Journal of Speech and Hearing Research*, *16*, 608-615.
- Beukelman, D. R., Fager, S., Ullman, C., Hanson, E., & Logemann, J. (2002). The impact of speech supplementation and clear speech on the intelligibility and speaking rate of people with traumatic brain injury. *Journal of Medical Speech-Language Pathology*, *10*, 237-242.
- Boersma, P., & Weenink, D. (2015). Praat 5.4.17: doing phonetics by computer [software]. Amsterdam, Netherlands: University of Amsterdam. Retrieved from <http://www.praat.org/>.
- Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech communication*, *20*(3), 255-272.
- Chen, F. R. (1980). *Acoustic characteristics and intelligibility of clear and conversational speech at the segmental level*. Doctoral dissertation, Massachusetts Institute of Technology: Massachusetts.
- Christensen, J. M., & Weinberg, B. (1976). Vowel duration characteristics of esophageal speech. *Journal of Speech, Language, and Hearing Research*, *19*(4), 678-689.

Christensen, J. M., Weinberg, B., & Alfonso, P. J. (1978). Productive voice onset time characteristics of esophageal speech. *Journal of Speech, Language, and Hearing Research, 21*(1), 56-62.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Crystal, T. H., & House, A. S. (1990). Articulation rate and the duration of syllables and stress groups in connected speech. *The Journal of the Acoustical Society of America, 88*(1), 101-112.

Diedrich, W. M., & Youngstrom, K. A. (1966). *Alaryngeal speech*. Charles C. Thomas Publisher.

Doyle, P. C., Danhauer, J. L., & Reed, C. G. (1988). Listeners' perceptions of consonants produced by esophageal and tracheoesophageal talkers. *Journal of Speech and Hearing Disorders, 53*(4), 400-407

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

Doyle, P.C., & Eadie, T.L. (2005). The perceptual nature of alaryngeal voice and speech. In P.C. Doyle, & R.L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing*. Austin, TX: Pro-Ed.

Ferguson, S. H., & Kewley-Port, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America, 112*(1), 259-271.

- Ferguson, S. H., & Kewley-Port, D. (2007). Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *Journal of Speech, Language, and Hearing Research, 50*(5), 1241-1255.
- Gandour, J., Weinberg, B., & Rutkowski, D. (1980). Influence of postvocalic consonants on vowel duration in esophageal speech. *Language and speech, 23*(2), 149-158.
- Gay, T. (1978). Effect of speaking rate on vowel formant movements. *The Journal of the Acoustical society of America, 63*(1), 223-230.
- Goldman-Eisler, F. (1954). On the variability of the speed of talking and on its relation to the length of utterances in conversations. *British Journal of Psychology. General Section, 45*(2), 94-107.
- Goldman-Eisler, F. (1956). The determinants of the rate of speech output and their mutual relations. *Journal of Psychosomatic Research, 1*(2), 137-143.
- Hillman, R. E., Walsh, M. J., Wolf, G. T., Fisher, S. G., & Hong, W. K. (1998). Functional outcomes following treatment for advanced laryngeal cancer. Part I - Voice preservation in advanced laryngeal cancer. Part II - Laryngectomy rehabilitation: The state of the art in the VA system. Research Speech-Language Pathologists. Department of Veterans Affairs Laryngeal Cancer Study Group. *Annals of Otology, Rhinology and Laryngology Supplement, 172*, 1-27.
- Hoops, H.R., & Noll, J.D. (1969). Relationship of selected acoustic variables to judgements of esophageal speech. *J. Comm. Dis. 2*, 3-13.
- House, A. S. (1961). On vowel duration in English. *The Journal of the Acoustical Society of America, 33*(9), 1174-1178.

- House, A. S., & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *The Journal of the Acoustical Society of America*, 25(1), 105-113.
- Jongmans, P., Hilgers, F. J. M., Pols, L. C. W., & van As-Brooks, C. J. (2006). The intelligibility of tracheoesophageal speech, with an emphasis on the voiced-voiceless distinction. *Logopedics Phoniatrics Vocology*, 31(4), 172-181.
- Kalb, M.B., & Carpenter, M.A. (1981). Individual speaker influence on relative intelligibility of esophageal speech and artificial larynx speech. *Journal of Speech and Hearing Disorders*, 46, 77-80.
- Kessinger R. H., & Blumstein S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English,” *J Phonetics*, 25, 143–168.
- Lam, J., & Tjaden, K. (2013). Intelligibility of clear speech: Effect of instruction. *Journal of Speech, Language, and Hearing Research*, 56(5), 1429-1440.
- Lam, J., Tjaden, K., & Wilding, G. (2012). Acoustics of clear speech: Effect of instruction. *Journal of Speech, Language, and Hearing Research*, 55(6), 1807-1821.
- Lindblom, B. E. (1963). Spectrographic study of vowel reduction. *The journal of the Acoustical Society of America*, 35(11), 1773-1781.
- Lindblom, B. E. (1964). Articulatory activity in vowels. *The Journal of the Acoustical Society of America*, 36(5), 1038-1038.
- Mazzoni, D., & Dannenberg, R. (2013). Audacity 2.0.5 [software]. Pittsburgh, PA: Carnegie Mellon University. Retrieved from: <http://audacityteam.org/>

- Meltzner, G. S., & Hillman, R. E. (2005). Impact of aberrant acoustic properties on the perception of sound quality in electrolarynx speech. *Journal of Speech, Language, and Hearing Research, 48*(4), 766-769.
- Meltzner, G. S., Kobler, J. B., & Hillman, R. E. (2003). Measuring the neck frequency response function of laryngectomy patients: Implications for the design of electrolarynx devices. *The Journal of the Acoustical Society of America, 114*(2), 1035-1047.
- Miller, J. L., Green, K. P., & Reeves, A. (1986). Speaking rate and segments: A look at the relation between speech production and speech perception for the voicing contrast. *Phonetica, 43*(1-3), 106-115.
- Miller, J. L., & Volaitis, L. E. (1989). Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics, 46*(6), 505-512.
- Moon, S. J., & Lindblom, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical society of America, 96*(1), 40-55.
- Öhman, S. E. (1967). Numerical model of coarticulation. *The Journal of the Acoustical Society of America, 41*(2), 310-320.
- Peterson, G. E. & Barney, H. L. (1952) Control methods used in the study of vowels, *Journal of the Acoustical Society of America, 24*(2), 175-184.
- Peterson, G. E., & Lehiste, I. (1960). Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America, 32*(6), 693-703.

- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing. I. Intelligibility differences between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 28, 96–103.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, 29, 434–446.
- Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *The Journal of the Acoustical Society of America*, 51(4B), 1296-1303.
- Raphael, L. J., Dorman, M. F., Freeman, F., & Tobin, C. (1975). Vowel and nasal duration as cues to voicing in word-final stop consonants: Spectrographic and perceptual studies. *Journal of Speech, Language, and Hearing Research*, 18(3), 389-400.
- Robbins, J., Christensen, J., & Kempster, G. (1986). Characteristics of speech production after tracheoesophageal puncture: Voice onset time and vowel duration. *Journal of Speech, Language, and Hearing Research*, 29(4), 499-504.
- Robbins, J., Fisher, H.B., Blom, E.C., & Singer, M.I. (1984). A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production. *Journal of Speech and Hearing Disorders*, 49(2), 202–210.
- Snidecor, J. C., & Curry, E.T. (1959). Temporal and pitch aspects of superior esophageal speech. *Ann Otol Rhinol Laryngol*, 68, 623-629.
- Sisty, N. L., & Weinberg, B. (1972). Formant frequency characteristics of esophageal speech. *Journal of Speech, Language, and Hearing Research*, 15(2), 439-448.

- Stevens, K. N., & House, A. S. (1955). Development of a quantitative description of vowel articulation. *The Journal of the Acoustical Society of America*, 27(3), 484-493.
- Theodore, R. M., Miller, J. L., & DeSteno, D. (2009). Individual talker differences in voice-onset-time: Contextual influences. *The Journal of the Acoustical Society of America*, 125(6), 3974-3982.
- Tjaden, K., Sussman, J.E., & Wilding, G. E. (2014). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research*, 57, 779-792.
- Trudeau, M. D., & Qi, Y. Y. (1990). Acoustic characteristics of female tracheoesophageal speech. *Journal of Speech and Hearing Disorders*, 55, 244-250.
- Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, 38(5), 1001-1013.
- Umeda, N. (1975). Vowel duration in American English. *The Journal of the Acoustical Society of America*, 58(2), 434-445.
- Umeda, N. (1977). Consonant duration in American English. *The Journal of the Acoustical Society of America*, 61(3), 846-858.
- Weiss, M S., & Basili, A. M. (1985). Electrolaryngeal speech produced by laryngectomized subjects: Perceptual characteristics. *Journal of Speech and Hearing Research*, 28, 294-300.

- Weiss, M. S., Yeni-Komshian, G. H., & Heinz, J. M. (1979). Acoustical and perceptual characteristics of speech produced with an electronic artificial larynx. *The Journal of the Acoustical Society of America*, 65(5), 1298-1308.
- Van den Berg, J. W., & Moolenaar-Bijl, A. J. (1959). Crico-pharyngeal sphincter, pitch, intensity and fluency in oesophageal speech. *ORL*, 21(4), 298-315.
- Yeni-Komshian, G. H., Weiss, M. S., & Heinz, J. M. (1975). Intelligibility of speech produced with an electronic artificial larynx. *The Journal of the Acoustical Society of America*, 58(S1), S112-S112.

Chapter 4

The Impact of Clear Speech on Auditory-Perceptual Judgments of Electrolaryngeal Speech

Contemporary voice and speech rehabilitation following total laryngectomy remains characterized by three postsurgical options, namely, esophageal speech, tracheoesophageal voice restoration, and use of the electronic artificial larynx or the electrolarynx (EL). Despite criticism regarding its use, information suggests that the EL remains a widely used postlaryngectomy method of communication (Hillman, Walsh, Wolf, Fisher, & Wong, 1998; Meltzner & Hillman, 2005; Mendenhall et al., 2002; Ward, Koh, Frisby, & Hodge, 2003). While the EL provides a readily accessible form of communication following laryngectomy for most individuals, EL speech demonstrates obvious acoustic and perceptual deviations compared to normal speech. For example, EL speech has been described as unnatural due to its mechanical quality (Bennett & Weinberg, 1973; Doyle & Eadie, 2005; Hillman et al., 1998; Meltzner & Hillman, 2005). In addition, speech produced using an EL contains numerous acoustic deficits in both intensity and frequency (Qi & Weinberg, 1991; Saikachi, Stevens, & Hillman, 2009; Verdolini, Skinner, Patton, & Walker, 1985). The resulting speech signal, even if highly intelligible to the listener, is one that is frequently judged as being monotone, characterized by a robotic voice quality that in itself may make communication challenging (Bennett & Weinberg, 1973; Cole, Sridharan, Moody, & Geva, 1997; Doyle, 1994; Hillman et al., 1998; Meltzner & Hillman, 2005). The goal of all communication, including postlaryngectomy speech produced using an EL, frequently centers on how well the speaker is understood (Goldstein, 1978). Characteristics of the EL sound source, however, create additional perceptual challenges. Although numerous approaches have

been used to improve the acoustic and perceptual aspects of EL speech (e.g., Espy-Wilson, Chari, Huang, & Walsh, 1998; Meltzner & Hillman, 2005), the electronic quality is clearly abnormal, which may place additional burden on the listener. Consequently, the present study is the first to examine the influence of clear speech (CS) on auditory-perceptual judgments of EL speech.

Briefly, CS is a style of speaking that involves a reduced rate of speech and over-articulation when compared to normal (or habitual) speech (Payton, Uchanski, & Braida, 1994; Picheny, Durlach, & Braida, 1985). CS has been used to facilitate improved speech intelligibility (SI) from 7 to 11% in individuals with a variety of communication disorders (Beukelman, Fager, Ullman, Hanson, & Logemann, 2002; Tjaden, Sussman, & Wilding, 2013). However, it is also important to consider the potential impact of CS on the listener's perception of EL speakers who use CS in an attempt to make their speech more intelligible. More specifically, because CS involves adjustments that are effected primarily at the temporal level speech production, such adjustments may alter the signal in a manner that also introduces new perceptual challenges to the listener.

Beyond how well a speaker's message is understood specific to SI (Kent, 1996), listeners often make judgments about what they hear and how the speaker's message is communicated. In addition, the "quality" of one's voice can influence listeners considerably (Kent, 1996; Kreiman, Gerratt, Kempster, Erman, & Berke, 1993). For example, listeners can decide how acceptable or pleasing a speaker's voice and speech are during communication, or how comfortable they are listening to the speaker. These types of auditory-perceptual judgments are often beneficial in determining the larger communication success of those who must rely on any method of postlaryngectomy

voice, including that of the EL. Therefore, a comparative evaluation of auditory-perceptual ratings of EL speakers when using habitual speech (HS) and CS may provide valuable information about the impact of such a speech production modification on listeners.

Speech acceptability (ACC) and listener comfort (LC) are two auditory-perceptual judgments that have been previously described in the communication disorders literature (Bennett & Weinberg, 1973; Doyle, 1999; Doyle & Eadie, 2005; Doyle et al., 2011; Eadie et al., 2007; Eadie et al., in press; O'Brian et al., 2003). Briefly, ACC refers to a listener's composite perceptual evaluation of pitch, rate, understandability, and voice quality (Bennett & Weinberg, 1973). In order to provide a judgment of ACC, listeners are asked to consider all four of these attributes without placing additional weight on a specific feature; it is in fact a collective perceptual assessment of the speech signal. In contrast, LC is a perceptual feature that assesses how comfortable listeners are when communicating with individuals who have a communication disorder (O'Brian et al., 2003). Even though the initial research on LC focused on individuals who stuttered, Eadie et al. (2007) have expanded the application of LC to include individuals with voice disorders. They concluded that auditory-perceptual ratings of LC might be useful for determining the impact of voice disorders on listeners, regardless of their experience in listening to disordered voice and speech (Eadie et al., 2007).

Collectively, ACC and LC would appear to be appropriate auditory-perceptual features to better understand the impact of CS on EL speakers. Since CS relies on a reduced rate of speech and over-articulation of speech sounds, auditory-perceptual

judgments of ACC would appear to be a natural fit to the application of CS in those who use the EL postlaryngectomy; that is, for both dimensions listeners are specifically asked to focus on the previously mentioned perceptual composite that includes speaking rate as one of its attributes. In addition, EL voice and speech presents unique pitch and quality characteristics, therefore, ACC and LC may permit a direct means of documenting the potentially disruptive effects of CS on an already ‘unnatural’ and mechanical voicing method. While research efforts must continue to focus on the effectiveness with which individuals are able to communicate using an EL, the impact of CS on listener perception remains unknown. Given the unique characteristics of the EL voice signal (i.e., low frequency, electronic), in addition to the speech alterations that occur as a direct result of CS (i.e., slowed speech rate, over-articulation), ACC and LC may serve as ideal auditory-perceptual features in an effort to assess the influence of CS on listeners. Thus, the purpose of this study sought to investigate the impact of CS on the auditory-perceptual judgments of normal-hearing listeners.

Method

Participant Speakers

Voice samples from ten adult males who served as participants for Experiment 1 and 2 were used in the present experiment. Each participant had undergone total laryngectomy and used an EL device as their primary mode of verbal communication served as participant speakers in this study. Speakers ranged in age from 59 to 87 years ($M_{\text{age}} = 74$ years) and were at least 24 months postlaryngectomy ($M_{\text{time}} = 133$ months; $\text{range} = 24\text{-}300$ months). Each speaker reported use an EL device since TL. Seven speakers had a neck dissection during their TL. All speakers had received radiation

therapy either prelaryngectomy (n= 4), postlaryngectomy (n= 5), or pre- *and* postlaryngectomy (n=1). Two speakers received concurrent chemoradiotherapy pre- (n=1) and postlaryngectomy (n=1). At the time of their participation in the study, all speakers indicated that they were in good general health with no known neurological, medical or psychological conditions, were native English speakers, and that they did not have hearing difficulties that prevented them from communicating with others in a quiet environment. Given the age and medical treatment related to TL, however, some level of hearing loss cannot be ruled out. Informed consent and demographic information was obtained from all speakers prior to their participation (Western University Research Ethics Board Approval #105382) (see Appendices B - D).

Preliminary Intelligibility Assessment

All speakers met a minimum consonant intelligibility criterion of at least 60% based on their production of an 18-item word list that was comprised of monosyllabic stimuli from Experiment 1. With one exception, stimulus items represented consonant-vowel-consonant constructions; the single exception was a consonant-vowel item. The intelligibility stimuli were derived from a longer word list that was originally presented by Weiss and Basili (1985) (see Appendix A). This subset of stimuli was selected so that each of the 17 consonants under assessment could be represented in both word initial and word final positions in the most efficient manner. The consonants represented were six plosives (/p/, /b/, /t/, /d/, /k/, and /g/), seven fricatives (/f/, /v/, /s/, /z/, /ʃ/, /θ/, and /ð/), two affricates (/tʃ/ and /dʒ/), and two nasals (/m/ and /n/). Digital recordings of stimuli using a sampling rate of 44 kHz were obtained during a single session that lasted approximately 20 minutes. Once the recording of word stimuli was completed, the

sentence stimuli used in the present experiment were obtained using the same recording equipment.

Upon completing the recording of word stimuli, individual items were digitally extracted and then randomized into multiple lists. These lists were then presented to 12 normal-hearing, naïve, young adult listeners who ranged in age from 19;10 to 33;08 years ($M_{\text{age}} = 24;05$ years). Listeners were instructed that they would be presented with a series of real English word and would then be requested to orthographically transcribe each word item that was heard. In cases where any consonant or vowel could not in any manner be identified, listeners were requested to draw a line through that item on the score sheet. Stimuli were presented to individual listeners under headphones in a quiet laboratory; a listener's entire series of transcriptions all obtained in a single session that ranged from 55 to 113 minutes ($M_{\text{time}} = 81$ minutes).

Once all listeners had completed the task, transcriptions were scored by an independent evaluator. Based the entire set of stimuli spoken by each speaker, an intelligibility score was calculated. This score was determined by identifying the correct number of correct listener judgments over the entire series of stimuli presented for each speaker. A summary of the individual speaker intelligibility scores is presented in Table 4.1. As can be seen, intelligibility scores ranged from 59.8 to 82.8% and from 62.3 to 97.1% for word initial and word final consonants, respectively.

Data Collection - Experimental Speech Samples

Recording of speech stimuli. Speech stimuli were obtained from all speakers using a unidirectional microphone (Shure PG-81, Niles, IL), a pre-amp (M-Audio, Avid Technology, Burlington, MA) and a laptop computer (Dell Inspiron, Round Rock, TX)

Table 4.1*Individual Speaker Intelligibility Scores for Word-Initial and Word-Final Stimuli*

Speaker	Word-Initial Consonants	Word-Final Consonants
1	78.4	92.6
2	80.9	75.0
3	82.8	97.1
4	82.8	83.8
5	59.8	62.3
6	78.4	78.9
7	63.2	62.3
8	73.0	85.8
9	80.9	87.7
10	75.0	86.3
Overall	75.5	81.2

Note. Speech intelligibility scores are shown as percentages.

that utilized Sona-Speech II software (Kay Pentax, Lincoln Park, NJ). All recordings were digitized at a sampling rate of 44.1 kHz.

To begin the recording session, participants were handed a printed copy of the Rainbow Passage (Fairbanks, 1960) and provided with the following instructions: “Please take a moment to look over the following paragraph. Once you are ready, please read it aloud. If you make a mistake, I will ask you to repeat the sentence(s) once you finish reading”. This will be referred to as the Habitual Speech (HS) condition. When speakers finished, they were provided with the following instructions: “Now I would like you to re-read the reading passage by speaking as clearly as possible. This will involve slowing down while speaking and over-articulating” (Picheny et al., 1985). This will be referred to as the CS condition. Because this study sought to comparatively assess auditory-perceptual dimensions between HS and CS speech conditions for EL speakers, the order of recording was not counterbalanced. We believed that had any of the CS samples been recorded first for any speaker, that this would increase the potential that some of those production behaviors may have been carried over to the HS condition. Thus, by recording all speakers using HS first, the likelihood of a CS confound was reduced if not fully eliminated. Recordings from each speaker were obtained in a quiet room free of background noise during a single session; the entire session lasted approximately 20 minutes.

Listener Stimuli

The audio files containing the first three sentences from each participant speakers’ Rainbow Passage served as auditory-perceptual stimuli for this study. All stimuli were initially edited on a desktop computer (Dell Optiplex, Round Rock, TX) using the

software program Audacity 2.0.5 (Mazzoni & Dannenberg, 2013) First, audible recording noise was removed from all 20 audio files using the ‘Noise Removal’ tool within Audacity. Specifically, a small window was highlighted at the beginning of each audio file (e.g., not involving speech stimuli) to capture a profile of track noise. The track noise was analyzed and then removed by the ‘Noise Removal’ tool, leaving speech stimuli unaltered in the process. Next, the first three sentences from each passage were extracted and used as stimuli for the present study.

Across all participant speakers, there were 20 experimental samples [1 speech sample x 10 speakers x 2 speaking conditions]. Additionally, 20% of the original samples (n=4) were randomly selected to assess reliability of judgments and these were included in the randomization of all speech stimuli presented to listeners. A single EL sample that was not produced by one of the 10 participant speakers also was included as an exposure sample to orient listeners to the types of samples they would be evaluating. Therefore, a total of 25 stimuli (1 exposure sample + [1 speech sample x 10 participant speakers x 2 speaking conditions] + 4 reliability samples) were generated for the auditory-perceptual phase of the study. Finally, all listener stimuli were randomized into 20 unique lists for participant listeners using a computer program written specifically for this project (Failla, 2014).

Participant Listeners

Twenty undergraduate and graduate students (eight males, 12 females) who had not participated in the intelligibility assessment phase of the project served as listeners in this study. Listeners ranged in age from 19;10 to 33;08 years ($M_{\text{age}} = 24;05$ years) and all were native English speakers. At the time of the study, participants indicated that they

were in good health, had no history of upper respiratory infections in the past week, and had no history of speech, voice, language, or hearing difficulties. Listeners participated voluntarily and were not reimbursed for their time or participation. Informed consent was obtained from all listeners prior to their participation (Western University Research Ethics Board Approval #105884) (see Appendices E and G).

All participant listeners were deemed to be naïve after confirming that they had no training in speech-language pathology and no formal experience listening to voice and/or speech disorders. Naïve listeners are representative of the population who laryngectomees are more likely to encounter on a daily basis (Eadie & Doyle, 2004; Kreiman et al., 1993; Tardy-Mitzell, Andrews, & Bowman, 1985). Further, research has demonstrated that naïve listeners are able to make reliable judgments pertaining to the differences between normal and alaryngeal speakers (Bennett & Weinberg, 1973; Eadie & Doyle, 2004). Therefore, the use of naïve speakers in the present experiment appeared worthwhile to obtain perceptual judgments of ACC and LC while EL speakers used HS and CS.

Auditory-Perceptual Rating Procedure

Participant listeners provided auditory-perceptual ratings over two sessions separated by approximately one week. Initial listening sessions were counterbalanced for the two perceptual dimensions under investigation, namely ACC and LC; during the first session half of the participant listeners made judgments of ACC while the other half of the participants were asked to make judgments of LC. In the second listening session, participants completed the same rating procedure for the remaining perceptual dimension

(e.g., if ACC was rated in the initial listening session, LC was rated in the second session and vice versa).

Each participant listener sat in front of a desktop computer (Dell, Optiplex, Round Rock, TX) and was provided with headphones and rating sheets. The listener was then instructed to click on a sound file corresponding to the one shown on the score sheets and then rate that sample. All samples were rated using a 100-mm visual analog scale (VAS) with the listener asked to bisect the scaled line at a point that best represented their evaluation of any given sample. The anchors for the ACC scale ranged from “Very Acceptable” to “Very Unacceptable”; for LC, the anchors ranged from “Very Comfortable” to “Very Uncomfortable”. Listeners were requested to make each rating independent of other samples. Further, listeners were permitted to listen to any sample as many times as they wished before making their rating, however, once entered on score sheet, they were instructed to not alter the rating or return to that sample again.

For ratings of ACC, participant listeners were provided with the following instructions at the beginning of each recording session:

In making your judgments about the speakers you are about to hear, give careful consideration to the attributes of pitch, rate, understandability, and voice quality. In other words, is the voice pleasing to listen to, or does it cause you some discomfort as a listener? (Bennett & Weinberg, 1973, p. 610).

Similarly, participant listeners were provided with the following instructions for LC:

How comfortable would you feel listening to the person’s speech in a social situation? Your rating should reflect your feelings about the way the person was speaking, not what the person was saying or how their personality affected you. (O’Brian et al., 2003, p. 509).

Once the initial rating session was completed, listeners were scheduled for the second session. Participant listeners typically returned one week later (*range* = 7-10 days) to

provide their ratings for the remaining perceptual dimension. This separation between listening sessions was done to control for any possible learning effects that might influence their judgments. The average time to complete the listening session for ACC ratings was 18 minutes and 35 seconds (*range* = 12-24 minutes) and 15 minutes and 9 seconds for LC ratings (*range* = 4 minutes 30 seconds-21 minutes).

Data Analysis

All VAS responses were scored using direct measurements (in mm) with final individual participant scores ranging from 1 (representing ‘Very Acceptable’ for ACC or ‘Very Comfortable’ for LC) to 100 (‘Very Unacceptable’ for ACC or ‘Very Uncomfortable’ for LC). Scaled scores were calculated using a ruler to determine the distance from the leftmost endpoint to the point of the listener’s response as indicated by a line crossing the scale. The resulting measurement was recorded for listener responses to all stimuli rated in both sessions.

A repeated measures analysis of variance (ANOVA) was used to examine the effect of speaking condition on ACC and LC listener scores. The magnitude of effect for each speaking condition was determined by calculating partial eta squared. Interpretation of effect size followed guidelines by Cohen (1988) (e.g., 0.01 = small effect, 0.06 = medium effect, and 0.14 = large effect). A Bonferroni correction was used for post-hoc testing, and an a priori significance level set at $p < 0.05$. Pearson Product-Moment correlations were used to describe the relationships amongst ACC and LC scores and the HS and CS speaking conditions.

Agreement and Reliability

Intra-rater reliability for ACC and LC ratings were calculated by comparing the first and second ratings of the four samples that were duplicated; this was achieved through the calculation of agreement. ACC and LC scores that fell within +/- 15 points of initial ratings were arbitrarily selected to demonstrate good levels of agreement. However, we also calculated agreement using +/- 5 and +/-10 scaled score criteria. This is more conservative than recent studies evaluating listeners judgments of EL speech, which used judgments within +/- 20 points (e.g., Nagle, Eadie, Wright, & Sumida, 2012). Finally, intra-class correlation coefficients (ICC; Shrout & Fleiss, 1979) were used to analyze inter-rater reliability.

Agreement data for listener ratings of ACC are shown in Table 4.2a. Intra-rater agreement by listeners for ACC ranged from 50% to 100% ($M = 72.5\%$). More specifically, 27/80 (33.75%) of listener judgments fell within +/-5 mm, 43/80 (53.75%) fell within +/-10 mm, and 58/80 (72.5%) fell within +/-15 mm of the initial sample ratings. The group mean average ICC for ACC was .941 with 95% confidence interval (CI) (0.896, .973).

Reliability data for listener ratings of LC are shown in Table 4.2b. For LC ratings, intra-rater agreements by listeners ranged from 50% to 100% ($M = 68.75\%$). For LC judgments, 20/80 (25%) fell within +/-5 mm, 39/80 (48.75%) fell within +/-10 mm, and 55/80 (68.75%) +/-15 mm of the initial sample ratings. The mean group ICC coefficient for LC was .933 with 95% CI (.882, .969). Given the complex nature of perceptual rating tasks, in addition to the present study's use of a more conservative approach to reliability

Table 4.2a*Agreement for Listener Ratings of Speech Acceptability (ACC)*

	+ / - 5	+ / - 10	+ / - 15	%
L1	2		2	100
L2	1	1	1	75
L3	2		1	75
L4	1	1		50
L5		3		75
L6	1	1	1	75
L7	2	1		75
L8	1	1		50
L9	1	1		50
L10	1		1	50
L11		2	2	100
L12	2		1	75
L13	4			100
L14	1		2	75
L15		1	1	50
L16	2		2	100
L17	3	1		100
L18		3		75
L19	2			50
L20	1		1	50
	27/80 (33.75%)	43/80 (53.75%)	58/80 (72.5%)	M = 72.5%

Note. L = listener.

Table 4.2b*Agreement for Listener Ratings of Listener Comfort (LC)*

	+/- 5	+/- 10	+/- 15	%
L1	1		1	50
L2	2	1		75
L3	1		1	50
L4	2	1		75
L5	3	1		100
L6	2	1	1	100
L7	1		2	75
L8		1	1	50
L9	1	1		50
L10	1		1	50
L11	1	1	2	100
L12	1	3		100
L13		3	1	100
L14	2	1		75
L15			2	50
L16	1		1	50
L17	1		1	50
L18		1	1	50
L19		3		75
L20		1	1	50
	20/80 (25%)	39/80 (48.75%)	55/80 (68.75%)	M = 68.75

Note. L = listener.

analysis using +/- 15 points on the VAS scaling procedures, intra-rater and inter-rater reliability were judged to be sufficient.

Results

Listener Ratings

Speech acceptability. The mean ACC rating was 60 ($SD = 15.1$) for the HS condition and 64 ($SD = 13.3$) in the CS condition. Results from a repeated measures ANOVA indicated that there was a significant effect of speaking condition on listener judgments of ACC, $F(8) = 6.96$, $p < .05$, partial $\eta^2 = .465$. The magnitude of the effect indicates that speaking condition demonstrated a large effect on ACC (Cohen, 1988). Post-hoc testing revealed that ACC scores were significantly higher when EL speakers used CS ($p < .05$), indicating that listeners judged CS to be more unacceptable.

Listener comfort. The mean LC rating by listeners for HS was 59 ($SD = 14.8$) and for the CS condition 61 ($SD = 12.4$). A repeated measures ANOVA indicated that speaking condition did not have a significant effect on listener judgments of LC.

Correlational Analyses. Data indicate a strong, statistically significant correlation between judgments of ACC in HS and CS, $r = 0.982$, $p < 0.001$ (see Figure 4.1). Similarly, data indicate a strong, statistically significant correlation between judgments of LC in HS and CS, $r = 0.962$, $p < 0.01$ (see Figure 4.2).

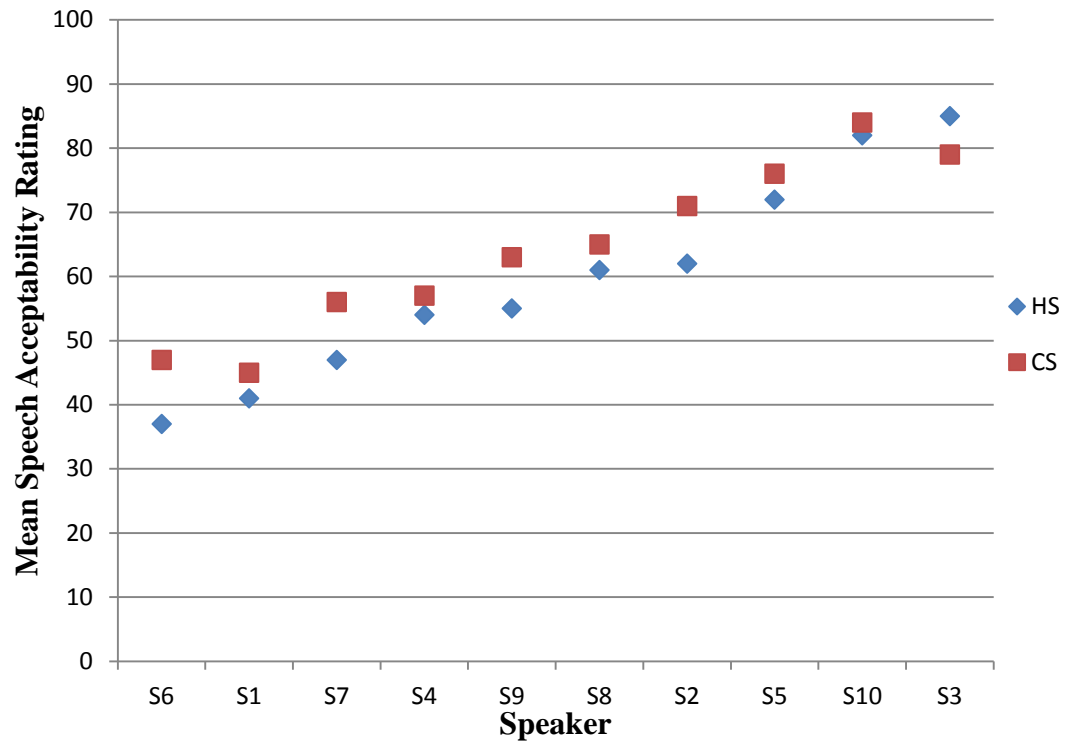


Figure 4.1. Mean listener ratings of speech acceptability (ACC) for electrolaryngeal speakers between habitual speech (HS) and clear speech (CS) conditions. Speaker ratings are arranged from lowest to highest.

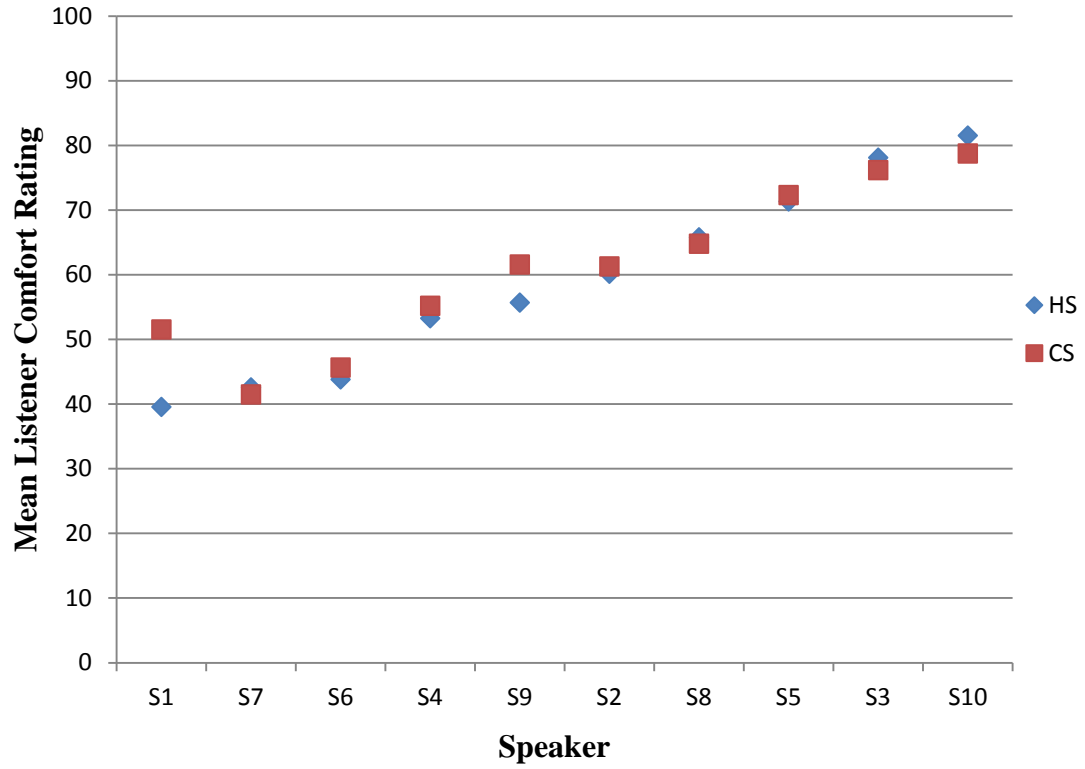


Figure 4.2. Mean listener ratings of listener comfort (LC) for electrolaryngeal speakers between habitual speech (HS) and clear speech (CS) conditions. Speaker ratings are arranged from lowest to highest.

Discussion

This study sought to determine whether listeners' auditory-perceptual ratings of ACC and LC differed when EL speakers produced speech using HS versus CS. This was achieved by having naïve listeners provide VA scaled judgments of voice recordings produced by EL speakers in both HS and CS conditions across two counterbalanced listening sessions. These findings indicate that when EL speakers are instructed to produce CS, listeners do not find the resulting speech to be less comfortable to listen to when compared to these speakers' HS. However, listeners did judge EL speakers to be less acceptable when they use CS. While these two perceptual features share some commonalities relative to the specific definitions as used in past studies (Bennett & Weinberg, 1973; O'Brian et al., 2003) as well as the present investigation, the significant findings for ACC but not LC indicate that ACC and LC might represent unique entities. This finding is important for several reasons.

First, the use of CS by EL speakers in the present experiment was based on previous research that has reported improvements in SI for individuals with a variety of communication disorders, as well as those with hearing impairment (Beukelman et al., 2002; Payton et al., 1994; Picheny et al., 1985; Tjaden et al., 2014). However, to date, the CS paradigm has not been applied to postlaryngectomy alaryngeal populations. Given the general nature of EL speech, one's use of CS as a production strategy would appear to be a viable therapeutic technique for these speakers. This is because laryngectomees who are learning to use an EL device are initially instructed to slow their rate of speech and over-articulate when learning how to produce speech (Doyle, 1994; 2005). A reduction in articulatory rate and the over-articulation of speech sounds are reasons why the use of

CS has been suggested to improve SI (Lam & Tjaden, 2013; Picheny, Durlach, & Braida, 1986). That is, reduction of one's overall articulatory rate is due to an increase in two factors; first, the lengthening of speech sound durations and second, the number of inserted pauses during a given utterance (Picheny et al., 1986). Ultimately, the productive modifications secondary to use of the CS are believed to enable improved coordination of the subsystems involved during speech (Tjaden et al., 2014). This in turn is believed to optimize the speech produced in an effort to aid the listener in understanding the speaker's intended message regardless of the category of speech disorder exhibited. Thus, CS is a phenomenon that offers potential advantages to both the speaker and the listener with the end product being improved communication. Although it is not anticipated that EL speech will be fully intelligible, isolated improvement in speech sound productions secondary to use of CS may hold considerable promise for improvements in the speaker's overall intelligibility.

Although the concept of CS has not been employed previously with postlaryngectomy alaryngeal speakers, its application would appear to be of some importance to laryngectomees who use an EL. Of particular concern here is the fact that when using an EL, the speaker must coordinate articulatory movements and speech rate while at the same time directly (and manually) manipulating an external, electronic voicing method. Research that has studied the EL source signal as it passes through neck tissues (i.e., the frequency response function) has suggested that this energy transfer also can impact its acoustic characteristics (e.g., attenuation of higher frequencies) (Meltzner, Kobler, & Hillman, 2003). A slower rate of speech and over-articulation, then, could assist EL speakers to maximally utilize a degraded speech signal to maximize signal

transmission into the vocal tract where articulation occurs. Given that EL speakers have been shown to consistently exhibit reduced SI related to normal speech and other alaryngeal speech methods (Barney, Haworth, & Dunn, 1959; Kalb & Carpenter, 1981; McCroskey & Mulligan, 1963; Shames, Font, and Matthews, 1963; Weiss & Basili, 1985; Weiss, Yeni-Komshian, & Heinz, 1979), attempts to improve or optimize the EL speech signal using CS appears to be warranted. Additionally, the EL generates a relatively consistent source signal, which has led to its identification by listeners as being monotonous and robotic (Bennett & Weinberg, 1973; Meltzner & Hillman, 2005). When these factors are combined, listeners are confronted with not only an unusual electronic speech signal that may be degraded at the phonemic level, but one that has inherent physical limitations that place greater demands on the listener. Consequently, the present work was designed to assess “quality” aspects of the EL voice and speech signal to determine if CS inadvertently creates another level of perceptual challenge for the listener.

Regardless of alaryngeal speech mode, any therapeutic attempt to improve postlaryngectomy speech should be mindful of potentially introducing features into modified speech that will further impact communication. Hanson, Beukelman, Fager, and Ullman (2004) stated that, “[i]f partner attitudes toward a communication strategy are negative, the behavioral tendency may be to reject the speaker” (p. 162). EL speakers must already rely on an alaryngeal voice source that is perceived as robotic and monotone in nature (Bennett & Weinberg, 1973). Thus, attempts to improve alaryngeal speech in general, and EL speech in specific, should not introduce additional changes that further challenge the listener’s ability to accurately receive the speaker’s message. If use of CS

further degrades EL communication, listeners could become more uncomfortable or perceive speech to be more unacceptable.

While using CS, EL speakers must seek to maintain a natural communication exchange while simultaneously making their speech clearer through a reduction in articulatory rate and over-articulation of speech sounds. The present work was the first study to investigate the relationship between the use of CS and its effect on listener judgments of ACC and LC in EL speakers. Previous research on articulatory rate has suggested that even when individuals with dysarthria are using a slower-than-normal rate, they can be perceived as less natural or acceptable (Dagenais, Brown, & Moore, 2006; Hanson et al., 2004). Tjaden et al. (2014) reported that individuals using CS or a slower-than-normal rate of speech can result in poorer speech severity ratings (i.e., a perceptual composite involving voice quality, resonance, articulatory precision, and speech rhythm), regardless of improvements in intelligibility. Thus, while modifications in one's speech can be modified using CS, a threshold may exist in which the results of this modification create other communication concerns relative to dyadic interactions. Therefore, the findings of the present study revealed that EL speakers are perceived to be less acceptable when they use CS compared to HS. This did not, however, translate to significant differences in ratings of LC.

While data also suggest a strong and significant correlation between listener judgments of ACC and LC in both HS and CS conditions, ACC ratings that ran contrary to LC ratings might suggest that listeners are able to differentiate ratings of ACC and LC despite the more global and somewhat overlapped definition for each feature. Given these findings, our data may provide evidence to suggest that the auditory-perceptual

dimensions measured ultimately address unique perceptual entities. That is, listeners are able to accurately and uniquely rate ACC and LC according to their unique descriptive properties within the definitions provided to listeners.

For ratings of ACC, listeners must make equally weighted judgments based on a perceptual composite involving pitch, rate, understandability, and voice quality. That is, listeners must make judgments that give similar consideration to each of the four perceptual features within this composite and not allow any individual feature to be the sole reason for their judgment. For ratings of LC, listeners are asked how comfortable they would be listening to speech in a social situation. This definition is similar to ACC in that listeners are forced to think about the manner that the speakers are speaking, a decision that could involve speech rate, understandability, and/or voice quality. LC, however, is unique in that it is broad enough to provide listeners with the freedom to make judgments based on the perceptual features they feel are most important without drawing their attention to a specific aspect of the speech being rated.

Second, judgments of ACC and LC provide information that may add to the potential effect of CS. For example, the definition of LC lends itself to a contextually-based social situation, whereby listeners must indicate if they would be comfortable speaking with the EL speakers using HS and CS. ACC on the other hand, targets specific aspects of voice that requires greater consideration for specific perceptual features on the listener's part. One of the hallmark features of CS is a volitional change to reduce speech rate; this is also one of the perceptual features that listeners must consider when making ACC judgments. In the present study, although speech rate was not controlled in the CS condition, EL speakers were allowed to modify their rate as needed to make their speech

“clearer”. A significant effect of CS on ACC ratings for EL speakers could indicate that listeners potentially penalized EL speakers for volitionally attempting to reduce their speaking rate further. That is, naïve listeners might have adjusted their ratings so as to focus primarily only one feature (e.g., speech rate) rather than all features of the composite ACC definition. However, anecdotal reports from individual participants after the completion of listening sessions indicated that listeners focused on several aspects of the signal. In fact, listeners’ main concerns related to the ACC of EL speech include device noise, rate, pitch, and intelligibility.

Third, research indicates that naïve listeners are able to make reliable judgments with consistent perceptual strategies (Kreiman et al., 1993). However, this same group might also make judgments according to an inherent metric based on normal (rather than pathological) voices (Kreiman et al., 1993). Given that the listeners in the present study lack training and experience in alaryngeal speech, it is important for future research to consider how ‘natural’ or ‘unnatural’ the voices could have been perceived, how this differs from ACC and LC. This could be examined by assessment of EL speech based on a perceptual feature termed ‘speech naturalness’ (NAT) (Martin, Haroldson, & Triden, 1984).

NAT was described by Martin et al. (1984) during the development of a 9-point Speech Naturalness Scale (*1* = highly natural, *9* = highly unnatural) for evaluating speech stimuli produced by persons who stutter (PWS). In order to make judgments of NAT, listeners were asked, “Make your judgment based on how natural or unnatural the speech sounds to you.” (Martin et al., 1984, p. 54). No further information or definition regarding NAT was typically provided in this prior work. More recently, O’Brian et al.

(2003) compared LC to NAT, and found that the LC scale used elicited a wider range of scores. The authors indicated that LC is a unique perceptual entity when compared to NAT, as LC involves more variables to consider than NAT (or, 'how speech sounds'). NAT, however, has been more clearly defined in several studies involving perceptual assessment of NAT in alaryngeal speakers. Specifically, Eadie and Doyle (2002) defined NAT as, "a perceptually derived, overall description of prosodic accuracy." (p. 1091). Given that Eadie and Doyle (2002) included prosodic characteristics of speech in their definition of NAT, this perceptual feature could share more similarities with ACC than LC; that is, ACC involves consideration for rate (i.e., a prosodic speech element) and pitch, which are key features concerning the prosody of speech (Lehiste, 1976). Although EL speech has been perceived as 'unnatural' according to listeners due to a monotone and robotic quality (Bennett & Weinberg, 1973), it is believed that NAT ratings of EL speakers using CS would produce similar ratings of ACC obtained in the present experiments. Since EL speech is deemed to be generally unacceptable to listeners as a result of a slower rate and reduced SI, alongside the presence of device noise and a robotic quality, this is very 'unnatural' compared to normal speech. Therefore, there are more perceptual features that can provide value in the perceptual assessment of alaryngeal speakers other than LC and ACC; NAT is one such example.

Overall, the findings from the current study suggest that listeners make distinct perceptual judgments of ACC of EL speech and how 'comfortable' they are while listening to EL speech using either HS or CS. In addition, the results suggest that ACC and LC are perceptually-unique, and therefore, our findings suggest that they serve to measure unique perceptual entities.

Clinical and Research Implications

Our findings reveal that a significant difference exists between HS and CS conditions for judgments of ACC, while no differences were noted for LC. These data would appear to provide initial evidence suggesting that volitional attempts to improve EL speech using CS do not result in any negative changes in all auditory-perceptual listener judgments of EL speech. Although speech intelligibility was not the target of this study, we did attempt to define it objectively through listener evaluation of a select set of stimuli. Current findings also suggest that the productive changes that occur in CS do not have a negative impact on the listener relative to certain composite assessments of speech (i.e., no differences found for LC). Employing CS as a remediation strategy to enhance EL speech may be of some benefit, but it may similarly introduce some new considerations into communication with others. This has been shown in previous research with other communication disorders (Beukelman et al., 2002; Picheny et al., 1985; Tjaden et al., 2014).

The strong relationship between listener judgments of ACC and LC also provides support for the use of scaled measurements to assess the impact of speech rehabilitation on individuals using the EL postlaryngectomy. Therefore, CS could potentially improve the SI of EL speakers without negatively impacting some auditory-perceptual listener judgments. In addition, although speech rate was not specifically controlled throughout the CS condition, previous research has found that there are distinct timing differences between CS and conversational speech (Picheny et al., 1986). That is, speakers using CS were found to have rates of 90 to 100 words per minute (wpm) when compared to conversational speech rates of 160 to 200 wpm (Picheny et al., 1986). Therefore, future

research should consider the specific articulatory rates during CS that might negatively impact listener perceptions of LC. For example, is there a specific threshold for speech rate (e.g., syllables per second) whereby listeners perceive CS to be significantly less acceptable and/or less comfortable to listen to when compared to conversational speech? Overall, the findings from this study suggest that CS might impact some auditory-perceptual listener judgments of EL speech.

References

- Barney, H. L., Haworth, F. E., & Dunn, H. K. (1959). An experimental transistorized artificial larynx. *Bell System Technical Journal*, 38(6), 1337-1356.
- Bennett, S., & Weinberg, B. (1973). Acceptability ratings of normal, esophageal, and artificial larynx speech. *Journal of Speech and Hearing Research*, 16, 608-615.
- Beukelman, D. R., Fager, S., Ullman, C., Hanson, E., & Logemann, J. (2002). The impact of speech supplementation and clear speech on the intelligibility and speaking rate of people with traumatic brain injury. *Journal of Medical Speech-Language Pathology*, 10, 237-242.
- Cole, D., Sridharan, S., Moody, M., & Geva, S. (1997). Applications of noise reduction techniques for alaryngeal speech enhancement. *Proceedings of the IEEE Telecon 1997 Region 10 Annual Conference, Speech and Image Technologies for Computing and Telecommunications*, 2, 491-494.
- Dagenais, P. A., Brown, G. R., & Moore, R. E. (2006). Speech rate effects upon intelligibility and acceptability of dysarthric speech. *Clinical linguistics & phonetics*, 20(2-3), 141-148.
- Doyle, P.C. (1994). Foundations of voice and speech rehabilitation following laryngeal cancer. San Diego: Singular.
- Doyle, P.C. (1999). Postlaryngectomy speech rehabilitation: Contemporary considerations in clinical care. *Journal of Speech-Language Pathology and Audiology*, 23, 109-116.

- Doyle, P. C. (2005). Clinical procedures for training use of the electronic artificial larynx. In P. C. Doyle & R. L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing*. Austin, TX: Pro-Ed.
- Doyle, P.C., & Eadie, T.L. (2005). The perceptual nature of alaryngeal voice and speech. In P.C. Doyle, & R.L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing*. Austin, TX: Pro-Ed.
- Doyle, P.C., Skidmore, E., Elliott, H., Senchuk, C., Day, A.M.B., Bornbaum, C., Dzioba, A., & Sleeth, L. (2011, November). Relationship between listener comfort and voice severity in tracheoesophageal speech. Presented at the American Speech and Hearing Association Convention, San Diego, CA.
- Eadie, T. L., & Doyle, P. C. (2002). Direct magnitude estimation and interval scaling of naturalness and severity in tracheoesophageal (TE) speakers. *Journal of Speech, Language, and Hearing Research, 45*(6), 1088-1096.
- Eadie, T. L., & Doyle, P. C. (2004). Auditory-perceptual scaling and quality of life in tracheoesophageal speakers. *The Laryngoscope, 114*(4), 753-759.
- Eadie, T. L., Nicolici, C., Baylor, C., Almand, K., Waugh, P., & Maronian, N. (2007). Effect of experience on judgments of adductor spasmodic dysphonia. *Annals of Otolaryngology, Rhinology and Laryngology, 116*, 695–701.
- Eadie, T. L., Otero, D., Cox, S. R., Johnson, J., Baylor, C., Yorkston, K., & Doyle, P. C. (in press). The relationship between communicative participation and post-laryngectomy speech outcomes. *Head and Neck*.

- Espy-Wilson, C. Y., Chari, V. R., MacAuslan, J. M., Huang, C. B., & Walsh, M. J. (1998). Enhancement of electrolaryngeal speech by adaptive filtering. *Journal of Speech, Language, and Hearing Research, 41*, 1253–1264.
- Failla, S. (2015). Randomizer [software]. London, Ontario.
- Goldstein, L. P. (1978). Part D: Intelligibility. In S.J. Salmon and L. P. Goldstein (Eds), *The artificial larynx handbook*. New York: Grune & Stratton.
- Hanson, E., Beukelman, D. R., Fager, S., & Ullman, C. (2004). Listener attitudes and speaking rate of people with traumatic brain injury. *Journal of Medical Speech-Language Pathology, 12(4)*, 161-166.
- Hillman, R. E., Walsh, M. J., Wolf, G. T., Fisher, S. G., & Hong, W. K. (1998). Functional outcomes following treatment for advanced laryngeal cancer. Part I - Voice preservation in advanced laryngeal cancer. Part II - Laryngectomy rehabilitation: The state of the art in the VA system. Research Speech-Language Pathologists. Department of Veterans Affairs Laryngeal Cancer Study Group. *Annals of Otology, Rhinology and Laryngology Supplement, 172*, 1-27.
- Kalb, M.B., & Carpenter, M.A. (1981). Individual speaker influence on relative intelligibility of esophageal speech and artificial larynx speech. *Journal of Speech and Hearing Disorders, 46*, 77-80.
- Kent, R. D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American Journal of Speech-Language Pathology, 5(3)*, 7-23.

- Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. *Journal of Speech, Language, and Hearing Research*, 36(1), 21-40.
- Lam, J., & Tjaden, K. (2013). Acoustic-perceptual relationships in variants of clear speech. *Folia phoniatrica et logopaedica: official organ of the International Association of Logopedics and Phoniatrics (IALP)*, 65(3), 148-153.
- Lehiste, I. (1976). Suprasegmental features of speech. In N. Lass. (Ed.), *Contemporary issues in experimental phonetics*. New York, NY: Academic Press.
- Martin, R. R., Haroldson, S. K., & Triden, K. A. (1984). Stuttering and speech naturalness. *Journal of Speech and Hearing Disorders*, 49(1), 53-58.
- Mazzoni, D., & Dannenberg, R. (2013). Audacity 2.0.5 [software]. Pittsburgh, PA: Carnegie Mellon University. Retrieved from: <http://audacityteam.org/>
- McCroskey, R. L., & Mulligan, M. (1963). The relative intelligibility of esophageal speech and artificial-larynx speech. *Journal of Speech and Hearing Disorders*, 28(1), 37-41.
- Meltzner, G. S., & Hillman, R. E. (2005). Impact of aberrant acoustic properties on the perception of sound quality in electrolarynx speech. *Journal of Speech, Language, and Hearing Research*, 48(4), 766-769.
- Meltzner, G. S., Kobler, J. B., & Hillman, R. E. (2003). Measuring the neck frequency response function of laryngectomy patients: Implications for the design of electrolarynx devices. *Journal of the Acoustical Society of America*, 114(2), 1035-1047.

- Mendenhall, W. M., Morris, C. G., Stringer, S. P., Amdur, R. J., Hinerman, R. W., Villaret, D.B., & Robbins, K. T. (2002). Voice rehabilitation after total laryngectomy and postoperative radiation therapy. *Journal of Clinical Oncology*, *20*, 2500–2505.
- Nagle, K. F., Eadie, T. L., Wright, D. R., & Sumida, Y. A. (2012). Effect of fundamental frequency on judgments of electrolaryngeal speech. *American Journal of Speech-Language Pathology*, *21*(2), 154-166.
- O'Brian, S., Packman, A., Onslow, M., Cream, A., O'Brian, N., & Bastock, K. (2003). Is listener comfort a viable construct in stuttering research? *Journal of Speech, Language, and Hearing Research*, *46*, 503-509.
- Payton, K. L., Uchanski, R. M., & Braida, L. D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America*, *95*, 1581–1592.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing. I. Intelligibility differences between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, *28*, 96–103.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, *29*, 434–446.
- Qi, Y. Y., & Weinberg, B. (1991) Low-frequency energy deficit in electrolaryngeal speech. *Journal of Speech, Language, and Hearing Research*, *34*(6), 1250-1256.

- Saikachi, Y., Stevens, K. N., & Hillman, R. E. (2009). Development and perceptual evaluation of amplitude-based F0 control in electrolarynx speech. *Journal of Speech, Language, and Hearing Research, 52*(5), 1360-1369.
- Shames, G. H., Font, J., & Matthews, J. (1963) Factors related to speech proficiency of the laryngectomized. *Journal of Speech and Hearing Disorders, 28*, 273-287.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.
- Tardy-Mitzell, S., Andrews, M. L., & Bowman, S. A. (1985). Acceptability and intelligibility of tracheoesophageal speech. *Archives of Otolaryngology, 111*(4), 213-215.
- Tjaden, K., Sussman, J.E., & Wilding, G. E. (2014). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research, 57*, 779-792.
- Verdolini, K., Skinner, M. W., Patton, T., & Walker, P. A. (1985). Effect of amplification on the intelligibility of speech produced with an electrolarynx. *Laryngoscope, 95*, 720-726.
- Watson, P. J., & Williams, S. E. (1987). Laryngectomees' and nonlaryngectomees' perceptions of three methods of alaryngeal voicing. *J Comm Disord., 20*, 295-304.
- Ward, E. C., Koh, S. K., Frisby, J., & Hodge, R. (2003). Differential modes of alaryngeal communications and long-term voice outcomes following pharyngolaryngectomy and laryngectomy. *Folia Phoniatica et Logopaedica, 55* (1), 39-49.

- Weiss, M. S., & Basili, A. M. (1985). Electrolaryngeal speech produced by laryngectomized subjects: Perceptual characteristics. *Journal of Speech and Hearing Research, 28*, 294-300.
- Weiss, M. S., Yeni-Komshian, G. H., & Heinz, J. M. (1979). Acoustical and perceptual characteristics of speech produced with an electronic artificial larynx. *Journal of the Acoustical Society of America, 65*(5), 1298-1308.

Chapter 5

General Discussion and Interpretation of Findings

This chapter will discuss and integrate the findings from the present investigations involving the therapeutic application of clear speech (CS) with electrolaryngeal (EL) speakers. The discussion to follow will begin with a brief summary of the findings from each of the three experiments. Findings from the current experiments will also be interpreted with specific reference to the literature reviewed in Chapter 1. Interpretation of findings will be followed by a discussion of the potential relationship between changes in speech intelligibility (SI), acoustic changes to the EL speech signal, and auditory-perceptual ratings of speech acceptability (ACC) and listener comfort (LC) as a result of CS. Lastly, this chapter will conclude with the limitations of the present work, its clinical implications, and directions for future research.

General Overview

The present work was comprised of three studies that directly focused on the influence of CS on EL speakers. These three studies were designed to specifically assess the impact of CS on SI (Experiment 1), the acoustic characteristics of EL speech (Experiment 2), and its influence on auditory-perceptual judgments of naïve, normal-hearing listener (Experiment 3). The specific research questions addressed in each of these studies were:

- (1) Does CS facilitate improved word intelligibility of EL speakers? (Chapter 2)
- (2) Does CS alter the acoustic characteristics of words and vowels in EL speech? (Chapter 3)

(3) Does CS result in altered auditory-perceptual ratings by listeners, namely ACC and LC, for EL speakers? (Chapter 4)

The present studies were the first to investigate the influence of CS on EL speech. The rationale for each of the present investigations began as a response to the nature of the EL speech signal, in addition to the historical controversy over the use of EL speech as an inferior communication option postlaryngectomy (Berry, 1978; Duguay, 1978; Gates et al., 1982; Lauder, 1968). For example, EL speech contains numerous deficits across SI, acoustic output, and auditory-perceptual judgments (Barney, Haworth, & Dunn, 1959; Bennett & Weinberg, 1973; Doyle & Eadie, 2005; Meltzner & Hillman, 2005). Much of the research describes deficits in EL speech as the result of numerous design and use characteristics. These characteristics have generally been present since their inception.

EL devices, however, have continued to be immediate and viable sources of postlaryngectomy voice and speech production since their introduction in the 1950s (Barney et al., 1959; Doyle, 1994; Hillman, Walsh, Wolf, Fisher, & Hong, 1998; Meltzner & Hillman, 2005). Unfortunately, deficits in EL signal properties continue to impact EL speakers from attaining high levels of SI (Meltzner & Hillman, 2005). In fact, the majority of research on EL speakers' SI indicates wide variability in performance, often centered on a mean SI between 50 to 60% with a documented range of 16 to 90% (Barney et al., 1959; Weiss & Basili, 1985; Weiss, Yeni-Komshian, & Heinz, 1979; Yeni-Komshian, Weiss, & Heinz, 1975). The limited number of attempts to experimentally modify EL speech have focused on improving the acoustic features in an effort to improve SI and auditory-perceptual aspects of EL speech. These studies have led to more favourable perceptual judgments of EL speech. However, these pursuits have not

been met with improvements in SI (Beaudin, 2002; Espy-Wilson, Chari, MacAuslan, Huang, & Walsh, 1998; Wong, 2003). Since SI has been viewed as one of the most important aspect of communication, especially for EL speakers (Goldstein, 1978), Experiment 1 (Chapter 2) directly focused on the therapeutic application of CS and its influence on the SI of EL speakers; in this study, both word SI and SI by consonant position were assessed.

To assist in understanding the inherent signal changes that may have occurred from volitional changes in EL speakers' articulation (i.e., alterations that evolved from use of a reduced speech rate, over-articulation and increased mouth opening), Experiment 2 (Chapter 3) assessed the acoustic changes associated with CS. More specifically, this experiment focused on word and vowel durations, fundamental frequency and formant frequencies of vowels. These findings were compared with EL speakers' 'habitual' speech (HS).

Lastly, Experiment 3 (Chapter 4) was conducted to assess potential auditory-perceptual challenges to the listener as a direct result of modifications to the acoustic signal secondary to EL speakers' use of CS. That is, this experiment was concerned with whether or not listeners deemed speech produced using CS as comfortable and/or acceptable to listen to compared to HS. Since listeners must adjust to an already 'mechanical' and 'monotone' voice with EL speakers, Experiment 3 (Chapter 4) compared listener perceptions of the volitional changes to articulation and speech rate when EL speakers used HS and CS.

Collectively, the present investigations were an important step to further efforts aimed at improving various aspects of EL speech, namely SI, acoustic characteristics, and

listener judgments. Since communication with an EL device has been shown to have numerous deficits in SI, acoustics, and auditory-perceptual judgments, in addition to CS not being previously studied in alaryngeal speakers, the present investigations appeared to be a worthwhile endeavour. Therefore, the following section will describe CS in greater detail and discuss the potential benefits for its use with EL speakers.

The basis of CS is for speakers to slow their rate of speech in an effort to make communication more intelligible for listeners. This is primarily achieved by instructing speakers to reduce their rate of speech and over-articulate (Picheny, Durlach, & Braida, 1985; Lam & Tjaden, 2013). These adjustments are therefore assumed to optimize the speech production process with its direct influence on the listener's perception of speech. Research on CS has been reported to improve SI up to 11% for individuals with speech impairment (Beukelman, Fager, Ullman, Hanson, and Logemann, 2002; Hanson, Beukelman, Fager, & Ullman, 2004; Tjaden, Sussman, & Wilding, 2014) and 18 to 26% for individuals with hearing impairment (Payton, Uchanski, & Braida, 1994; Picheny et al., 1985). Thus, CS has been shown to not only be an effective strategy in the retrieval of the message in those with hearing loss, but also to improve the understandability of those who have deficits in the production of speech. Interestingly, there is a close connection between principles of CS and the initial training of laryngectomees to use an EL device; that is, in order to provide laryngectomees with effective and intelligible speech, they are initially instructed to use a slower rate of speech and over-articulate during communication (Doyle, 1994; 2005). Therefore, this therapeutic modification may assist in explaining how a reduced speaking rate for EL speakers may be ancillary to changes in articulation patterns while using an EL device (Doyle & Eadie, 2005).

Generally, alaryngeal speakers speak at a slower than normal rate than laryngeal speakers (Doyle & Eadie, 2005; Hillman et al., 1998; Robbins, Fisher, Blom, & Singer, 1984). For EL speakers, this is primarily due to speech-language pathologists' (SLPs) therapeutic emphasis on ensuring accurate articulation and potentially improve a listeners' ability to process EL speakers' messages (Doyle, 1994; Ward & Van As-Brooks, 2014). Amongst all three commonly used alaryngeal speaking options, EL and tracheoesophageal (TE) speakers speak around ~130 words per minute (wpm), while ES speakers often have speech rate of 90 to 114 wpm (Hoops & Noll, 1969; Robbins et al., 1984; Snidecor & Curry, 1959). Research comparing speech rate between EL and TE speakers has confirmed that listeners might judge EL speech rate to be perceptually slower than TE speakers (Williams & Watson, 1985). Overall, a rate of 130 wpm is generally considered to be slower than normal, laryngeal speakers (Robbins et al., 1984). Attempts to improve EL speech using CS which involves a reduced rate of speech may also foster the associated act of over-articulation. Thus, because of the interaction between varied elements of the CS process, alterations of speech rate may further facilitate improved communicative effectiveness in some EL speakers. Therefore, the following sections will describe and integrate the findings from the present experiments to understand if a CS benefit exists for EL speakers. Following the integration of findings, the clinical implications and directions for future research will be discussed.

Integration of Findings

Findings from the present work indicate that CS did not have a significant impact on the overall speech of EL speakers. In particular, Experiment 1 (Chapter 2) demonstrated that CS does not significantly impact the SI of words or WF consonants by

word position for the present EL speakers. On the other hand, two important findings emerged from Experiment 2 (Chapter 3). First, Experiment 2 revealed that CS had a significant impact on overall word durations. In addition, CS appeared to improve the voicing characteristics of EL speech to a certain degree for WI and WF consonants. Finally, Experiment 3 (Chapter 4) indicated that the use of CS can result in reductions in the perceived “acceptability” of EL voice when judged by naïve listeners.

Numerous reasons may potentially account for the lack of statistical significance across several areas investigated in the present experiments. First, the SI of words was only 1.3% greater in CS compared to HS, in addition to 1.6% WI consonants and 0.9% WF consonants. These results are in stark contrast to the previous research on CS demonstrating that a benefit of up to 11% in SI may be observed for individuals with dysarthria (Tjaden et al., 2014). The discrepancy between these two clinical populations suggests that on average, EL speakers did not modify their speech to an extent that was different from HS. Although the amount of speech rehabilitation following each participant’s laryngectomy is unknown, there is the potential that EL speakers were already producing a reduced speaking rate previously emphasized in their postlaryngectomy speech rehabilitation (Doyle, 1994). Therefore, some EL speakers did not derive further benefit in their SI from the use of CS.

Speech rate is one of the hallmark features of CS (Picheny et al., 1985). When speakers reduce their rate, not unexpectedly, phoneme and syllable durations also have been shown to increase (Kessinger & Blumstein, 1997; Miller, Green, & Reeves, 1986; Theodore, Miller, & DiSteno, 2009). Previous research also has indicated that speech rate while using CS is markedly different than habitual speech (HS); for example,

speaking rates using CS are reported to range from 90 to 100 wpm, while HS speech rates are approximately 160 to 200 wpm (Picheny, Durlach, & Braidia, 1986). Increases in phoneme and syllable durations are additive and, therefore, will systematically lengthen word and utterance durations. This has been shown to impact the perception of specific phonemes and members of phonetic categories (Miller & Volaitis, 1989). In addition, increasing word and utterance duration further highlights the relative importance of rate reduction as a key feature of CS to improve SI. Research has indicated that reductions in speech rate as a result of training normal speakers to use CS have nearly doubled vowel and sentence durations (Ferguson & Kewley-Port, 2002; Picheny et al., 1985). Relative to the findings from Experiment 2 (Chapter 3), EL speakers' were unable to use CS with a similar proficiency described in the literature. As a result, EL speakers' use of CS did not result in similar increases in the durational properties of phonemes and words.

Collectively, findings from Experiment 1 and Experiment 2 revealed that EL speakers produced significantly longer word durations in CS compared to HS. However, EL speakers were unable to improve SI as reported in previous work (Beukelman et al., 2002; Tjaden et al., 2014). Previous research investigating the use of CS led to an 11% increase in SI of individuals with dysarthria (Beukelman et al., 2002) and approximately 18 to 26% for individuals with hearing impairment (Payton et al., 1994; Picheny et al., 1985). It is interesting to note that CS is not known to bring about general, uniform changes in rate of speech (Picheny et al., 1986) and SI (Lam & Tjaden, 2013). This might partly explain the lack of significant differences in SI across words and all word positions for EL speakers. That is, half of the present EL speakers improved their SI, while the remaining half maintained or even slightly decreased their SI following instructions to

produce CS. Varied performance has been confirmed in previous research indicating that the general instructions to produce CS might be unclear to some speakers (Ferguson & Kewley-Port, 2002). For the purpose of the present investigations, however, it appears that the varied performance in SI was unavoidable, given that the instructions to produce CS was similar for all EL speakers. Further, no EL speaker appeared to misunderstand or express concerns regarding the CS instructions.

It is important to recognize that research also has shown that a reduced rate of speech is not the only important factor for improving SI while using CS (Lam, Tjaden, & Wilding, 2012; Lam & Tjaden, 2013; Krause & Braida, 2004). In fact, research suggests that speakers must be instructed to reduce speaking rate *in addition to* over-articulate (Lam et al., 2012; Lam & Tjaden, 2013) and increase mouth opening (Picheny et al., 1985) to derive the greatest benefit from CS. In the present series of investigations, it is possible that EL speakers only reduced their rate of speech, rather than using a combination of rate reduction, over-articulation, and mouth opening. This can be partly explained by the lack of significant differences in SI of words and identification of consonants in WF position in Experiment 1 (Chapter 2), in addition to the slight lengthening of word and phoneme durations when EL speakers used CS in Experiment 2 (Chapter 3). Although mouth opening was not directly measured, acoustic findings from Experiment 2 (Chapter 3) might assist in determining such general differences between CS compared to HS.

Experiment 2 (Chapter 3) focused on CS and its influence on acoustic characteristics of the EL signal (e.g., temporal, frequency, etc.). One method of indirectly assessing EL speakers' degree of mouth opening might lie in measures of vowel formant

frequency that were obtained from this experiment. Research has indicated the strong link that exists between mouth opening and proportional changes in the first formant frequencies (Stevens & House, 1955). More specifically, a wider the mouth opening will produce a larger, proportional increase in the first formant frequency. In order to produce CS, EL speakers were provided with instructions that emphasized a reduced rate of speech, over-articulation with an increase in mouth opening. If EL speakers over-articulated and opened their mouths to a greater degree in CS compared to HS, then proportional increases in vowel formant frequencies would be expected to occur in CS only.

After further examination of formant frequencies across all vowel stimuli, it appears that the F1 in HS *and* CS were higher than data for normal, laryngeal speakers reported by Peterson and Barney (1952). No notable changes in F1 values were observed between speaking conditions and F1 values were highly variable between speakers. It is possible that EL speakers were already demonstrating an increased mouth opening in HS, especially when compared to normative data on normal, laryngeal speakers. For example, Peterson and Barney (1952) reported an average first formant frequency of 270 Hz for /i/ when produced by male speakers. Data from the present group of EL speakers include the following first formant frequencies for /i/: 726.3 Hz for Servox speakers in HS and 723.3 Hz in CS and 616.3 Hz for Trutone users in HS and 619.4 Hz in CS. This data strongly suggests that EL speakers did not produce the expected articulatory changes when moving from HS to CS. Granted, Sisty and Weinberg (1972) presented vowel data for esophageal speakers, whose first formant frequencies for /i/ were 401 Hz. This study

demonstrated that formant frequencies are higher in alaryngeal speakers due to a shortened vocal tract postlaryngectomy

Aside from significant durational changes in CS, there were minimal changes in the acoustic structure between speaking conditions observed in the present series of experiments. There is a possibility that when EL speakers were provided with instructions to produce CS, some implemented a reduced rate in addition to an articulation pattern that was more exaggerated than that observed using HS. Contrarily, the remaining EL speakers might have only reduced their rate. This was observed acoustically given that some speakers (slightly) increased their F1 frequency when moving from HS to CS, whereas others slightly decreased F1 values across conditions. This data contributes to the literature indicating that there is also variability in performance when speaker use CS (Picheny et al., 1985). In addition, F1/F2 vowel plots were generally overlapped across all speakers in Experiment 2 across HS and CS. Research indicates that individuals with larger vowel spaces tend to have higher SI (Bradlow et al., 1996; Turner et al., 1995). In the current series of investigations, EL speakers had smaller vowel spaces alongside reduced SI. Therefore, vowel formant data from Experiment 2 and SI data from Experiment 1 provide a clearer picture of the relationship between SI and acoustics of EL speech, while highlighting the lack of predicted vowel trajectories in the absence of anticipated articulatory changes due to CS (e.g., lack of F1 increase suggests EL speakers did not use a wider mouth-opening).

Variability has been observed in applications of CS involving individuals with hearing impairment listening to normal speakers using CS (Ferguson & Kewley-Port, 2002; Picheny et al., 1985), and studies involving individuals with dysarthria (Beukelman

et al., 2002; Hanson et al., 2004; Lam & Tjaden, 2013; Tjaden et al., 2014). In order to justify the use of CS and to compare results from the present investigations to literature focused on individuals with dysarthria, it is important to understand the obvious differences (and similarities) between EL speakers and those with dysarthria.

The most obvious difference between EL speakers and individuals with dysarthria are the etiologies and resulting communication deficits in these populations. Generally, individuals with dysarthria speak with an anatomically-intact vocal tract and larynx. Although laryngeal deficits might be present in subtypes of dysarthria (e.g., flaccid dysarthria), the primary voice and speech deficits are based in the neuromuscular control over the speech mechanism (Darley, Aronson, & Brown, 1969). CS has been assessed in a variety of dysarthric speakers including those with Parkinson's disease and Multiple Sclerosis (Tjaden et al., 2014), in addition to TBI (Beukelman et al., 2002). Speech deficits for dysarthria often include a consistently reduced rate of speech and imprecise consonants (Darley et al., 1969). In contrast, EL speakers produce speech using an external, electronic speech aid that has 'mechanical' signal properties. In addition, EL voice and speech involves a degraded acoustic signal that is unlike laryngeal speech in its frequency, intensity, and harmonic-to-noise ratio (Bennett & Weinberg, 1973; Doyle & Eadie, 2005; Hillman et al., 1998; Meltzner & Hillman, 2005; Watson & Schlauch, 2009). Ultimately, these acoustic deficits can greatly impact perception by the listener (Bennett & Weinberg, 1973; Doyle & Eadie, 2005; Meltzner & Hillman, 2005; Williams & Watson, 1985). For example, the continuously voiced nature of the EL can lead to listener confusions at the phoneme-level (e.g., voiced-for-voiceless phoneme errors), in addition to listeners' ability to understand only 50 to 60% of an EL speakers' intended

message (Barney et al., 1959; Weiss & Basili, 1985; Weiss et al., 1979). There is, however, at least one commonality between individuals with dysarthria and EL speakers; both groups can benefit from general modification of their speech rate in order to achieve the most intelligible speech possible (Beukelman et al., 2002; Doyle, 1994, 2005). Although CS is often implemented with individuals with dysarthria, unfortunately, there is a lack of clinical research focused on rate modification for laryngectomees using an EL. Given the numerous deficits in dysarthric and EL speech, it is not surprising that there is wide variability in SI data obtained from previous research and the current investigation. Therefore, the next section will examine the variability in SI observed in EL speakers.

Variability in EL SI has been observed in previous investigations examining the relationship between acoustic modification of the EL speech signal and the resulting listener judgments (e.g., voice quality, listener preference) and overall SI (Beaudin 2002; Espy-Wilson et al., 1998; Meltzner & Hillman, 2005). In essence, even if attempts to improve acoustic and/or auditory-perceptual aspects of EL speech are undertaken, the electronic signal quality remains abnormal compared to laryngeal speech (Meltzner & Hillman, 2005). This may create challenges for the listener, especially those that have minimal experience communicating with alaryngeal speakers. Listeners must adjust to the collective differences in the frequency, intensity, and rate of EL speech (Qi & Weinberg, 1991; Saikachi, Stevens, & Hillman, 2009; Watson & Schlauch 2009; Weiss et al., 1979), which results in a general reduction of SI (Barney et al., 1959; Hillman et al., 1998; Weiss & Basili, 1985; Weiss et al., 1979). In addition, the parametric differences in the EL signal have negative implications for auditory-perceptual judgments

of listeners (Doyle & Eadie 2005). Therefore, the impact of CS on auditory-perceptual judgments of EL speech as well as findings from Experiment 3 (Chapter 4) will be addressed in the subsequent section.

Auditory-Perceptual Assessment Following the Application of CS

The general nature of EL speech presents numerous challenges for listeners. Doyle and Eadie (2005) have described that auditory-perceptual assessment is vital toward understanding the therapeutic success of postlaryngectomy rehabilitation. In the present investigation, Experiment 3 sought to assess listener judgments of ACC (Bennett & Weinberg, 1973) and listener comfort (LC) (O'Brian et al., 2003). While Experiments 1 and 2 demonstrated minimal changes in the SI and overall acoustic characteristics of EL speech, an important finding emerged from Experiment 3. That is, it appears listeners might have been sensitive to the durational changes in words that were found to be significantly different between speaking conditions.

First, ratings of LC required listeners to make judgments based on how comfortable they were while listening to a speaker in a suggested social situation (O'Brian et al., 2003). Data indicate that listeners' LC ratings were similar when EL speakers spoke during HS and CS conditions. Using a visual analogue scale (i.e., '0' representing 'very comfortable' to '100' representing very uncomfortable), this was demonstrated by listeners rating EL speech with a mean score of '59' when speaking in HS and a mean score of '61' while speaking in CS. In addition, the present study suggests that EL speakers are not penalized to a greater extent when using CS compared to HS. Although findings from the present study suggest that there are no differences in LC levels while listening to EL speakers use HS or CS, Experiment 4 (Chapter 3) revealed

that CS might negatively impact ACC. The following paragraph, then, will describe listener judgments of ACC, in addition to providing evidence to suggest that CS might negatively impact this perceptual feature when EL speakers use CS.

Ratings of ACC are based on a perceptual composite involving numerous considerations on part of the listener. For example, each listener is required to make judgments of EL speakers' pitch, rate, understandability, and voice quality (Bennett & Weinberg, 1973). Experiment 3 (Chapter 4) found that listeners deem EL speech produced using CS significantly less acceptable when compared to HS samples.

Overall, Experiment 3 supports the notion that listeners are sensitive to the 'unnatural' and 'mechanical' nature of EL speech, which is acoustically and perceptually different than normal, laryngeal speech. To account for the significant effect of speaking condition on ACC, however, it is important to evaluate the potential changes that occurred during CS. First and foremost, EL speech is generally known to have deficits in each of the areas described by the definition of ACC (e.g., pitch, rate, understandability, and voice quality). When considering the pitch of EL speech, listeners must assess a speech signal that is introduced via an external device that has been shown to have deficits in frequency energy, range and variation (Meltzner & Hillman, 2005; Nagle, Eadie, Wright, & Sumida, 2012; Watson & Schlauch, 2009; Weiss et al., 1979). Further, EL speakers are required to speak slower and over-articulate during communication with an EL device, and this is the focus of speech rehabilitation postlaryngectomy (Doyle, 1994). The general deficits in EL speech impact upon EL users' prosody, including their ability to produce intonation, stress, rhythm, and appropriate word junctures during speech. Each of these suprasegmental features are often realized through variations in F_0 ,

intensity, and lengthening of speech sounds (Lehiste, 1976). Given that naïve listeners are considered to have an inherent perceptual metrics well-matched to normal (rather than pathological) voices (Kreiman et al., 1993), it is possible that the numerous limitations of EL speech (e.g., frequency, device noise) impacted perceptual judgments. For example, one listener claimed that “a low acceptability” was “shared across [voice] samples”. When asked to comment further about why the listener gave low ACC ratings, it was because the voices were “robotic” with “not much pitch differences in voice”. Since F_0 is an important aspect of realizing all aspects of prosody, in addition to the fact that intensity levels were monitored during recording and playback to listeners, the frequency deficits in EL speech, alongside device noise and overall robotic quality, proved to be less acceptable, and particularly in CS. This is likely due to the fact that the robotic and monotone nature of EL speakers’ voices are far removed from the prosodic normal voices that naïve listeners are used to hearing. Lastly, some listeners even commented that the “slower voices” were deemed to be “less acceptable” and “more uncomfortable” during the listening session. Durational changes in CS have been often cited to account for prosody-related SI benefits (Mayo, Aubanet, & Cooke, 2012; Picheny et al., 1986). Further, the present studies highlight the importance of Meltzner and Hillman’s (2005) that improvements in voice quality occur when EL speakers use a device that has an improved low-frequency component, can vary frequency, and produce speech with less noise radiating from the device. These are enhancements that enable EL speakers to approximate more typical prosodic aspects of normal speech.

EL speakers are known also to have reduced SI when compared to other alaryngeal speakers and normal, laryngeal speakers. As part of initial speech therapy, EL

speakers are initially trained to use a slower rate of speech in order to be effective communicators with their EL device. However, EL speech is perceived to have numerous acoustic deficits that contribute to listener descriptions of ‘robotic’ and ‘monotone’ (Bennett & Weinberg, 1973; Hillman et al., 1998). Since the devices used in the present investigation were not modified by the investigator in any way, findings from the present investigation suggest that the production of CS is responsible for the increased levels of unacceptability amongst listeners. This is the result of listeners rating the same 10 EL speakers in HS and CS, in addition to each EL speaker using the same device for both conditions. For example, if a speaker used a Servox in HS, they used it again in CS. The only modification was the instructions to produce CS. Furthermore, the changes in speech rate introduced through CS instruction facilitated a greater divide between EL speakers and normal, laryngeal speakers. That is, word duration data from Experiment 2 suggest that EL speakers spoke significantly slower in CS compared to HS. As a result, the change in speech rate further challenged listeners, and thereby, impacted listener judgments of ACC, but not LC. Ratings of ACC, by definition, increase the likelihood that listeners specifically consider speaking rate when forming their judgments. Thus, the slower speech rate used by EL speakers while producing CS would appear to have been more readily perceived by listeners while making judgments of ACC. Even though wide variability was noted in individual speaker performance, it is the collective impact of CS (e.g., slower rate of speech, over-articulation, and increased mouth opening) that typically has been shown to result in negative listener judgments. This is not only the case for the present experiments, but in previous research involving individuals with dysarthria (Hanson et al., 2004).

Previous research has indicated that, although speech enhancement or supplementation strategies may improve aspects of an individual's speech (e.g., SI), this does not translate to being the most preferred or acceptable strategy (Hanson et al., 2004). Hanson et al. (2004) found that 60 judges (i.e., 15 naïve listeners, 15 SLPs, 15 allied health professionals, and 15 family members) rated videotaped samples of nine individuals with moderate-to-severe dysarthria using various supplementation strategies (e.g., alphabet supplementation, clear speech, topic supplementation, and habitual speech). Findings indicated that listeners rated the most beneficial strategy (e.g., alphabet supplementation) as 'unacceptable', even in the presence of improved SI (Hanson et al., 2004). This is potentially the case for CS and the effect it had on SI in the present study; that is, while CS provided an improvement in SI for EL speakers, CS produced less acceptable speech based on listener judgments. Therefore, there are some data to suggest that listeners could be sensitive to productive changes when EL speakers use CS. This in turn can lead listeners to negatively perceive even the most effective strategies intended to improve SI.

Due to complex nature of EL speech, some research has even indicated the opposite effect. Specifically, experimental modifications to the acoustic signal of EL speech have led to more favorable listener ratings, but this was not matched by improvements in SI. This was demonstrated by Wong (2003), who studied the SI of EL speakers using a Servox EL and a modified prototype EL. The modified device involved acoustic adjustments to frequency, device noise, and variable frequency control. Previous research had indicated that listeners preferred the modified device in terms of overall voice quality (Beaudin, 2002). However, a follow-up study by Wong (2003) revealed

that, in spite of improved listener judgments related to quality, speakers using the unmodified device were judged to be more intelligible (e.g., SI score of 66% using the unmodified Servox vs. SI of 59% for the modified prototype EL). We can compare the SI and perceptual findings from the research by Beaudin (2002) and Wong (2003) to the current investigation in several ways. Collectively, Beaudin (2002) and Wong (2003) found favourable improvements in voice quality alongside an overall decrease in SI using the modified EL device. The present investigation, however, found that CS did not negatively impact LC judgments while at least maintaining SI for HS and CS. The biggest difference is the negative impact CS had on ACC. While the modified device in Beaudin's (2002) study favourably impacted listener perceptions, EL speakers were deemed less acceptable to listen to as a result of the volitional changes to speech production in the present investigation. Therefore, understanding these perceptual differences might allow further refinement of attempts (both therapeutic and experimental) to improve alaryngeal speech in general.

The above-mentioned comparison recognizes the difficulty in improving EL speech and further recognizes the need to improve this alaryngeal communication method; where one aspect of EL speech can be modified (e.g., acoustic signal properties, productive aspects of speech), other aspects due to the numerous deficits in parametric measurements and transmission characteristics are limited. Further inquiries must be made toward improving EL voice and speech through device modification and/or volitional changes to EL speakers' communication. Given the resulting negative change in listener judgments as a result of the current or previous attempts to improve EL speech, the additional challenges presented to the listener must be considered. The discussion to

follow, then, will describe the collective findings from the Experiment 3 (Chapter 4) relative to how EL speakers' use of CS might present more challenges to listeners.

When attempting to improve the complex acoustic and perceptual characteristics of EL speech, consideration must be given to the notion that additional changes are being introduced with CS instruction. For example, EL speakers in the present investigation were instructed to slow their rate of speech, over-articulate, and increase mouth opening in an effort to speak more clearly (Lam & Tjaden, 2013; Picheny et al. 1985). These changes add additional challenges as listeners attempt to accurately retrieve an EL speaker's message. Given that listeners might have been more sensitive to changes in speaking rate, CS can be viewed as a further degradation to EL speech, and consequently, result in a negative impact on listener judgments of both LC and ACC. Thus, the interaction between changes in the acoustic signal of EL speech and their impact on perceptual judgments must also be considered when attempting to improve any aspect of EL communication.

In summary, the previous section discussed findings from Experiment 4, which involved listener judgments of EL speakers using CS; more specifically, listeners made judgments of LC and ACC as EL speakers used HS and CS. It was revealed that listeners did not rate their comfort levels differently between EL speakers' productions in HS and CS. However, listeners judged CS to be less acceptable than HS. Findings suggest that modifying EL speakers' speech rate, one of the hallmark features that defines the ACC composite, might have contributed to the difference in listener judgments between speaking conditions. Together, findings from all three experiments suggest that CS might potentially be useful in improving EL speakers' SI. Even though CS has the ability to

significantly lengthen the durations of words and some vowels with the potential to alter vowel formant frequencies of some EL speakers, this cannot overcome the general nature of the EL speech signal. Listener's accuracy in identifying words in the light of durational improvements, however, does not result in favorable listener judgments.

Limitations of the Present Work

The most notable limitations of the present work are based on the elicitation of CS, limited practice time to produce CS, and the use of word stimuli. First, the instructions used to elicit CS from EL speakers were provided in the absence of controlling or modifying speech rate in any other way. This allowed EL speakers to alter their rate based solely on the single set of instructions. The similar word SI scores between HS and CS in Experiment 1 (Chapter 2) and similar temporal and frequency data between HS and CS from Experiment 2 (Chapter 3) suggest that EL speakers produced relatively similar speech patterns in both speaking conditions. Therefore, EL speakers could have potentially benefitted from further training and instruction to elicit CS. Additionally, similar SI scores and acoustic data suggest that EL speakers could have benefitted from more guided practice while attempting to produce CS.

Each speaker produced CS after instructions were provided to them during each recording session. Contrarily, rather than a lack of instruction or practice, consideration must be provided for the fact that participant speakers received postlaryngectomy speech rehabilitation using an EL device. This often involves a reduced rate of speech and highlights the importance of over-articulation during speech production with an EL device. Since speakers range from 24 to 300 months postlaryngectomy. Alternatively, if

principles from EL speech rehabilitation were maintained, could EL speakers execute CS instructions if were speaking with a reduced rate and over-articulating.

Although EL speakers were at least two years postlaryngectomy and were deemed proficient users of an EL device, consideration must be provided for the cognitive and effort demands associated with CS. When using an EL device, speakers must consider the maintenance of proper placement of the EL device on the neck, use of a wider mouth opening alongside a slower rate of speech, and coordination of the on/off operation of the device during speech production (Doyle, 1994). While one of the potential benefits of CS is suggested to be improved coordination of various subsystems for speech production (Tjaden et al., 2014), research has indicated that speakers exert greater effort during CS (Perkell, Zandipour, Matthies, & Lane, 2002). Given the results of the present work (e.g., only 50% of the participants exhibited improvements in SI), it is not surprising that Perkell et al. (2002) previously discovered also that ~40% of speakers in their study produced greater changes in articulation with greater effort. The remaining ~60% of speakers, however, only increased vowel durations and/or intensity without increasing effort (Perkell et al., 2002). Further, it is important to note that the EL speakers in the present investigation were asked to further reduce their rate and over-articulate beyond how they were already speaking in HS. Therefore, the additional cognitive and effort demands required to produce CS in association with the basic tasks involved in producing speech with an EL device, could lead to two interpretations of the present findings.

First, the additional cognitive and effort demands were too great for the EL speakers and attempts to further modify speech might have involved modification of only one aspect within the CS instructions (e.g., modifying speech rate as observed through

vowel and word duration increases without modifying articulation). This could be supported further by the increased vowel durations alongside the unchanged first formant frequency data in Experiment 2. The other possibility is that at least 50% of speakers could have abandoned the CS instructions altogether (e.g., did not attempt to further reduce rate or over-articulate), resulting in at least half of EL speakers decreasing their SI when producing CS in Experiment 1.

A limitation of the present series of experiments must consider how the word stimuli posed several challenges for analysis and generalizing results during Experiments 1 and 2. For Experiment 1, listeners were tasked with having to identify single, isolated words in order to determine the overall SI of EL speakers. While the intent of using isolated words was to identify the impact CS in this challenging and decontextualized context, it is difficult to generalize the results in SI scores beyond the present investigation. More specifically, it is difficult to extend findings from investigations of word SI to communication involving connected speech. In addition, further difficulty is met when attempting to compare the present data to previous research. The majority of the research investigating CS has used sentence stimuli to assess the influence of CS on SI and auditory-perceptual measures (Hanson et al., 2004; Krause & Braida, 2002; Picheny et al., 1985; Tjaden et al., 2014). Even in studies that have analyzed the effect of CS on words, the stimulus words were often extracted from recorded sentences (Ferguson & Kewley-Port, 2002; Uchanski, Choi, Braida, & Durlach, 1996). This study, however, is the first to examine the influence on EL speakers, in addition to one of few studies examining the application of CS on words.

Lastly, if sentences were used in the present investigation, it is believed that the SI of EL speakers (and the overall effect of CS) would have been greater. Research by Egan (1948) suggested that there is a distinct relationship between SI (or, ‘articulation scores’) of words and sentences. In their investigation of EL speaker SI using word stimuli, Barney et al. (1959) commented:

...it has been found that a 60 per cent articulation from such isolated words corresponds to a sentence intelligibility of more than 95 per cent, and that even 40% in the word score means that more than 90 per cent of sentences would be understood. (p. 1355).

Therefore, the mean word SI of 53% achieved by the present EL speakers while using CS could correspond to a sentence SI score of more than 90%. Findings from Experiment 1, then, did not potentially illustrate the full impact of CS on EL speakers’ SI given that sentences could have portrayed a very different effect. In addition, the actual structure of stimulus words could have been balanced more carefully for the acoustic analyses of vowels in Experiment 2.

The word stimuli were chosen to ensure equal representation of consonants in WI and WF positions in Experiment 1. From the 18 words used in Experiment 1, four monophthongs and two diphthongs were represented and further analyzed for Experiment 2. These vowels occurred in a variety of phonetic contexts (e.g., voiceless WI consonant and voiceless WF consonant, Voiceless WI consonant and voiced WF consonant, voiced WI consonant and voiced WF consonant). Generally, the overall number of vowel stimuli was unequal in presentation, in addition to the representation in the number of times each vowel appeared in specific phonetic contexts (e.g., more occurrences of /æ/ than any other vowel). Research has shown that phonetic context plays a role in determining the acoustic characteristics of vowels for normal speakers, especially vowel duration

(Raphael, 1972; Umeda, 1975). In addition, previous research investigating the impact of CS on vowel SI provided a /bVd/ context (Ferguson & Kewley-Port, 2002). Ultimately, this type of experimental control enabled researchers to remove the possible effects of phonetic context on the SI and acoustic analyses of vowels. While basic analyses of vowel duration and frequency were consistent across speakers for the present study, the general selection of word stimuli make it difficult for specific comparisons to be made to previous research. Therefore, future investigations must ensure equal representation of vowel stimuli so that more finite conclusions can be drawn regarding the impact of CS on the acoustic characteristics of vowels in EL speech.

Clinical Implications

EL speakers are initially trained to use an EL device by using a set of general principles that include a slower rate of speech, over-articulation, and proper on/off timing during communication (Doyle, 1994, 2005). The ultimate goal, of course, is to produce speech that is as intelligible as possible. An important clinical consideration prior to any pursuits that seek to improve acoustic or auditory-perceptual characteristics of EL speech is two-fold. First, it is important to acknowledge the ramifications of the unique and complex acoustic structure of the EL speech signal and any modifications that are pursued. That is, further modification of speech rate, for example, can pose increased challenges for the listener. For example, research by Goldstein and Rothman (1976) indicated that EL speakers might be penalized if they speak too slowly (e.g., EL speakers judged as ‘poor’ spoke nearly twice the duration as EL speakers judged to be ‘good’) (as cited in Rothman, 1978). Overall, speaking rate was determined to be the biggest predictor for EL speech proficiency. Since CS involves a reduced rate of speech

alongside over-articulation, it is important to determine the slowest rate of speech that EL speakers can produce to improve SI without reductions in listener perceptions. Hanson et al. (2004) demonstrated that even highly effective speech supplementation strategies that improve SI in dysarthria speakers can result in negative listener judgments. This speaks to the inherent difficulty, but necessity, to compare the current investigations of CS to those involving participants with dysarthria.

When compared to dysarthria, EL speech is unique in that it is not, and has never been, perceived to be of a ‘human-like’ quality; rather, it has been deemed ‘noisy’, ‘mechanical’ and ‘robotic’ (Barney et al., 1959; Bennett & Weinberg, 1973; Hillman et al., 1998; Meltzner & Hillman, 2005). As a result, EL speakers’ communicate with an external voicing source that is characterized by reduced SI (Barney et al., 1959; Weiss & Basili, 1985; Weiss et al., 1979), numerous acoustic deficits in the monotony of frequency (Cole, Sridharan, Moody, & Geva, 1997; Meltzner & Hillman 2005), and resultantly unfavourable listener perceptions (Bennett & Weinberg, 1972; Doyle & Eadie, 2005; Williams & Watson, 1985). To integrate findings from previous research and those from the present investigation, the discussion to follow will examine the clinical utility of CS toward improving EL voice and speech.

In the context of the present experiments, it is important to consider the influence of CS on EL speech, and in particular, the impact of CS on SI, the complex acoustic nature of the EL signal, and listener judgments. The present experiments revealed relatively similar SI scores when EL speakers produced words using HS and CS. In addition, while there were significant changes to the durational properties of words and some vowels, there were minimal overall changes to the frequency characteristics of the

EL speech. Due to the continuously voiced nature of EL speech, voicing confusions are a predominant class of errors (Weiss & Basili, 1985; Weiss et al., 1979). Generally, the present findings suggest that CS can lead to a small reduction of ~ 2.2% of voicing errors in WI position and 6.9% voicing errors in the WF position. The application of CS, however, cannot overcome the electronic, continuously voiced source used by EL speakers. Therefore, data indicate that these voicing errors persisted for EL speech in the presence of these types of articulatory modifications. Other concerns that cannot be overcome by CS pertain to neck placement of EL devices and the altered vocal tract in which the EL speech signal resonates following laryngectomy. The following paragraph will describe the general concerns regarding the transmission of the EL speech signal across neck tissue postlaryngectomy, in addition to the impact of an altered vocal tract on acoustic characteristics of EL speech.

Treatment of laryngeal cancer can include laryngectomy combined with radiation and/or chemotherapy. Following laryngectomy and radiation treatment (with or without chemotherapy), EL speakers often have surgical scarring, in addition to fibrotic or lymphedematous neck tissue (Doyle, 1994). Investigations directed toward improving EL speech, then, must consider the difficulties associated with sound transmission across neck tissues into an altered vocal tract postlaryngectomy (Meltzner, Kobler, & Hillman, 2003; Sisty & Weinberg, 1972). Meltzner et al. (2003) commented on the difficulty of transmitting the EL signal across neck tissue, which is often thought to contribute to the frequency deficits observed in EL speech. Furthermore, Sisty and Weinberg (1972) demonstrated that for laryngectomized individuals who use ES speech, the postlaryngectomy vocal tract is reduced in effective length which can then alter the

formant frequency characteristics of ES speech. Together, these anatomical changes can impact the frequency characteristics of EL speech and further contribute to difficulty in overcoming these deficits with CS alone. That is, without manipulating the acoustic signal, Experiments 1 through 3 have demonstrated that minimal changes occur as a direct result of EL speakers using CS in isolation. Given all of the anatomical changes secondary to total laryngectomy, including scarring and fibrosis, consideration must be provided for what is deemed the ‘expected performance norms’ when using an EL device. That is, a general understanding of a speakers’ anatomical limits using an EL must be respected in relation to the amount of side effects that exist postlaryngectomy (Doyle & Eadie, 2005).

Research has provided a general indication of the established levels of communication proficiency for EL speakers (e.g., Rothman, 1978; Rothman & Goldstein, 1976). This includes levels of speaker performance that are not easily overcome by the most sophisticated attempts to modify current EL devices; including, the removal of device noise, enhancement of low-frequency deficits, and modification of intonation patterns (Espy-Wilson et al., 1998; Meltzner & Hillman, 2005). Since Experiment 1 centered on SI, it is important to understand that EL speakers have a mean SI between 50 to 60% with a range of 16 to 90% (Barney et al., 1959; Weiss & Basili, 1985; Weiss et al., 1979, Yeni-Komshian et al., 1975). Generally, numerous factors can contribute to variability of SI, including speaking rate and degree of over-articulation (Doyle, 1994). Given the range of word and vowel durations observed, data from the present study suggest that EL speakers vary in their use of CS, which is in agreement with previous work on CS (Picheny et al., 1985). In addition, indirect assessment of formant frequency

data of vowels suggest that speakers also varied their degree of mouth opening. Together, variability in speech rate and mouth opening could have impacted SI scores, especially since CS has been generally shown to improve SI when speakers properly reduce rate, over-articulate, and increase mouth opening (Ferguson & Kewley-Port, 2002; Lam & Tjaden, 2013; Picheny et al., 1985, 1986).

Experiment 1 demonstrated a drastic range in EL speaker performance in both speaking conditions (e.g., 29- 67% in CS and 30-69% in HS), which highlights another important implication for clinical intervention aimed at improving EL speech. This calls attention to the need for clinical monitoring of individuals' speech production while using an EL device. This could involve the consistent use of CS instruction (i.e., a reduced rate, over-articulation and increased mouth opening) and monitoring by SLPs to assist in preventing poor levels of SI to be reached. Whether monitoring occurs at follow-up SLP and otolaryngology visits with or without scheduled speech rehabilitation sessions, it is the duty of each SLP to ensure that alaryngeal speakers are speaking with the highest level of SI. Findings indicate that brief instruction and limited practice in the current investigations resulted in a 1.3% increase in SI. Further, it is important to note that these speakers had at least two years of experience using an EL device. Therefore, this style of speaking might facilitate larger improvements for individuals undergoing laryngectomy and/or learning to acquire EL speech. Lastly, if SLPs incorporate CS, research has suggested that instructions must include 'a slow rate of speech', 'over-articulation' and 'increased mouth opening' to facilitate the best possible productions from the clinical population (Lam & Tjaden, 2013; Picheny et al., 1985). While these directions were followed in the current investigations, it appears that EL speakers were unable to

significantly benefit from CS instructions. This is possibly due to their experience with previous EL training, or not fully adjusting the productive aspects of their speech according to CS instructions. Further development of a CS criterion for EL speakers is required. These clinical considerations lend themselves to important research applications and directions.

Directions for Future Research

First, it is important for future research investigations to establish a criterion that separates CS from a general reduction in speech rate and over-articulation following EL speech rehabilitation. Several steps are essential in order to establish such a criterion. First, research has indicated that instructions to produce CS must include explicit directions for speakers to over-articulate in addition to reducing speech rate and increasing mouth opening (Lam & Tjaden, 2013; Picheny et al., 1985). Consideration should be given to allow research participants to practice with the instructions for longer periods; specifically, Krause and Braida (2002) indicated that speakers in their study were provided with one hour of practice with CS after a thorough discussion of the technique. Furthermore, control of speaking rates can occur through the use of a metronome (Krause & Braida, 2002). One application of a metronome in therapy could require speakers to produce a given number of stimuli in between metronome 'clicks' at varying rates (Krause & Braida, 2002). This procedure would assist in maintaining the speech rate of speakers throughout therapy (Krause & Braida, 2002). Finally, the criteria for accurate production of CS could be established through direct measurement of mouth opening to ensure that all speakers are able to achieve a relatively similar increase mouth opening. While obtaining a direct measurement is not always clinically reasonable for

each production, basic measurement practices would enable further definition of an acceptable range of mouth-opening during CS.

Once CS criteria are established, detailed analyses of the prosodic differences can assist in identifying how EL speakers' change their speech when producing HS and CS. It is important to highlight that, for example, many of the prosodic changes in CS have been shown to occur as a result of the insertion of pauses at word boundaries and lengthening of speech sounds (Picheny et al., 1986). Therefore, future research should consider the analysis of at least sentence-level assessments. Such assessments can include the influence of changes to intonation, stress, rhythm and juncture on SI and the acoustics and perceptual aspects of EL. Further, a range of expected outcomes measures in SI and frequency data relative to degree of mouth opening can be obtained and compared to future productions. Of course, outcome measures would be unique to laryngectomees with similar characteristics following TL (e.g., treatment characteristics involving neck dissection, radiation, additional surgical procedures, etc.). Ideally, this entire process would permit formal description of CS when compared to HS.

Specific to assessment of SI, EL speakers could be guided to use the established criterion while producing a given set of the stimuli (as chosen by the SLP). The clinician would monitor speaker performance to ensure that the principles of the CS criterion are followed (i.e., reduction of rate according to metronome, over-articulation and mouth opening to a specific measurement). All speakers would be provided with an explicit set of instructions alongside the criterion due to the importance of instruction for producing CS (Lam & Tjaden, 2013). A similar study to Experiment 1 in the present work can then be conducted and SI levels can be determined for EL speakers using the CS criterion.

After SI data are collected, acoustic analyses would permit an understanding of the alterations to the EL speech signal after the CS criterion is used; for example, analysis of vowel formant frequencies between HS and CS. An example template to conduct such a project can be found in Experiment 2. Lastly, auditory-perceptual research can assess the impact of CS using the established criterion on listener judgments. Similar to the methodology in Experiment 3 (Chapter 4), researchers can use LC and ACC to assess the influence of the CS criterion on these perceptual judgments. Overall, the goal of this process is to arrive at a refined assessment of CS (e.g., the criterion), and the effectiveness of this criterion to ensure all speakers produce CS to a similar degree. If, by using the criterion, it is determined that instruction of CS requires speakers to be given a specific tempo (e.g., in syllables per second), would EL speakers be able to improve overall SI (e.g., of words)? In addition, for the EL speaker who has already been formally trained to use an EL device as outlined by Doyle (1994; 2005), questions arise in regard to how they perceive the basic instructions to produce CS? Would they require further, in-depth training? Contrarily, if proficient EL speakers are told to ‘speak as clearly as possible’ by ‘over-articulating’ and ‘slowing down while speaking’, would they continue in their ‘habitual’ manner of speaking due to an assumed level of proficiency with talking in this manner?

In contrast, if EL speakers are trained to use EL speech but do not happen to fully adapt to the initial EL speech rehabilitation protocol, do they possess the ability to successfully adapt to CS instructions at a later period? In other words, would CS benefit the EL speaker any differently than the set of instructions already provided to produce speech in a clear and effective manner with their EL device? While these instructions

include important terminology toward improving SI (e.g., ‘over-articulate’), it appears that these instructions did not lead to an appreciable improvement in SI in the present study (e.g., 1.3% for words, 1.6% for WI consonants, and 0.9% for WF consonants). This might suggest that EL speakers could benefit from longer training sessions in order to properly produce CS (Krause & Braida, 2002). The creation of thorough criteria that facilitates proper utilization of CS might be warranted.

One final direction for future research can be drawn from collective work by Beaudin (2002), Meltzner and Hillman (2005), and Wong (2003). First, Meltzner and Hillman’s (2005) work demonstrated that improving aspects of EL speech involves a combination of several features (e.g., low-frequency enhancement, device noise reduction, *and* frequency variation) rather than single, isolated acoustic features. The present investigations serve as examples whereby the focus on voluntary modification to EL speakers’ productions was geared toward understanding the influence of CS on numerous aspects of EL speech; this includes SI, acoustics, and listener judgments. While this involves a slower rate of speech and over-articulation, the acoustic structure of the EL speech signal is far too complex to be overcome by modification in speech rate and articulatory patterning without altering the EL signal itself. Further support is provided by results from Experiment 2 showing no significant acoustic changes occurred when EL speakers attempted to use CS. Again, while Beaudin (2002) found that listeners might indicate a preference for signals generated from modified devices, Wong’s (2003) research indicates that even the most sophisticated and promising modifications to an EL device are unable to simultaneously improve the SI of EL speakers. Therefore, it is suggested that future investigations seeking to improve communication for EL speakers

consider the simultaneous use of a modified EL speech signal alongside the modification (and monitoring) of the productive aspects of EL speech (e.g., reduced speech rate, over-articulation, and increased mouth opening). Lastly, when attempting to modify articulation, not only is it essential to provide the appropriate instructions (e.g., slow rate, over-articulate, increase mouth opening), but there is a need for a criterion to be established as to what constitutes CS in EL speakers, and that each speaker meets this criterion as measured by a specific tempo and mouth-opening measurement.

The present series of experiments demonstrate that individuals may require greater training, refinement, and monitoring in CS to facilitate significant differences in this speaking condition compared to HS. In addition, careful consideration must be given to the threshold whereby articulatory rate begins to negatively impact the speaker. For example, is there a threshold for speech rate (e.g., reduced syllables per second) whereby listeners perceive CS to be significantly less acceptable and/or less comfortable to listen to when compared to HS? The findings of the present three experiments suggest that the general significance between speaking rates in CS and HS impacted ACC, but not LC. Therefore, it is important to consider that these two perceptual judgments are potentially sensitive enough to detect changes in reduced speaking rate, over-articulation, and/or increased mouth opening while EL speakers use CS.

Conclusions

Results from the present study suggest that, while word durations were significantly longer in CS compared to HS, the group of EL speakers were unable to derive a significant improvement in SI scores and alterations to the frequency components of the EL signal while using CS. Alterations to EL speakers' articulation led

to significant differences in listener judgments related to ACC, but did not impact judgments of LC. Findings are inconsistent with previous research that examined the use of CS involving individuals with variety of speech and hearing impairments. However, this is the first study to report the application of CS in EL speakers, a group of individuals who have received speech rehabilitation involving the use of a slower rate of speech and instruction to over-articulate. Additionally, previous research has indicated that stimuli spoken using CS (e.g., vowels and sentences) are twice as long as stimuli spoken in conversational (or, habitual) speech (Picheny et al., 1985; Ferguson & Kewley-Port, 2002). The present investigation, however, found that EL speakers were unable to increase the duration of words to a similar degree. In addition, Ferguson and Kewley-Port (2002) reported that CS led to an expanded formant frequency vowel space. In the current investigation, the vowel space in EL speakers did not drastically change between speaking conditions. In fact, much overlap was observed in F1 and F2 formant frequencies between HS and CS.

Finally, the present investigation was consistent with previous research indicating that the most effective strategy at facilitating communication may not be the most preferred or acceptable strategy as perceived by listeners (Hanson et al., 2004). Future research efforts should be focused on improving EL speech by addressing the acoustic aspects of the signal (e.g., frequency, intonation, intensity, etc.) alongside the implementation of a CS criterion. Ultimately, criteria to ensure that EL speakers are meeting a minimum performance standard relative to accurately producing CS. If the criteria permits greater improvements in SI, then the criteria will be able to facilitate each EL speaker's highest level of communication proficiency.

References

- Barney, H. L., Haworth, F. E., & Dunn, H. K. (1959). An experimental transistorized artificial larynx. *Bell system technical Journal*, 38(6), 1337-1356.
- Beaudin, P.G. (2002). A paired-comparison evaluation of electronic artificial laryngeal device preference for vowel stimuli. Unpublished master's thesis, Western University, Ontario, Canada.
- Bennett, S., & Weinberg, B. (1973). Acceptability ratings of normal, esophageal, and artificial larynx speech. *Journal of Speech and Hearing Research*, 16, 608-615.
- Berry, W. R. (1978). Attitudes of speech language pathologists and otolaryngologists about artificial larynges. In S.J. Salmon and L. P. Goldstein (Eds). *The artificial larynx handbook*. New York: Grune & Stratton.
- Beukelman, D. R. Fager, S. Ullman, C. Hanson, E., & Logemann, J. (2002). The impact of speech supplementation and clear speech on the intelligibility and speaking rate of people with traumatic brain injury. *Journal of Medical Speech-Language Pathology*, 10, 237-242.
- Bradlow, A. R., Torretta, G. M., & Pisoni, D. B. (1996). Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics. *Speech communication*, 20(3), 255-272.
- Cole, D., Sridharan, S., Moody, M., & Geva, S. (1997). Applications of noise reduction techniques for alaryngeal speech enhancement. Proceedings of the IEEE Telecon 1997 Region 10 Annual Conference, Speech and Image Technologies for Computing and Telecommunications, 2, 491-494.

- Darley, F. L., Aronson, A. E., & Brown, J. R. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech, Language, and Hearing Research*, 12(2), 246-269.
- Doyle, P.C. (1994). Foundations of voice and speech rehabilitation following laryngeal cancer. San Diego: Singular.
- Doyle, P. C. (2005). Clinical procedures for training use of the Electronic artificial larynx. In P.C. Doyle, & R.L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing*. Austin, TX: Pro-Ed.
- Doyle, P.C., & Eadie, T.L. (2005). The perceptual nature of alaryngeal voice and speech. In P.C. Doyle, & R.L. Keith (Eds.), *Contemporary considerations in the treatment and rehabilitation of head and neck cancer: Voice, speech, and swallowing*. Austin, TX: Pro-Ed.
- Duguay, M. J. (1978). Why not both? In S.J. Salmon and L. P. Goldstein (Eds.), *The artificial larynx handbook*. New York: Grune & Stratton.
- Eadie, T. L., & Doyle, P. C. (2004). Auditory-perceptual scaling and quality of life in tracheoesophageal speakers. *The Laryngoscope*, 114(4), 753-759.
- Espy-Wilson, C. Y., Chari, V. R., MacAuslan, J. M., Huang, C. B., & Walsh, M. J. (1998). Enhancement of electrolaryngeal speech by adaptive filtering. *Journal of Speech, Language, and Hearing Research*, 41, 1253–1264.
- Gates, G., Ryan, W., Cooper, J., Lawlis, G., Cantu, E., Hyashi, T.,...Hearne, E. (1982). Current status of laryngectomy rehabilitation: I. Results of therapy. *American Journal of Otolaryngology*, 3, 1-17.

- Goldstein, L. P., & Rothman, H. B. (1976, November). Analysis of speech produced with an artificial larynx. Presented at the American Speech and Hearing Association Convention, Houston, TX.
- Hanson, E., Beukelman, D. R., Fager, S., Ullman, C. (2004). Listener attitudes and speaking rate of people with traumatic brain injury. *Journal of Medical Speech-Language Pathology, 12*(4), 161-166.
- Hillman, R. E., Walsh, M. J., Wolf, G. T., Fisher, S. G., & Hong, W. K. (1998). Functional outcomes following treatment for advanced laryngeal cancer. Part I - Voice preservation in advanced laryngeal cancer. Part II - Laryngectomy rehabilitation: The state of the art in the VA system. Research Speech-Language Pathologists. Department of Veterans Affairs Laryngeal Cancer Study Group. *Annals of Otology, Rhinology and Laryngology Supplement, 172*, 1-27.
- Hoops, H.R., & Noll, J.D. (1969). Relationship of selected acoustic variables to judgements of esophageal speech. *Journal of Communication Disorders, 2*, 3-13.
- Kessinger R. H., & Blumstein, S. E. (1997). Effects of speaking rate on voice-onset time in Thai, French, and English. *Journal of Phonetics, 25*, 143-168.
- Krause, J. C., & Braida, L. D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *Journal of the Acoustical Society of America, 112*(5), 2165-2172.
- Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *The Journal of the Acoustical Society of America, 115*(1), 362-378.

- Kreiman, J., Gerratt, B. R., Kempster, G. B., Erman, A., & Berke, G. S. (1993). Perceptual evaluation of voice quality: Review, tutorial, and a framework for future research. *Journal of Speech, Language, and Hearing Research*, 36(1), 21-40.
- Lam, J., & Tjaden, K. (2013). Intelligibility of clear speech: Effect of instruction. *Journal of Speech, Language, and Hearing Research*, 56(5), 1429-1440.
- Lauder, E. (1968). The laryngectomy and the artificial larynx. *Journal of Speech and Hearing Disorders*, 33(2), 147-157.
- Lehiste, I. (1976). Suprasegmental features of speech. In N. Lass. (Ed.), *Contemporary issues in experimental phonetics*. New York, NY: Academic Press.
- Mayo, C., Aubanel, V., & Cooke, M. (2012). Effect of prosodic changes on speech intelligibility. *Interspeech*, 1708-1711.
- Meltzner, G. S., & Hillman, R.E. (2005). Impact of aberrant acoustic properties on the perception of sound quality in electrolarynx speech. *Journal of Speech, Language, and Hearing Research*, 48(4), 766-769.
- Meltzner, G. S., Kobler, J. B., & Hillman, R. E. (2003). Measuring the neck frequency response function of laryngectomy patients: Implications for the design of electrolarynx devices. *Journal of the Acoustical Society of America*, 114(2), 1035-1047.
- Miller, J. L., Green, K. P., & Reeves, A. (1986). Speaking rate and segments: A look at the relation between speech production and speech perception for the voicing contrast. *Phonetica*, 43(1-3), 106-115.

- Miller, J. L., & Volaitis, L. E. (1989). Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics*, *46*(6), 505-512.
- Nagle, K. F., Eadie, T. L., Wright, D. R., & Sumida, Y. A. (2012). Effect of fundamental frequency on judgements of electrolaryngeal speech. *American Journal of Speech-Language Pathology*, *21*(2), 154-166.
- O'Brian, S., Packman, A., Onslow, M., Cream, A., O'Brian, N., & Bastock, K. (2003). Is listener comfort a viable construct in stuttering research? *Journal of Speech, Language, and Hearing Research*, *46*, 503-509.
- Payton, K.L., Uchanski, R.M., & Braida, L.D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America*, *95*, 1581–1592.
- Perkell, J. S., Zandipour, M., Matthies, M. L., & Lane, H. (2002). Economy of effort in different speaking conditions. I. A preliminary study of intersubject differences and modeling issues. *The Journal of the Acoustical Society of America*, *112*(4), 1627-1641.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing. I. Intelligibility differences between clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, *28*, 96–103.
- Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard of hearing. II. Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language, and Hearing Research*, *29*, 434–446.
- Qi, Y. Y., & Weinberg, B. (1991) Low-frequency energy deficit in electrolaryngeal speech. *Journal of Speech, Language, and Hearing Research*, *34*(6), 1250-1256.

- Raphael, L. J. (1972). Preceding vowel duration as a cue to the perception of the voicing characteristic of word-final consonants in American English. *Journal of the Acoustical Society of America*, 51(4B), 1296-1303.
- Robbins, J., Fisher, H.B., Blom, E.C., & Singer, M.I. (1984). A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production. *Journal of Speech and Hearing Disorders*, 49(2), 202–210.
- Rothman, H. B. (1978). *Analyzing artificial electronic larynx speech*. In S. J. Salmon and L. P. Goldstein. (Eds.), *The artificial larynx handbook*. New York, NY: Grune. pp. 87-111.
- Saikachi, Y., Stevens, K. N., & Hillman, R. E. (2009). Development and perceptual evaluation of amplitude-based F0 control in electrolarynx speech. *Journal of Speech, Language, and Hearing Research*, 52(5), 1360-1369.
- Sisty, N. L., & Weinberg, B. (1972). Formant frequency characteristics of esophageal speech. *Journal of Speech, Language, and Hearing Research*, 15(2), 439-448.
- Smiljanić, R., & Bradlow, A. R. (2009). Speaking and hearing clearly: Talker and listener factors in speaking style changes. *Language and linguistics compass*, 3(1), 236-264.
- Snidecor, J. C., & Curry, E.T. (1959). Temporal and pitch aspects of superior esophageal speech. *Ann Otol Rhinol Laryngol*, 68, 623-629.
- Stevens, K. N., & House, A. S. (1955). Development of a quantitative description of vowel articulation. *The Journal of the Acoustical Society of America*, 27(3), 484-493.

- Tardy-Mitzell, S., Andrews, M. L., & Bowman, S. A. (1985). Acceptability and intelligibility of tracheoesophageal speech. *Archives of Otolaryngology*, *111*(4), 213-215.
- Theodore, R. M., Miller, J. L., & DeSteno, D. (2009). Individual talker differences in voice onset time: Contextual influences. *Journal of the Acoustical Society of America*, *125*(6), 3974-3982.
- Tjaden, K., Sussman, J.E., & Wilding, G. E. (2014). Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in parkinson's disease and multiple sclerosis. *Journal of Speech, Language, and Hearing Research*, *57*, 779-792.
- Turner, G. S., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, *38*(5), 1001-1013.
- Uchanski, R. M., Choi, S. S., Braida, L. D., Reed, C. M., & Durlach, N. I. (1996). Speaking clearly for the hard of hearing IV: Further studies of the role of speaking rate. *Journal of Speech, Language, and Hearing Research*, *39*(3), 494-509.
- Umeda, N. (1975). Vowel duration in American English. *Journal of the Acoustical Society of America*, *58*(2), 434-445.
- Ward, E. C., & van As-Brooks, C. J. (Eds.). (2014). *Head and neck cancer: treatment, rehabilitation, and outcomes* (2nd ed.). San Diego, CA: Plural Publishing Inc.
- Weiss, M S., & Basili, A. M. (1985). Electrolaryngeal speech produced by laryngectomized subjects: Perceptual characteristics. *Journal of Speech and Hearing Research*, *28*, 294-300.

Weiss, M. S., Yeni-Komshian, G. H., & Heinz, J. M. (1979). Acoustical and perceptual characteristics of speech produced with an electronic artificial larynx. *Journal of the Acoustical Society of America*, 65(5), 1298-1308.

Wong, J.H.C. (2003). *An experimental evaluation of speech intelligibility using a prototype electrolaryngeal device*. Unpublished master's thesis, Western University, Ontario, Canada.

Yeni-Komshian, G. H., Weiss, M. S., & Heinz, J. M. (1975). Intelligibility of speech produced with an electronic artificial larynx. *Journal of the Acoustical Society of America*, 58(S1), S112-S112.

Appendix A

Stimulus Word List (Weiss & Basili, 1985)

1. leave	23. feel	45. <i>badge</i>
2. cane	24. witch	46. sheath
3. jog	25. near	47. gab
4. cheap	26. <i>dab</i>	48. <i>gain</i>
5. <i>catch</i>	27. sag	49. thigh
6. meal	28. hun	50. path
7. thy	29. bad	51. game
8. tab	30. zack	52. edge
9. five	31. ease	53. chad
10. <i>mass</i>	32. rich	54. <i>vet</i>
11. veal	33. <i>teeth</i>	55. sheathe
12. rice	34. bat	56. <i>chief</i>
13. <i>pad</i>	35. deer	57. <i>these</i>
14. wedge	36. hung	58. <i>fish</i>
15. teethe	37. leaf	59. zing
16. hail	38. <i>jeep</i>	60. jaw
17. came	39. <i>shave</i>	61. <i>theme</i>
18. dope	40. <i>zag</i>	62. gnash
19. <i>sack</i>	41. seek	63. thou
20. ice	42. veer	64. <i>know</i>
21. pat	43. thing	65. <i>loathe</i>
29. mash	44. rise	66. way

**Italicized words indicate words used in the current investigation*

Appendix B



Research Ethics

**Western University Health Science Research Ethics Board
HSREB Delegated Initial Approval Notice**

Principal Investigator: Dr. Philip Doyle
Department & Institution: Health Sciences/Communication Sciences & Disorders, Western University

HSREB File Number: 105382
Study Title: The application of clear speech in alaryngeal speakers.
Sponsor:

HSREB Initial Approval Date: July 30, 2014
HSREB Expiry Date: January 31, 2016

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Instruments	VRQOL Questionnaire	2014/05/20
Data Collection Form/Case Report Form	Demographic Collection	2014/06/16
Recruitment Items	Revised Recruitment Poster	2014/06/16
Western University Protocol	Revised Western Protocol (pdf)	2014/07/09
Letter of Information & Consent	Revised LOI/Consent (pdf)	2014/07/03

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the above named study, as of the HSREB Initial Approval Date noted above.

HSREB approval for this study remains valid until the HSREB Expiry Date noted above, conditional to timely submission and acceptance of HSREB Continuing Ethics Review. If an Updated Approval Notice is required prior to the HSREB Expiry Date, the Principal Investigator is responsible for completing and submitting an HSREB Updated Approval Form in a timely fashion.

The Western University HSREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use Guideline for Good Clinical Practice Practices (ICH E6 R1), the Ontario Personal Health Information Protection Act (PHIPA, 2004), Part 4 of the Natural Health Product Regulations, Health Canada Medical Device Regulations and Part C, Division 5, of the Food and Drug Regulations of Health Canada.

Members of the HSREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Ethics Officer, on behalf of Dr. M. Kremenchutzky, HSREB Vice Chair

Ethics Officer to Contact for Further Information

Erika Basile	Grace Kelly	Miss Mikhail	<input checked="" type="checkbox"/> Yuki Tran
--------------	-------------	--------------	---

This is an official document. Please retain the original in your files.

Appendix C

Letter of Information and Consent Form

Study Title: *The application of clear speech in alaryngeal speakers.*

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

Co-Investigators: Steven R. Cox, PhD(c)

Introduction

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

You are being invited to take part in this study because you use a method of alaryngeal speech as a result of your total laryngectomy.

Purpose of Study

The purpose of this study is to collect voice samples and voice-related quality of life data from individuals who use an alaryngeal method of voice production. Specifically, the purpose of voice sample collection is to investigate whether or not providing guided instructions to a speaker can make their alaryngeal speech as understandable as possible, a process termed 'clear speech'. In doing so, speakers will be requested to provide samples in their typical manner, and then in a clear speech mode during the production of sounds, words, and/or reading passages. It is anticipated that attempts to create clear speech will facilitate communication exchanges between alaryngeal speakers and their communication partners. Additionally, your data will be used to explore how one's voice-related quality of life is impacted as a result of a voice-disorder or use of an alaryngeal method of speech.

Inclusion Criteria

If you are over the age of 18 years old and can read, write, and speak English, you can choose to participate in this study.

Exclusion Criteria

If you are unable to read, write, and speak English, you should not participate in this study.

Description of the Research

This study will require you speak into a microphone so that a variety of voice/speech samples can be recorded. This will involve the recording of several sustained vowels such as "ah", "ee", and "ooh", repeating some short sentences, and the reading aloud of a short paragraph with guided instructions. The recording will require approximately 20-30 minutes and will be done in a formal recording suite or quiet room within a private setting. As well, you will be asked to complete two written questionnaires, 1) a simple document that gathers demographic information from you (e.g., age, time since laryngectomy, etc.) and 2) the Voice-Related Quality of Life Questionnaire which is a 10-item questionnaire that seeks information regarding problems you may experience as a result of your postlaryngectomy voice/speech method.

Participation in this study will require keeping your voice samples and questionnaire data in a secure database for up to ten (10) years for the purposes of this research study.

Risks & Harms

Because of the nature of these tasks, there are no known or anticipated physical, psychological, or emotional risks or discomforts associated with completing this study. However, if you do experience any problems or discomfort, you can discontinue the task at any time.

Benefits

You may not directly benefit from participating in this study but information gathered may provide benefits to others in the laryngectomy community relative to their experiences using alaryngeal methods of postlaryngectomy communication.

Voluntary Participation

Participation in this study is voluntary. You may refuse to participate, refuse to answer any questions, refuse to complete a voice task, or withdraw your study data at any time, even in the future. You will not be compensated for your participation in this research.

Refusal to Participate & Discontinuing Participation

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

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

Confidentiality

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

All of your personal data will be stored electronically in a password protected and encrypted file and as a hard copy in a locked filing cabinet at a locked laboratory at Western University. This locked file is only accessible to the study investigators. Also, a unique identifier will be used instead of your name on all study materials and instruments to protect your confidentiality. If the results of the study are published, your name will not be used and information that discloses your identity will not be released or published. Each participant's full name will be collected and retained to allow us to contact them to invite them to participate in future research. Further, because opportunities to collect additional voice and VRQOL data often occur over time (e.g., future attendance at future national meetings/conferences, etc.), it is important that we are able to reference individuals by name in the database so that additional data can be attributed to the same individual, and not entered as new participant.

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

Contact Information

If you have any questions about your rights as a research participant, the conduct of the study, or the status or maintenance of our database you may contact Steven Cox, Co-Investigator via email, or Dr. Philip Doyle, Principal Investigator, by phone or email.

If you would like to receive a copy of any potential study results, please contact Steven Cox or Dr. Doyle.

If you wish, you may also contact Dr. David Hill, Scientific Director, Lawson Health Research Institute if you have any questions about this research relative to LHSC, or The Office of Research Ethics if you have any other questions about this research.

Waiver of Rights

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

This letter and the consent statement are yours to keep.

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

Consent Statement – Investigators’ Copy

Project Title: *The application of clear speech in alaryngeal speakers.*

Study Investigators:

Philip C. Doyle, Ph.D.

Steven R. Cox, Ph.D(c)

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

Do you agree to be contacted for future research? Yes No

 Participant’s Signature, or
 Legally Authorized Representative

 Investigator’s Signature

 Participant’s Name

 Investigator’s Name

 Date

 Date

Appendix D

Demographic Information Questionnaire

Voice Production and Perception Laboratory Demographic Information Questionnaire

CODE					

Questions about your treatment:

- Neck dissection:** Y | N If yes, which side: Left | Right | Both
- Radiation:** Y | N If yes, pre or post surgery: Pre | Post | Both
- Chemotherapy:** Y | N If yes, pre or post surgery: Pre | Post | Both

Questions about your voice:

- Primary Speech Mode:** Tracheoesophageal (TE) | Esophageal (ES) | Electrolarynx (EL)
- If TE, primary (at time of surgery) or secondary (after surgery): Primary | Secondary
- If TE, which prosthesis: Blom-Singer - InHealth | Atos - Provox | Other
- If "other", please specify: _____
- If TE, size _____ indwelling device: Y | N

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

Very poor | Poor | Fair | Good | Excellent

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

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

Other treatment or health related notes: _____

Appendix E



Research Ethics

Western University Health Science Research Ethics Board HSREB Delegated Initial Approval Notice

Principal Investigator: Dr. Philip Doyle
Department & Institution: Health Sciences/Communication Sciences & Disorders, Western University

HSREB File Number: 105884
Study Title: Impact of Clear Speech on Listener Perception of Electrolaryngeal Speech
Sponsor:

HSREB Initial Approval Date: November 25, 2014
HSREB Expiry Date: August 31, 2015

Documents Approved and/or Received for Information:

Document Name	Comments	Version Date
Recruitment Items	Class Announcement	2014/10/31
Data Collection Form/Case Report Form	Experiment 3 - Listener Comfort Data Collection Sheet	2014/09/26
Data Collection Form/Case Report Form	Experiment 1 Data Collection Sheet (sample)	2014/09/26
Data Collection Form/Case Report Form	Experiment 2 Data Collection Sheet (sample)	2014/09/26
Data Collection Form/Case Report Form	Experiment 3 - Speech Acceptability Data Collection Sheet	2014/09/26
Recruitment Items	Recruitment Poster	2014/10/31
Recruitment Items	SCox Recruitment Script	2014/10/31
Recruitment Items	PDoyle Email/Telephone Script	2014/10/31
Western University Protocol	Revised Western Protocol	2014/10/31
Letter of Information & Consent	LOI Phase 1	2014/10/31
Letter of Information & Consent	LOI Phase 2	2014/10/31
Letter of Information & Consent	LOI Phase 3	2014/10/31

The Western University Health Science Research Ethics Board (HSREB) has reviewed and approved the above named study, as of the HSREB Initial Approval Date noted above.

HSREB approval for this study remains valid until the HSREB Expiry Date noted above, conditional to timely submission and acceptance of HSREB Continuing Ethics Review. If an Updated Approval Notice is required prior to the HSREB Expiry Date, the Principal Investigator is responsible for completing and submitting an HSREB Updated Approval Form in a timely fashion.

The Western University HSREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use Guideline for Good Clinical Practice Practices (ICH E6 R1), the Ontario Personal Health Information Protection Act (PHIPA, 2004), Part 4 of the Natural Health Product Regulations, Health Canada Medical Device Regulations and Part C, Division 5, of the Food and Drug Regulations of Health Canada.

Members of the HSREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB.

The HSREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 0000940.

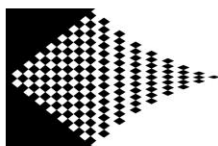
Ethics Officer, on behalf of Dr. Joseph Gilbert, HSREB Chair

Ethics Officer to Contact for Further Information

Erika Beale	Grace Kelly	Hee Mahall	Yvonne Tran
-------------	-------------	------------	-------------

This is an official document. Please retain the original in your files.

Appendix F



**Voice
Production &
Perception
Laboratory**



**VOICE PRODUCTION AND PERCEPTION LABORATORY
REHABILITATION SCIENCES
WESTERN UNIVERSITY**

**Letter of Information
*Participant Listeners***

Project Title: “The Impact of Clear Speech on Listener Perception of Electrolaryngeal Speech”

Investigators:

Principal Investigator: Philip C. Doyle, Ph.D., *Voice Production and Perception Laboratory, Rehabilitation Sciences and Department of Otolaryngology- Head & Neck Surgery, Western University*

Co-Investigator: Steven R. Cox, PhD(c), *Voice Production and Perception Laboratory, Rehabilitation Sciences, Western University*

Introduction

You are invited to participate in this research because you have responded to an announcement or poster on campus. This letter contains information to help you decide whether or not to participate in this research study. It is important for you to understand why the study is being conducted and what it involves. Please read this letter carefully and feel free to ask questions if anything that is presented is unclear or if there is something you do not understand.

Purpose of the Study:

This study will examine auditory-perceptual judgments of voice and speech characteristics of individuals who use electrolaryngeal (EL) speech. The EL is a method of restoring speech to a person who has been treated for cancer and who has lost their voice box. Auditory-perceptual assessment involves having listeners evaluate voice samples by orthographically identifying words. We are interested in how normal hearing listeners perceive speech samples as part of this study. Hence, because you are a normal hearing listener without formal training related to voice or voice disorders, we are seeking to recruit you as a naïve listener for this study.

Benefits to Society:

The findings from this project may permit Speech-Language Pathologists and Otolaryngologists to provide evidence-based information to potential users of the artificial electrolarynx for those who lose their voice box (larynx) due to cancer. Thus, any potential benefit relates only to those who have undergone laryngectomy.

Activities You Will Participate In:

You will be required to attend two, 90 minute listening sessions in the Voice Production and Perception Laboratory at [REDACTED] Western University. If you agree to participate, you will be asked to listen to and orthographically transcribe words presented through headphones. After completing the first listening session, a second session will be scheduled within 7 days.

Inclusion Criteria

Participants will be of good general health with normal hearing at the time of the study. All participants will be 18 years of age or older, and must be able to read, write, and understand written and spoken English.

Exclusion Criteria:

If you have had prior exposure to or training in voice disorders (formal coursework or clinical experience), previous experience with auditory-perceptual research, or a personal history of any speech, voice, language, or hearing difficulties, you will not be able to be a participant in this study. Also, if you have or have had an upper respiratory infection within the past week that may have influenced your hearing due to congestion, you will not be able to participate.

Voluntary Participation:

Participation in this study is voluntary. You may refuse to participate, refuse to answer any question(s), or withdraw from the study at any time without penalty to you or your academic standing. You can also choose to withdraw any data that you provide to the investigators in the event you decide to withdraw from the study.

Any Possible Risks or Discomforts:

There are no known risks or discomforts associated with participation in this research study.

Any Possible Benefits:

Due to the nature of this study, you will not directly benefit from the data obtained and you will not be compensated for your participation in this research.

Confidentiality:

All data obtained will remain confidential; specifically, all paper documentation used in this study will be stored in a locked cabinet within the Voice Production and Perception Laboratory and electronic files will be stored on a USB key encrypted with TrueCrypt. All study data will be kept for a maximum of 10 years. After which time, paper documents will be shredded in the appropriate area within the Health and Rehabilitation Sciences department. If the results of this study are published, your name will not be used and no information that discloses your identity will be released or published. Representatives of Western University's Health Sciences Research Ethics Board may contact you or require access to your study-related records to monitor the conduct of the research.

If you have any questions about the conduct of this study or your rights as a research subject you may contact the Director of the Office of Research Ethics at [REDACTED], or email at [REDACTED]. Should you have additional questions about the study, you can contact Dr. Philip Doyle at [REDACTED].

Waiver of Rights

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

REB#105884

This letter is yours to keep for future reference.



**Voice
Production &
Perception
Laboratory**



**VOICE PRODUCTION AND PERCEPTION LABORATORY
REHABILITATION SCIENCES
WESTERN UNIVERSITY**

Consent
Participant Listener

Project Title: “The Impact of Clear Speech on Listener Perception of Electrolaryngeal Speech”

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 _____

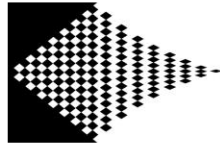
Participant’s Signature _____ Date: _____

Person Obtaining Informed Consent

Printed Name _____

Signature _____ Date: _____

Appendix G



**Voice
Production &
Perception
Laboratory**



**VOICE PRODUCTION AND PERCEPTION LABORATORY
REHABILITATION SCIENCES
WESTERN UNIVERSITY**

**Letter of Information
*Participant Listeners***

Project Title: “The Impact of Clear Speech on Listener Perception of Electrolaryngeal Speech”

Investigators:

Principal Investigator: Philip C. Doyle, Ph.D., *Voice Production and Perception Laboratory, Rehabilitation Sciences and Department of Otolaryngology- Head & Neck Surgery, Western University*

Co-Investigator: Steven R. Cox, PhD(c), *Voice Production and Perception Laboratory, Rehabilitation Sciences, Western University*

Introduction

You are invited to participate in this research because you have responded to an announcement or poster on campus. This letter contains information to help you decide whether or not to participate in this research study. It is important for you to understand why the study is being conducted and what it involves. Please read this letter carefully and feel free to ask questions if anything that is presented is unclear or if there is something you do not understand.

Purpose of the Study:

This study will examine auditory-perceptual judgments of voice and speech characteristics of individuals who use electrolaryngeal (EL) speech. The EL is a method of restoring speech to a person who has been treated for cancer and who has lost their voice box. Auditory-perceptual assessment involves having listeners evaluate voice samples by making judgments of speech samples using a scale on which you mark your judgments. We are interested in how normal hearing listeners perceive speech samples as part of this study. Hence, because you are a normal hearing listener without formal training related to voice or voice disorders, we are seeking to recruit you as a naïve listener for this study.

Benefits to Society:

The findings from this project may permit Speech-Language Pathologists and Otolaryngologists to provide evidence-based information to potential users of the artificial electrolarynx for those who lose their voice box (larynx) due to cancer. Thus, any potential benefit relates only to those who have undergone laryngectomy.

Activities You Will Participate In:

You will be required to attend two, 30 minute listening sessions in the Voice Production and Perception Laboratory at ██████████ Western University. If you agree to participate, you will be asked to make judgments on the samples for a dimension called “speech acceptability” and “listener comfort”. A definition of speech acceptability and listener comfort will be provided to you before beginning the experiment. After completing the first listening session, a second session will be scheduled within 7 days. All listening sessions will be completed while wearing headphones in a quiet listening environment.

Inclusion Criteria

Participants will be of good general health with normal hearing at the time of the study. All participants will be 18 years of age or older, and must be able to read, write, and understand written and spoken English.

Exclusion Criteria:

If you have had prior exposure to or training in voice disorders (formal coursework or clinical experience), previous experience with auditory-perceptual research, or a personal history of any speech, voice, language, or hearing difficulties, you will not be able to be a participant in this study. Also, if you have or have had an upper respiratory infection within the past week that may have influenced your hearing due to congestion, you will not be able to participate.

Voluntary Participation:

Participation in this study is voluntary. You may refuse to participate, refuse to answer any question(s), or withdraw from the study at any time without penalty to you or your academic standing. You can also choose to withdraw any data that you provide to the investigators in the event you decide to withdraw from the study.

Any Possible Risks or Discomforts:

There are no known risks or discomforts associated with participation in this research study.

Any Possible Benefits:

Due to the nature of this study, you will not directly benefit from the data obtained and you will not be compensated for your participation in this research.

Confidentiality:

All data obtained will remain confidential; specifically, all paper documentation used in this study will be stored in a locked cabinet within the Voice Production and Perception Laboratory and electronic files will be stored on a USB key encrypted with TrueCrypt. All study data will be kept for a maximum of 10 years. After which time, paper documents will be shredded in the appropriate area within the Health and Rehabilitation Sciences department. If the results of this study are published, your name will not be used and no information that discloses your identity will be released or published. Representatives of Western University’s Health Sciences Research Ethics Board may contact you or require access to your study-related records to monitor the conduct of the research.

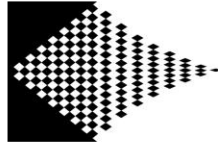
If you have any questions about the conduct of this study or your rights as a research subject you may contact the Director of the Office of Research Ethics at [REDACTED], or email at [REDACTED]. Should you have additional questions about the study, you can contact Dr. Philip Doyle at [REDACTED].

Waiver of Rights

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

REB#105884

This letter is yours to keep for future reference.



**Voice
Production &
Perception
Laboratory**



**VOICE PRODUCTION AND PERCEPTION LABORATORY
REHABILITATION SCIENCES
WESTERN UNIVERSITY**

Consent
Participant Listener

Project Title: “The Impact of Clear Speech on Listener Perception of Electrolaryngeal Speech”

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: _____

CURRICULUM VITAE

STEVEN RANDALL COX

Assistant Professor, Department of Communication Sciences and Disorders

Director, Voice and Quality of Life Laboratory

Ruth S. Ammon School of Education

Adelphi University

Garden City, New York

EDUCATION

Western University, London, Ontario

PhD, Rehabilitation Sciences, 2011- 2016

Central Michigan University, Mount Pleasant, Michigan

MA, Speech-Language Pathology, 2007-2009

Brock University, St. Catharines, Ontario

BA (Hons) with First-Class Standing, Speech and Language Sciences, 2003-2007

POSITIONS HELD

Adelphi University, Garden City, New York

Assistant Professor, Department of Communication Sciences and Disorders

2015 to Present

Adelphi University, Garden City, New York

Director, Voice and Quality of Life Laboratory

2015 to Present

Western University, London, Ontario

Research Associate, Voice Production and Perception Laboratory

2011 – 2016

Western University, London, Ontario

Graduate Teaching Assistant, Resonance Disorders

2012 - 2015

Western University, London, Ontario

Graduate Teaching Assistant, Voice Disorders

2014 - 2015

Western University, London, Ontario

Graduate Teaching Assistant, Speech Science

2014

Western University, London, Ontario

Work Study Student, Department of Communication Sciences and Disorders

2012 – 2013

Brock University, St. Catharines, Ontario

Research Assistant, Department of Education

2011 – 2013

Western University, London, Ontario

Graduate Teaching Assistant, Developmental Language Disorders II

2012

Western University, London, Ontario

Graduate Teaching Assistant, Fluency

2012

Fairview Care Centre at Bethlehem Pike, Philadelphia, Pennsylvania

Clinical Fellow-Speech-Language Pathology, Rehabilitation Department

2010

Roswell Park Cancer Institute, Buffalo, New York

Graduate SLP Intern, Head & Neck and Dental Center

2009

TEACHING EXPERIENCE

GRADUATE COURSEWORK

2. Assistant Professor/Course Coordinator, SPH 630: Voice Disorders, Adelphi

University

Garden City, NY, 2015 to Present

1. Assistant Professor, SPH 600: Speech and Hearing Science, Adelphi University

Garden City, NY, 2015

UNDERGRADUATE COURSEWORK

2. Assistant Professor, SPE 220: Introduction to Speech Science, Adelphi University

Garden City, NY, Present

1. *Instructor*, LING 4P31: Augmentative and Alternative Communication, Brock University,

St. Catharines, ON, 2011

INVITED LECTURES

1. **Cox, S. R.**, & Doyle, P.C. (2013, March). *Caring for the human voice*. Faculty of Music, Western University, London, ON.

MANUSCRIPTS

3. Eadie, T. L., Otero, D., **Cox, S. R.**, Johnson, J., Baylor, C., Yorkston, K., & Doyle, P. C. (2016). The relationship between communicative participation and post-laryngectomy speech outcomes. *Head and Neck*.
2. **Cox, S. R.**, Theurer, J. A., Spaulding, S. J., & Doyle, P. C. (2015). The multidimensional impact of total laryngectomy on women. *Journal of communication disorders*, 56, 59-75.
1. **Cox, S. R.**, & Doyle, P. C. (2014). The influence of electrolarynx use on postlaryngectomy voice-related quality of life. *Otolaryngology - Head & Neck Surgery*, 150(6), 1005-1009.

OTHER PROFESSIONAL CONTRIBUTIONS

3. Bartlett D, Skarakis-Doyle E and members of the Rehabilitation Sciences Journal Club, Health and Rehabilitation Sciences Program at Western (**Cox S.R.**, Davis E, Gregory M, Hope A, Izaryk K, Jeevanantham D, Kogutec D, Lee J, Lutz S, Ready E, and Doyle P). (2014). Response to the World Health Organization's request for comments on the document: How to Use the ICF: A Practical Manual for using the International Classification of Functioning, Disability and Health.

2. **Cox, S.R.**, & Doyle, P.C. (2014). *Why not an electrolarynx?* Voice Prints: WebWhispers.

<http://www.webwhispers.org/news/documents/WhisperontheWeb-May2014.pdf>

1. Doyle, P. C., Caty, M. E., & **Cox, S. R.** (2012). *Use of the artificial electronic larynx in postlaryngectomy speech rehabilitation*. In *Educational and support kit for the laryngectomee and their family*. Santa Clara, CA: The Foundation for Voice Restoration.

PRESENTATIONS AT PROFESSIONAL MEETINGS

19. **Cox, S. R.**, & Doyle, P. C. (2015, November). *Influence of clear speech on the word intelligibility of electrolaryngeal speakers*. Acoustical Society of America, Jacksonville, FL.
18. **Cox, S. R.**, Failla, S., & Doyle, P. C. (2015, November). *The impact of clear speech on listener judgements of electrolaryngeal speech*. American Speech-Language-Hearing Association Annual Convention, Denver, CO.
17. Doyle, P. C., Davis, E., **Cox, S. R.**, Sui, T., Day, A., Bornbaum, C., & Dzioba, A. (2015, November). *Using the Voice-Related Quality of Life questionnaire to examine voice disability following total laryngectomy*. American Speech-Language-Hearing Association Annual Convention, Denver, CO.
16. Reitzel, K., **Cox, S. R.**, Failla, S., & Doyle, P. C. (2015, November). *Temporal modifications of electrolaryngeal speech during a clear speech task*. American Speech-Language-Hearing Association Annual Convention, Denver, CO.

15. Smith, N., **Cox, S. R.**, Day, A., Failla, S., & Doyle, P. C. (2015, November). *Situationally-bound judgements of listener comfort for postlaryngectomy voice & speech*. American Speech-Language-Hearing Association Annual Convention, Denver, CO.
14. Eadie, T. L., Sawin, D., **Cox, S.**, Baylor, C., Yorkston, K., & Doyle, P. C. (April, 2015). *The relationship between post-laryngectomy speech outcomes and communicative participation*. Podium presentation at the American Head and Neck Society (AHNS) Translational Research Meeting, Boston, MA.
13. **Cox, S. R.**, Jeremic, G., Fung, K., & Doyle, P. C. (2014, November). *Patient-reported Mandibular function secondary to radiation treatment for head & neck cancer*. American Speech-Language-Hearing Association Annual Convention, Orlando, FL.
12. Wilkinson, J., Doyle, P.C, **Cox, S. R.**, Failla, S., & Leblanc, C. (2014,November). Determining the presence of a gender bias in auditory-perceptual evaluations of tracheoesophageal speakers. American Speech-Language-Hearing Association Annual Convention, Orlando, FL.
11. 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. American Speech-Language-Hearing Association Annual Convention, Orlando, FL.

10. Doyle, P.C., **Cox, S.R.**, Khalili, S., Fung, K., & Theurer, J. (2014, July). The influence of flap reconstruction on auditory-perceptual evaluation of voice acceptability for tracheoesophageal speakers. 2014 Annual Meeting of the American Head and Neck Society, New York City, NY.
9. Doyle, P.C., Nash, M. M., Scott, G. M., **Cox, S. R.**, Davis, E., Izaryk, K., Bornbaum, C. C., Day, A. M. B., Dzioba, A., & Skarakis-Doyle, E. (2014, May). Aligning the Voice-Related Quality of Life Measure with the ICF: An Examination of Voice Disability Secondary to Laryngeal Cancer. Pacific Rim International Conference on Disability and Diversity, Honolulu, HI.
8. Fung, K, Jeremic, G., **Cox. S. R.**, & Doyle, P. C. (2014, January). *Mandibular metrics, perceived disability and quality of life associated with trismus*. 2014 Triological Society Combined Sections Meeting, Miami Beach, FL.
7. **Cox, S.R.**, Jeremic, G., Fung, K., & Doyle, P. C. (2013, November). *Clinical evaluation of trismus in head & neck cancer: Implications for the Speech-Language Pathologist*. Presented at the American Speech-Language-Hearing Association Annual Convention, Chicago, IL.
6. **Cox, S.R.**, Jeremic, G., Fung, K., & Doyle, P. C. (2013, November). *Clinical evaluation of trismus in head & neck cancer: Implications for the Speech-Language Pathologist*. Presented at the American Speech-Language-Hearing Association Annual Convention, Chicago, IL.

5. Scott, G., **Cox, S.R.**, Nash, M., & Doyle, P.C. (2013, November). *Auditory-perceptual evaluation of speech acceptability in tracheoesophageal speakers*. Presented at the American Speech-Language-Hearing Association Annual Convention, Chicago, IL.
4. Doyle, P. C., Scott, G., & **Cox, S. R.** (2013, November). *Assessment of listener comfort in esophageal, tracheoesophageal, & electrolaryngeal speakers*. Presented at the American Speech-Language-Hearing Association Annual Convention, Chicago, IL.
3. **Cox, S. R.**, Day, A.M.B., Bornbaum, C., & Doyle, P. C. (2013, April). *The impact of the electrolarynx (EL) on Voice-Related Quality of Life (V-RQOL)*. Presented at the 9th Annual Communication Sciences and Disorders Research Day, Western University, London, ON.
2. **Cox, S. R.**, Day, A.M.B., Bornbaum, C., & Doyle, P. C. (2012, November). *The impact of the electrolarynx on voice-related quality of life*. Presented at the American Speech-Language-Hearing Association Annual Convention, Atlanta, GA.
1. Fung, K., **Cox, S. R.**, & Doyle, P. C. (2012, April). *The influence of voice breaks on listener perception of laryngeal overpressure in adductor spasmodic dysphonia*. Presented at the 115th Annual Triological Society's Continuing Medical Education Meeting, San Diego, CA.

UNIVERSITY-AFFILIATED PRESENTATIONS

-
12. **Cox, S. R.** (2015, November). *Assessing voice and quality of life outcomes in individuals with voice disorders*. Adelphi University, Garden City, NY.

11. **Cox, S.R.** (2015, March). The impact of clear speech on electrolaryngeal speakers. Rehabilitation Sciences Journal Club. Western University, London, ON.
10. Doyle, P.C., Dzioba, A, **Cox, S. R.**, Brandt, M., Caty, M. E., Nash, M., Scott, G., Davis, E., Jackman, K., & Wilkinson., J. (2014, March). Two labs, one shared passion: Partners in cancer-related health, well-being, and quality of life. 2014 Faculty of Health Sciences Research Day, Western University, London, ON.
9. **Cox, S. R.**, Jeremic, G., Fung, F., & Doyle, P.C. (2014, February). *Measuring Mandibular Function after head and neck cancer treatment*. 2014 HRS Graduate Research Forum, London, ON.
8. Scott, G., **Cox, S. R.**, Nash, M., & Doyle, P.C. (2014, February). *Factors impacting speech acceptability of tracheoesophageal speakers*. 2014 HRS Graduate Research Forum, London, ON,
7. Khalili, S., **Cox, S.R.**, Fung, K., Parsa, V., & Doyle, P. C. (2013, May). *Correlates of Tracheoesophageal voice restoration outcomes: a surgical, acoustic and perceptual investigation*. Presented at the Schulich School of Medicine and Dentistry Resident's Research Day, Western University, London, ON.
6. Angel, D., **Cox, S. R.**, Fung, K., & Doyle, P.C. (2012, May). *Autologous plasma adhesives and wound healing in split thickness skin graft donor and recipient sites*. Presented at the Schulich School of Medicine and Dentistry Resident's Research Day, Western University, London, ON.

5. Jeremic, G., **Cox, S. R.**, Fung, K., & Doyle, P. C. (2012, May). *Evaluating trismus in head and neck patients treated with radiotherapy: A prospective study*. Presented at the Schulich School of Medicine and Dentistry Resident's Research Day, Western University, London, ON.
4. Khalili, S., **Cox, S. R.**, Fung, K., & Doyle, P.C. (2012, May). *An experimental investigation of voice quality in tracheoesophageal puncture and voice restoration*. Presented at the Schulich School of Medicine and Dentistry Resident's Research Day, Western University, London, ON.
3. Paradis, J., Zimmerman, K., Doyle, P. C., Brandt, M., Agrawal, S., Parnes, L..., **Cox, S. R.** (2012, May). *Cutaneous scars following simultaneous bilateral cochlear implants: A comparison of scalpel versus electrocautery incisions*. Presented at the Schulich School of Medicine and Dentistry Resident's Research Day, Western University, London, ON.
2. **Cox, S. R.**, Day, A.M.B., Bornbaum, C., Sleeth, L., Dzioba, A., & Doyle, P. C. (2012, February). *Voice-related disability: Uncovering the impact of the electrolarynx (EL) on Voice-Related Quality of Life (V-RQOL)*. Presented at the 5th Annual Graduate Research Forum, Western University, London, ON.
1. **Cox, S. R.**, Day, A.M.B., Bornbaum, C., Sleeth, L., Dzioba, A., & Doyle, P. C. (2012, February). *Electrolaryngeal use on Voice-Related Quality of Life (V-RQOL): Understanding the potential for Voice-Related Disability following total laryngectomy*. Presented at the ARGC/FHS Symposium, Parkwood Hospital, London, ON.

AWARDS AND ACCOMPLISHMENTS

2011 - 2015	Western Graduate Research Scholarship, Western University
2015	Meritorious Poster Award, American Speech-Language and Hearing Association (ASHA)
2014	Meritorious Poster Award, American Speech-Language and Hearing Association (ASHA)
2012 - 2014	Health and Rehabilitation Sciences (HRS) Travel Award
2012 - 2014	Faculty of Health Sciences (FHS) Travel Award
2013	Western University Ontario Student Opportunity Trust Fund Bursary
2007 - 2009	Out-of-State Merit Tuition Award, Central Michigan University
2004 - 2007	Brock University Entrance Scholarship Renewal, Brock University
2003 - 2007	Edward J. Freeland Award, Town of Fort Erie
2007	Dean's Honour List, Brock University
2003	Brock University Entrance Scholarship, Brock University
2003	Brock Scholars Award, Brock University
2003	Ontario Scholar Award
2003	Queen Elizabeth II Aiming for the Top Scholarship

SCHOLARLY AND PROFESSIONAL ACTIVITIES

2015 - Present	American Speech-Language-Hearing Association
2015 - Present	Acoustical Society of America, Student Member
2008 - 2016	National Student Speech Language and Hearing Association (NSSLHA), Student Member

- 2011 - 2015 Rehabilitation Sciences Journal Club, Student Member
- 2011 - 2015 Head Injury Association of Fort Erie (HIAFE),
Consulting Board Member