TOTAL LARYNGECTOMY

Exploring voice outcomes and functional issues

Klaske Elisabeth van Sluis

Total laryngectomy

Exploring voice outcomes and functional issues

Published by Ridderprint | www.ridderprint.nl

Cover design: Femke Zandberg

ISBN: 978-94-6416-238-7

Total laryngectomy

Exploring voice outcomes and functional issues

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 16 december 2020, te 14.00 uur

door

Klaske Elisabeth van Sluis

geboren te Leeuwarden

Promotiecommissie:

Promotores:	Prof. dr. M.W.M. van den Brekel	Universiteit van Amsterdam
	Prof. dr. P.P.G. Boersma	Universiteit van Amsterdam
Copromotores:	Dr. L. van der Molen	Universiteit van Amsterdam
	Dr. R.J.J.H. van Son	Universiteit van Amsterdam
Overige leden:	Prof. dr. E. Gerrits	Universiteit Utrecht
	Prof. dr. G. Van Nuffelen	Universiteit Antwerpen
	Prof. dr. J.C.M. van Weert	Universiteit van Amsterdam
	Prof. dr. J.E. Rispens	Universiteit van Amsterdam
	Prof. dr. I.M. Verdonck-de Leeuw	Vrije Universiteit Amsterdam
	Prof. dr. A.J.M. Balm	Universiteit van Amsterdam

Faculteit der Geesteswetenschappen

Het Nederlands Kanker Instituut heeft een financiële bijdrage ontvangen van Atos Medical AB (Hörby, Zweden), wat bijdraagt aan de bestaande infrastructuur van de afdeling hoofd-halsoncologie en -chirurgie om kwaliteit van leven onderzoek mogelijk te maken.
Het drukken van dit proefschrift werd mede mogelijk gemaakt door steun van Castor EDC, ChipSoft, Meditop Medical Products B.V., Atos Medical B.V., Science Plus Group, Pro Education, de Nederlandse Vereniging voor Logopedie en Foniatrie, Logoscientia, Prof. dr. Eelco Huizinga Stichting en de Patiëntenvereniging Hoofd-Hals.

Draai met de zon mee
zodanig dat je de warmte
in je hart kan blijven voelen
Patiënt na hoofd-halskanker

Contents

Author Contributions

Abbreviations of contributing authors

Klaske E. van Sluis	KvS
Rob J.J.H. van Son	RvS
Lisette van der Molen	LvdM
Michiel W.M. van den Brekel	MvdB
Paul P.G. Boersma	PB
Frans J.M. Hilgers	FH
Patrick A. Bhairosing	PAB
Anthony J. McGuinness	AM
Carsten E. Palme	CP
Daniel Novakovic	DN
Danielle Stone	DS
Lydia Natsis	LN
Emma Charters	EC
Kelly Jones	KJ
Richard Dirven	RD
Marijn Kapitein	MK
Maartje Leemans	ML
Anne F. Kornman	AK
Wim G. Groen	WG
Bari Hoffman-Ruddy	BH
Martijn M. Stuiver	MS
Gili Yaron	GY

Chapter 1

Manuscript preparation	KvS
Manuscript editing and review	RvS, LvdM, MvdB, PB

Chapter 2

van Sluis, K.E., van der Molen, L., van Son, R.J.J.H., Hilgers, F.J.M., Bhairosing, P.A., van den Brekel, M.W.M. "Objective and subjective voice outcomes: a systematic review." *European Archives of Oto-Rhino-Laryngoly* 275,1: 11-26, 2018.

Study concepts and design	KvS, LvdM, FH, RvS, MB
Data acquisition	KvS, PAB
Data analysis and interpretation	KvS, RvS, LvdM
Manuscript preparation	Kvs, LvdM, FH
Manuscript editing and review	KvS, LvdM, FH, RvS, MB

Chapter 3

van Sluis, K.E., van Son, R.J.J.H., van der Molen, L., McGuinness, A.J., Palme, C.E., Novakovic, D., Stone, D., Natsis, L., Charters, E., Jones, K., Dirven, R., van den Brekel, M.W.M. "Multidimensional evaluation of voice outcomes following total laryngectomy: a prospective multicenter cohort study." *European Archives of Oto-Rhino-Laryngoly*, 1-14, 2020.

Study concepts and design	KvS, RvS, LvdM, MB, RD
Data acquisition	KvS, LvdM, AM, CP, DN, DS, LN, EC, KJ, RD
Data analysis and interpretation	KvS, RvS, LvdM, MB, RD
Statistical analysis	KvS, RvS
Manuscript preparation	KvS, RvS
Manuscript editing and review	KvS, LvdM, RvS, AM, CP, DN, DS, LN, EC, KJ, RD, MB

Chapter 4

van Sluis, K., Kapitein, M., van Son, R., Boersma, P. "The acoustic contrast between the Dutch consonants /t/ and /d/ is reduced in tracheoesophageal speech." *Proceedings of International Congress of Phonetic Sciences ICPhS 2019, 5-9 August, Melbourne, Australia*, 2019.

Study concepts and design	KvS, RvS, MK
Data acquisition	KvS, RvS, MK
Data analysis and interpretation	KvS, RvS, MK, PB
Statistical analysis	KvS, RvS, PB
Manuscript preparation	KvS
Manuscript editing and review	KvS, RvS, PB

Chapter 5

van Sluis, K.E., van den Brekel, M.W.M., Hilgers, F.J.M., van Son, R.J.J.H. "Long-term stability of tracheoesophageal voices." *Proceedings of Interspeech 2016, San Francisco*: 102-106, 2016.

Study concepts and design	KvS, RvS, FH
Data acquisition	KvS, RvS
Data analysis and interpretation	KvS, RvS, FH, MvdB
Statistical analysis	KvS, RvS
Manuscript preparation	KvS, RvS
Manuscript editing and review	KvS, RvS, FH, MvdB

Chapter 6

Leemans, M, van Sluis, K.E., van Son, R.J.J.H., van den Brekel, M.W.M. "Interaction of functional and participation issues on quality of life after total laryngectomy." *Laryngoscope Investigative Otolaryngology* 5,3: 453-460, 2020.

Study concepts and design	KvS, ML, RvS, MvdB
Data acquisition	KvS, ML, RvS
Data analysis and interpretation	KvS, ML, RvS
Statistical analysis	KvS, ML, RvS
Manuscript preparation	KvS, ML
Manuscript editing and review	KvS, ML, RvS, MvdB

ML and KvS contributed equally to this chapter

Chapter 7

van Sluis, K.E., Kornman, A.F., Groen, W.G., van den Brekel, M.W.M., van der Molen, L., Hoffman-Ruddy, B., Stuiver, M.M. "Expiratory muscle strength training in patients after total laryngectomy; a feasibility pilot study.: *Annals of Otology, Rhinology & Laryngology*, epub ahead of print, 2020.

Study concepts and design	KvS, AK, FH, BH, WG, LvdM, MS, MvdB
Data acquisition	KvS, AK, WG, LvdM
Data analysis and interpretation	KvS, AK, MS, WG
Statistical analysis	KvS, AK, MS
Manuscript preparation	KvS, AK
Manuscript editing and review	KvS, AK, BH, WG, LvdM, MS, MvdB

AK and KvS contributed equally to this chapter

Chapter 8

van, Sluis, K.E., Kornman, A.F., van der Molen, L., van den Brekel, M.W.M., and Yaron, G. "Women's perspective on life after total laryngectomy: a qualitative study." *International Journal of Language & Communication Disorders.* 55,2: 188-199, 2020.

Study concepts and design	KvS, AK, MvdB, GY
Data acquisition	KvS, AK, GY
Data analysis and interpretation	KvS, AK, LvdM, MvdB, GY
Manuscript preparation	KvS, AK, GY
Manuscript editing and review	KvS, AK, MvdB, LvdM, GY

CHAPTER 1

General introduction

1.1 Introduction

This thesis focuses on individuals who have had head and neck cancer and had to undergo a surgical procedure called total laryngectomy. This is an organ-sacrificing surgery which leads to lifelong consequences in anatomy and physiology of the body. During the total laryngectomy procedure, the larynx is removed, which leads to a loss of the natural voice. Our mission with this thesis is to contribute to better understanding of voice and speech outcomes after total laryngectomy as well as the functional and participation issues which affect well-being. Based on these findings I aim to identify rehabilitation gaps and formulate recommendations to improve clinical practice and ongoing research.

The thesis is divided into two sections. The first chapters focus on voice and speech outcomes after total laryngectomy. The second part of this thesis focuses on functional issues and psychosocial functioning after total laryngectomy. This introductory chapter will help the reader understand the consequences of a total laryngectomy. The available speech rehabilitation options will be discussed as well as the tools to evaluate quality of voice and speech outcomes. This is followed by a section on the functional and participation issues which total laryngectomy patients face.

1.2 Head and Neck Cancer

Head and neck cancer accounts for approximately 5% of all malignant tumours and includes cancer of the oral cavity, pharynx, larynx and oropharynx [1, 2]. Every year, 140.000 people in Europe are diagnosed with head and neck cancer, of which 40.000 are diagnosed with larynx cancer [3]. In the Netherlands both laryngeal and hypo-pharyngeal cancers are typically squamous cell carcinomas. The peak age for these cancers is between 55 and 74 years. Both cancers are more often diagnosed in men than in women (4:1), although the incidence among men is decreasing and the incidence among women is stable or slightly increasing [4].

1.3 Functions of the larynx

The larynx has a crucial part in basic physical functions including breathing, speaking and swallowing. Air which is inhaled via the nose or mouth passes the larynx when entering the trachea and lungs. The air is humidified and warmed by the upper respiratory tract. The larynx can produce sound; it is also referred to as the voice box, since the vocal folds are situated within it. During the phonatory process, air is ejected from the lungs through the glottis; the vocal folds move closer together which results in oscillation of the vocal cords [5]. The volume of the voice is determined by the pressure of the pulmonary air blown through the vocal folds, whilst the fundamental frequency (pitch) is defined by the frequency of the mucosal waves [5]. The quality of the voice depends on the myoelastic characteristics of the vocal folds, degree of vocal fold closure, irregularities during oscillation, and resonance characteristics [5, 6]. The generated voice can be articulated into speech with help of the structures in the pharynx, mouth and lips. This process of speech production can be described according to the source-filter theory [7]. This theory states that speech is a representation of the produced sound source (e.g. voice) which is formed into speech by the resonators (oral cavities, nasal cavaties) and the articulators (tongue, teeth and lips) functioning as a filter. For intake of food and drinks, the larynx has an important role during swallowing. Within the act of swallowing the epiglottis covers the entrance of the larynx which protects the airway for aspiration. In Figure 1.1 a schematic drawing of the lateral view of the normal anatomy is shown with the larynx in situ.

1.4 Total laryngectomy

A total laryngectomy refers to an organ-sacrificing surgery that can be indicated for laryngeal and hypo-pharyngeal cancers. An indication depends on the degree of invasiveness (indicated with T-stage), size, site and recurrence of the cancer, laryngeal functionality, the patient's condition and patient's and doc-

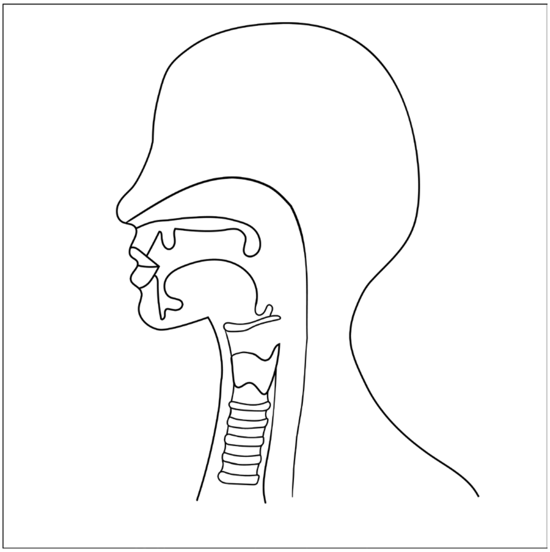

Figure 1.1: Schematic drawing of a lateral view of the normal anatomy with the larynx in situ.

tor's preferences [8]. Working towards an educated treatment choice is referred to as shared decision making. Three main indications for total laryngectomy are present. Firstly, a total laryngectomy can be a primary treatment for advanced laryngeal and hypo-pharyngeal cancers, the T3 and T4 staged cancers. These cancers can be primary, the first diagnosed cancer, or secondary primary; the latter referring to a new primary cancer in a patient that already had cancer. A total laryngectomy for primary cancer is typically combined with post-laryngectomy (chemo)radiotherapy, whereas patients with secondary primary cancer often have already received (chemo)radiotherapy for their first cancer [4, 9, 10]. Secondly, a total laryngectomy can be a salvage treatment in case of recurrent or residual disease that is not successfully treated with organ-preserving treatment, such as (chemo)radiotherapy [10]. In approximately one fourth of the patients with advanced larynx cancer, this salvage treatment needs to be performed since the (chemo)radiotherapy or minimally invasive procedures were not sufficient [11, 12]. Finally, laryngeal functionality can become so impaired due to previous (chemo)radiotherapy that a total laryngectomy is required [13, 14]. In this case the procedure is performed as a result of a dysfunctional larynx. Patients with a dysfunctional larynx often suffer ongoing aspiration and/or obstruction in breathing which makes them dependent on a tracheotomy, e.g. a canulla [13, 14].

The surgical procedure of a total laryngectomy generally involves removal of the larynx, resection of the trachea, removal of the hypopharynx in accordance with the extent of invasion and reconstruction of the neopharynx. In case of limited invasion a primary closure of the pharynx can be performed, but in more extensive resections, reconstruction of the lost tissue is necessary by reinforcing tissue from another part of the body, e.g. a pectoralis major flap, radial forearm flap or gastric pull-up. For the patient to be able to breathe, a tracheal stoma is made by bringing the trachea to the skin in the inferior anterior area of the neck. When possible, a primary insertion of voice prosthesis is performed allowing early vocal rehabilitation with tracheoesophageal speech.

1.5 Speech rehabilitation after total laryngectomy

Due to the removal of the larynx, one of the immediate consequences is that the patient loses his ability to generate voice; the patient will need a substitute voice. Substitute voicing is defined as 'voicing without the true vocal folds'. After total laryngectomy the three most used options for creating a substitute-voice are tracheoesophageal speech, artificial larynx, and esophageal speech.

1.5.1 Tracheoesophageal speech

The majority of this thesis is focused on outcomes in tracheoesophageal speech, which is considered the preferred speech rehabilitation option in the Western world. A fistula is created between the trachea and esophagus, in which a valve (voice prosthesis) is placed. This re-establishes the connection between the lungs and the mouth. When the stoma is occluded, air from the lungs enters the esophagus, which results in vibrations in the pharyngo-esophageal segment. The pharyngoesophageal segment (PE-segment), also referred to as the neoglottis, serves as the new vibratory voicing source in tracheoesophageal speech. In contrast to the quasi-symmetrical vocal folds, the vibrating neoglottis consists of amorphic vibrating elements. The whole vibrating segment is larger since it has more mass and therefore its vibrations are typically less regular compared to laryngeal speech. Tracheoesophageal speech is generalized by a lower fundamental frequency, a rougher voice quality, reduced voicing distinctions in consonants and abnormal prosody. Tracheoesophageal speech can be perceptually hoarse or breathy, it can be noisy, caused by air seeping from the stoma during speech or the gargling noise of saliva or mucus flowing down the throat [15, 16]. The general speaking rate is reduced with fewer words and syllables per minute and longer or greater number of pauses [16–18]. In speech production, problems occur in producing voiced and voiceless consonants, voicing vowels, maintaining pitch, and producing specific phonemes [19]. In Figure 1.2 a schematic drawing of tracheoesophageal speech is visualized.

1.5.2 Esophageal speech

Esophageal speech is performed by administering air into the esophagus, which is subsequently expelled, causing mucosal vibrations in the PE-segment. Prior

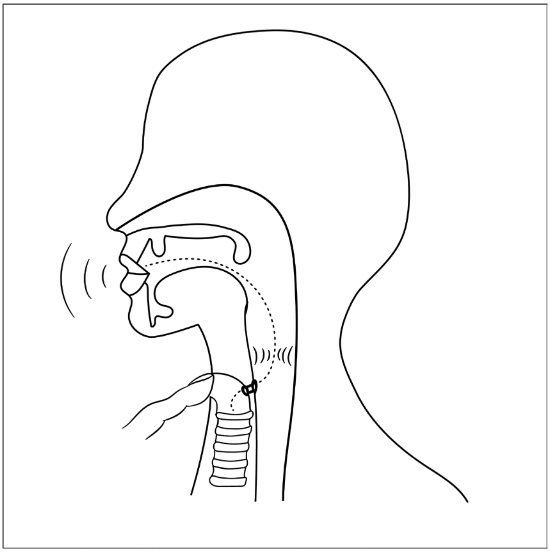

Figure 1.2: Schematic drawing of tracheoesophageal speech. With manual occlusion of the stoma, air from the lungs passes through the voice prosthesis to allow the pharyngo-esophageal segment to vibrate and generate voice.

to the introduction of tracheoesophageal speech in the early 1980s, esophageal speech played a prominent role in speech rehabilitation after total laryngectomy. Esophageal speech does offer advantages over tracheoesophageal speech and artificial larynx speech: it does not require a mechanical or prosthetic device and the hands are not occupied during voicing. Disadvantages of esophageal speech are the short phonation time, reduced loudness, and lower success rate, and it has the longest training time of all of the speech rehabilitation options. Esophageal speech is produced by injection of air into the esophagus. The esophagus serves as the air reservoir; the PE-segment is set into vibration. The reported volume of this esophageal reservoir is approximately 40 to 80cc [20]. There are two main ways to get air into the esophagus for voicing purposes which are classified as a positive pressure approach and a negative pressure approach. In the positive pressure approach, air is injected into the esophagus on a consonant or with a tongue pumping method. In the negative pressure approach an inhalation technique is used to create a negative air pressure into the upper esophagus. When the air is released, the PE-segment vibrates which creates voicing. Both methods require oral, motor and cognitive skills of the patient. Both approaches to get the air into the esophagus can disrupt articulatory movements and reduce fluency of speech. Therefore, longer training periods are required, often up to months [21]. Since tracheoesophageal speech is considered as the preferred method, nowadays, fewer speech and language pathologists (SLP's) are proficient at teaching esophageal speech. In Figure 1.3

esophageal speech is visualized.

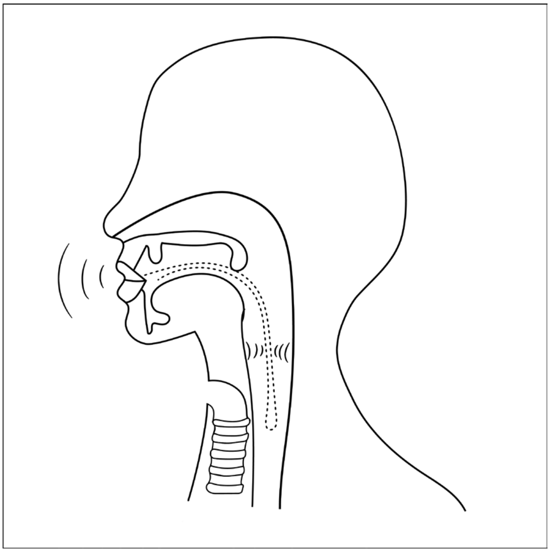

Figure 1.3: Schematic drawing of esophageal speech. Air from the mouth is injected in the esophagus and subsequently expelled, which results in mucosal vibrations of the pharyngo-esophageal segment and generates voice.

1.5.3 Electrolarynx speech

Speech rehabilitation with an artificial larynx involves a substitute sound source. This can be a pneumatic device, a pipe leading from the stoma to the mouth housing a rubber membrame which vibrates on pulmonary air stream. Or an electrolarynx is used, a (mostly) handheld device which can be placed against the neck or cheek. This electronic device mechanically generates sound which is transmitted through the tissue, the tissue turning air within the vocal tract into vibration. Speech is generated by articulatory movements. Typically these devices have separate buttons for pitch and loudness manipulations. An advantage of electrolarynx speech is that it can be utilized within the first few days after the laryngectomy surgery, it is low-cost and has a fairly high success rate. Disadvantages are that the sound of the voice is mechanical and perceived as unnatural and that intelligibility is less than tacheoesophageal speech and esophageal speech [21]. In Figure 1.4 electrolarynx speech is visualized.

1.5.4 Functional issues following total laryngectomy

To help patients adapt to post-laryngectomy life and to minimize negative consequences, multidisciplinary rehabilitation is provided. Apart from the changes

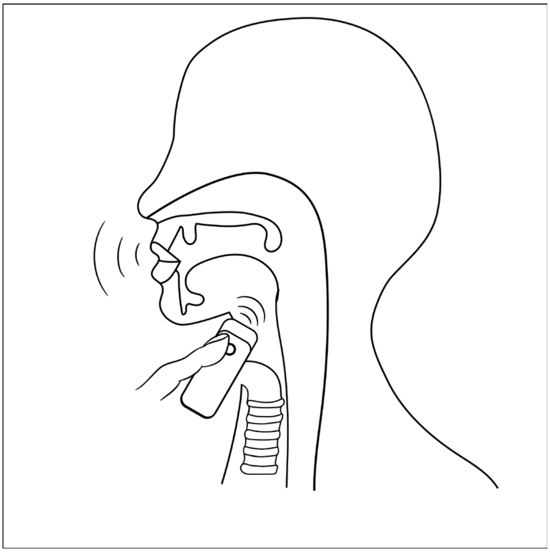

Figure 1.4: Schematic drawing of electrolarynx speech. The electrolarynx is placed against the neck, the generated sound transmitted through the tissue into the mouth and articulated into speech.

in voice and speech, the total laryngectomy procedure leads to lifelong changes in physical, psychological and social functioning. The surgical procedure leads to changes in respiration, swallowing, olfaction, and general health. This can severely impact the patients' normal daily functioning and their general well-being.

1.6 Multidimensional analysis of voice and speech

To evaluate voice and speech it is recommended to perform a multidimensional analysis, which evaluates both objective and subjective aspects. Multidimensional analysis of voice and speech includes analysis of perceptual, acoustic, and patient-reported outcome measures (PROM) [22]. Substitute voices are generally characterized by strong irregular voice quality compared to healthy speakers; therefore, a well-thought-out approach is required to evaluate substitute speech [23].

1.6.1 Perceptual analysis

Perceptual evaluation includes standardized listening experiments. As communication is mostly a perceptual matter, auditory-perceptual evaluations are considered as the "gold standard". For laryngeal voices perceptual evaluation is standardly performed with the GRBAS scale, including the five parameters Grade, Roughness, Breathiness, Asthenia, and Strain. Unfortunately, this scale is not sufficient for evaluating substitute voicing, as these largely irregular voicing types often score the highest grade of dysphonia. For substitute voices, judgements of experienced SLP's are considered as more proficient than judgements of naïve raters [24]. In the literature various variables are applied to evaluate substitute voices by perceptual ratings. Moerman and colleagues developed the IINFVo scale to assess substitute speech perceptually. The five IINFVo scale parameters are: overall impression (I), impression of intelligibility (I), unintended additive noise (N), fluency (F) and voicing (Vo) [25]. It appears that the variables can be reduced to two perceptual dimensions which are the most reliable: acceptability and intelligibility [25, 26]. Overall impression or acceptability of the voice refers to the degree to which speech is agreeable to the ear and has a pleasant quality. Intelligibility refers to the clarity and understandability of words and sentences [25, 26].

1.6.2 Acoustic analysis

Acoustic analysis of the voice includes analysis of the signal of audio recordings of the voice and speech. Acoustic outcomes provide an objective measure of vocal functioning. The characteristics of the voice signal can be analyzed with the help of software; frequently used software is Praat [27]. Standard acoustic measures of a sustained vowel /a/ include perturbation measures, providing information about the period-to-period and amplitude variation of the voice signal, jitter and shimmer, respectively. A limitation of this perturbation analysis is that the systems often cannot analyze strongly aperiodic signals which are present in substitute voicing, therefore the validity and clinical utility of many of these acoustic measures has been strongly debated [28–30]. Thus, additional information can be provided by measuring the proportion of voiced frames and signal-to-noise ratio computations. Signal-to-noise ratio computations include harmonics-to-noise-ratio and cepstrum peak. Over the last decade automatic speech and voice analysis has become field of interest [30–32]. Automatic analysis is promising in providing consistent ratings and analyzing trends within a single speaker. To perform an automatic analysis which combines several acoustic parameters in one score the Acoustic Voice Quality Index can be used [30]. The AVQI is a measurement of six acoustic domains in time, frequency and quefrency modeled in in a linear regression formula.

1.6.3 Patient-reported outcome measures

Patient-reported outcome measures (PROM's) are ratings performed by the patient; patients evaluate their own voice quality, understandability, and functioning with the voice. This evaluation is subjective by definition, but very important in clinical practice. It is the patient who has to function in everyday life with the changed voice; therefore these questionnaires often assess the difficulties or handicaps which are experienced. Simple but effective assessment of the voice can be performed with a visual analogue scale of 100mm to obtain an impression of the voice quality, intelligibility and/or the impact the voice functioning has on performing everyday (social) activities. Different questionnaires have been developed that focus on voice functioning, e.g. the voice handicap index is a widely used questionnaire, which is available in a 30-item and a 10-item version [33, 34]. The questionnaire has been validated for use in laryngeal speakers experiencing voice problems but not specifically for total laryngectomy patients.

1.7 Aims and scope of this thesis

In this PhD thesis I aim to investigate voice and speech outcomes as well as functional outcomes and functional consequences after total laryngectomy. The thesis is divided in two sections; the first section includes chapters 2 to 5 and presents outcomes on voice and speech after total laryngectomy. **Chapter 2** presents a systematic literature review which was conducted to evaluate comparative acoustic, perceptual, and patient-reported outcomes for esophageal speech, tracheoesophageal speech, and electrolarynx speech in comparison with healthy speakers. In **Chapter 3** we report on our study which is unique in prospectively assessing the course of acoustic voice outcomes, reported voice problems, and quality of life in total laryngectomy patients from pre-laryngectomy up to one year post-surgery. In **Chapter 4** an investigation of the acoustic changes of the consonants /t/ and /d/ in Dutch speaking individuals before and after total laryngectomy is presented. **Chapter 5** presents the outcomes of longitudinally collected voice recordings in individuals with a time range of seven to eighteen years, assessing if voice quality and intelligibility is stable on the long-term in tracheoesophageal speakers. The second section, covering Chapters 6 to 9, presents studies on functional issues, rehabilitation strategies for pulmonary complaints and women-specific problems. **Chapter 6** presents the results of a worldwide survey assessing functional and participation issues in a group of 1705 respondents with total laryngectomy. In **Chapter 7** we report on the first Expiratory Muscle Strength Training we provided to patients after total laryngectomy which is aimed to reduce their pulmonary complaints. **Chapter 8** describes the results of a qualitative study in which we interviewed female patients on their perspective of life after total laryngectomy. In **Chapter 9** the results described in this thesis are discussed and recommendations for clinical practice and suggestions for future research are given. In **Chapter 10** a summary of this thesis in Dutch and English is presented.

References

[1] M. M. Öztürk, A., "Determination of problems in patients with post-laryngectomy," *Scandinavian Journal of Psychology*, vol. 54, no. 2, pp. 107–111, 2013.

[2] T. D. Woodard, A. Oplatek, and G. J. Petruzzelli, "Life after total laryngectomy: a measure of long-term survival, function, and quality of life," *Archives of Otolaryngology–Head & Neck Surgery*, vol. 133, no. 6, pp. 526–532, 2007.

[3] M. Wells, M. Cunningham, H. Lang, S. Swartzman, J. Philp, L. Taylor, and J. Thomson, "Distress, concerns and unmet needs in survivors of head and neck cancer: a cross-sectional survey," *European Journal of Cancer Care*, vol. 24, no. 5, pp. 748–760, 2015.

[4] A. J. Timmermans, C. J. de Gooijer, O. Hamming-Vrieze, F. J. Hilgers, and M. W. van den Brekel, "T3-t4 laryngeal cancer in the netherlands cancer institute; 10-year results of the consistent application of an organ-preserving/-sacrificing protocol," *Head & Neck*, vol. 37, no. 10, pp. 1495–1503, 2015.

[5] I. R. Titze, *Vocal fold physiology: Frontiers in basic science*, vol. 7. Singular Publishing Group, 1993.

[6] I. Jacobi, L. van der Molen, H. Huiskens, M. A. Van Rossum, and F. J. Hilgers, "Voice and speech outcomes of chemoradiation for advanced head and neck cancer: a systematic review," *European Archives of Oto-Rhino-Laryngology*, vol. 267, no. 10, pp. 1495–1505, 2010.

[7] G. Fant, "The source filter concept in voice production," *STL-QPSR*, vol. 1, no. 1981, pp. 21–37, 1981.

[8] P. Sheahan, "Management of advanced laryngeal cancer," *Rambam Maimonides medical journal*, vol. 5, no. 2, p. e0015, 2014.

[9] O. Ceachir, R. Hainarosie, and V. Zainea, "Total laryngectomy–past, present, future," *Mœdica*, vol. 9, no. 2, pp. 210–216, 2014.

[10] J. L. Lefebvre, Y. Pointreau, F. Rolland, M. Alfonsi, A. Baudoux, C. Sire, D. de Raucourt, O. Malard, M. Degardin, C. Tuchais, *et al.*, "Induction chemotherapy followed by either chemoradiotherapy or bioradiotherapy for larynx preservation: the tremplin randomized phase ii study," *J Clin Oncol*, vol. 31, no. 7, pp. 853–859, 2013.

[11] H. F. A. N. B. A. Ackerstaff, A.H., "Communication, functional disorders and lifestyle changes after total laryngectomy," *Clinical Otolaryngology & Applied Sciences*, vol. 19, no. 4, pp. 295–300, 1994.

[12] L. van der Molen, A. F. Kornman, M. N. Latenstein, M. W. van den Brekel, and F. J. Hilgers, "Practice of laryngectomy rehabilitation interventions: a perspective from europe/the netherlands," *Current Opinion in Otolaryngology & Head and Neck Surgery*, vol. 21, no. 3, pp. 230–238, 2013.

[13] K. A. Hutcheson, C. P. Alvarez, D. A. Barringer, M. E. Kupferman, P. R. Lapine, and J. S. Lewin, "Outcomes of elective total laryngectomy for laryngopharyngeal dysfunction in disease-free head and neck cancer survivors," *Otolaryngology–Head and Neck Surgery*, vol. 146, no. 4, pp. 585–590, 2012.

[14] M. C. Topf, L. C. Magaña, K. Salmon, J. Hamilton, W. M. Keane, A. Luginbuhl, J. M. Curry, D. M. Cognetti, M. Boon, and J. R. Spiegel, "Safety and efficacy of functional laryngectomy for end-stage dysphagia," *The Laryngoscope*, vol. 128, no. 3, pp. 597–602, 2018.

[15] N. Deore, S. Datta, R. Dwivedi, R. Palav, R. Shah, S. Sayed, M. Jagde, and R. Kazi, "Acoustic analysis of tracheo-oesophageal voice in male total laryngectomy patients," *The Annals of The Royal College of Surgeons of England*, vol. 93, pp. 523–527, Oct. 2011.

[16] T. Drugman, M. Rijckaert, C. Janssens, and M. Remacle, "Tracheoesophageal speech: A dedicated objective acoustic assessment," *Computer Speech & Language*, vol. 30, no. 1, pp. 16–31, 2015.

[17] J. Robbins, H. B. Fisher, E. C. Blom, and M. I. Singer, "A Comparative Acoustic Study of Normal, Esophageal, and Tracheoesophageal Speech Production," *Journal of Speech and Hearing Disorders*, vol. 49, pp. 202–210, May 1984.

[18] M. H. Bellandese, J. W. Lerman, and H. R. Gilbert, "An Acoustic Analysis of Excellent Female Esophageal, Tracheoesophageal, and Laryngeal Speakers," *Journal of Speech, Language, and Hearing Research*, vol. 44, pp. 1315–1320, Dec. 2001.

[19] P. Jongmans, F. Hilgers, L. Pols, and C. van As-Brooks, "The intelligibility of tracheoesophageal speech, with an emphasis on the voiced-voiceless distinction," *Logopedics Phoniatrics Vocology*, vol. 31, no. 4, pp. 172–181, 2006.

[20] J. Van den Berg and A. J. Moolenaar-Bijl, "Cricopharyngeal sphincter, pitch, intensity, and fluency in oesophageal speech.," *Folia Phoniatrica*, vol. 10, pp. 65–84, 1959.

[21] E. C. Ward and C. J. van As-Brooks, *Head and neck cancer: treatment, rehabilitation, and outcomes.* Plural Publishing, 2014.

[22] P. H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. Van De Heyning, M. Remacle, and V. Woisard, "A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques," *European Archives of Oto-rhino-laryngology*, vol. 258, no. 2, pp. 77–82, 2001.

[23] M. Moerman, J.-P. Martens, and P. Dejonckere, "Multidimensional assessment of strongly irregular voices such as in substitution voicing and spasmodic dysphonia: A compilation of own research," *Logopedics Phoniatrics Vocology*, vol. 40, pp. 24–29, Jan. 2015.

[24] C. J. van As, F. J. Koopmans-van Beinum, L. C. Pols, and F. J. Hilgers, "Perceptual evaluation of tracheoesophageal speech by naive and experienced judges through the use of semantic differential scales," *Journal of Speech, Language, and Hearing Research*, vol. 46, no. 4, pp. 947–959, 2003.

[25] M. Moerman, J.-P. Martens, and P. Dejonckere, "Application of the voice handicap index in 45 patients with substitution voicing after total laryngectomy," *European Archives of Oto-Rhino-Laryngology*, vol. 261, no. 8, pp. 423–428, 2004.

[26] T. Most, Y. Tobin, and R. C. Mimran, "Acoustic and perceptual characteristics of esophageal and tracheoesophageal speech production," *Journal of Communication Disorders*, vol. 33, no. 2, pp. 165–181, 2000.

[27] P. P. G. Boersma and D. J. M. Weenink, *Praat: Doing Phonetics by Computer: Version 3.4*. [computer program] Instituut voor Fonetische Wetenschappen, 1996.

[28] M. S. De Bodt, F. L. Wuyts, P. H. Van de Heyning, and C. Croux, "Test-retest study of the grbas scale: influence of experience and professional background on perceptual rating of voice quality," *Journal of Voice*, vol. 11, no. 1, pp. 74–80, 1997.

[29] I. R. Titze, *Workshop on acoustic voice analysis: Summary statement*. National Center for Voice and Speech, 1995.

[30] Y. Maryn, P. Corthals, P. Van Cauwenberge, N. Roy, and M. De Bodt, "Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels," *Journal of Voice*, vol. 24, no. 5, pp. 540–555, 2010.

[31] R. Clapham, C. Middag, F. Hilgers, J.-P. Martens, M. Van Den Brekel, and R. Van Son, "Developing automatic articulation, phonation and accent assessment techniques for speakers treated for advanced head and neck cancer," *Speech Communication*, vol. 59, pp. 44–54, 2014.

[32] C. Middag, R. Clapham, R. Van Son, and J.-P. Martens, "Robust automatic intelligibility assessment techniques evaluated on speakers treated for head and neck cancer," *Computer Speech & Language*, vol. 28, no. 2, pp. 467–482, 2014.

[33] C. A. Rosen, A. S. Lee, J. Osborne, T. Zullo, and T. Murry, "Development and validation of the voice handicap index-10," *The Laryngoscope*, vol. 114, no. 9, pp. 1549–56, 2004.

[34] B. H. Jacobson, A. Johnson, C. Grywalski, A. Silbergleit, G. Jacobson, M. S. Benninger, and C. W. Newman, "The voice handicap index (vhi) development and validation," *American Journal of Speech-Language Pathology*, vol. 6, no. 3, pp. 66–70, 1997.

CHAPTER 2

Objective and subjective voice outcomes after total laryngectomy: a systematic review

Abstract

Background: Esophageal speech (ES), tracheoesophageal speech (TES) and/or electrolarynx speech (ELS) are three speech rehabilitation methods which are commonly provided after total laryngectomy (TL).

Methods: A systematic review of the literature was conducted to evaluate comparative acoustic, perceptual, and patient-reported outcomes for ES, TES, ELS and healthy speakers.

Results: Twenty-six articles could be included. In most studies, methodological quality was low. It is likely that an inclusion bias exists, many studies only included exceptional speakers. Significant better outcomes are reported for TES compared to ES for the acoustic parameters, fundamental frequency, maximum phonation time and intensity. Perceptually, TES is rated with a significant better voice quality and intelligibility than ES and ELS. None of the speech rehabilitation groups reported clearly better outcomes in patient-reported outcomes.

[0] Appeared as: van Sluis, K.E., van der Molen, L., van Son, R.J.J.H., Hilgers, F.J.M., Bhairosing, P.A., van den Brekel, M.W.M. "Objective and subjective voice outcomes: a systematic review." *European Archives of Oto-Rhino-Laryngoly* 275,1: 11-26, 2018.

Conclusions: Studies on speech outcomes after TL are flawed in design and represent weak levels of evidence. There is an urge for standardized measurement tools for evaluations of substitute voice speakers. TES is the favorable speech rehabilitation method according to acoustic and perceptual outcomes. All speaker groups after TL report a degree of voice handicap. Knowledge of caretakers and differences in health care and insurance systems play a role in the speech rehabilitation options that can be offered.

2.1 Introduction

As a consequence of total laryngectomy (TL), patients lose their natural voice, making speech rehabilitation with a substitute sound source a major rehabilitation goal. The three main rehabilitation options are esophageal speech (ES), tracheoesophageal speech (TES), and electrolarynx speech (ELS) [1]. ES and TES have in common that the substitute sound source is internal, i.e. the voice is produced in the pharyngoesophageal (PE) segment. ES is performed by administering air into the esophagus, which is subsequently expelled, causing mucosal vibrations in the PE-segment. In TES, pulmonary air channeled through a voice prosthesis or tracheoesophageal (TE) fistula. The voice prosthesis enables pulmonary air to enter the esophagus, and prevents esophageal content from entering the airway. In TES, pulmonary air is the driving force for the mucosal vibrations in the PE-segment. In ELS the substitute sound source is external: an electrolarynx is a sound producing, mostly handheld device, which can be placed against the neck or cheek [1].

Worldwide, no evidence-based consensus exists on which speech rehabilitation method is best for restoring oral communication. It is often assumed that for TL patients a better voice quality is associated with an improved quality of life [2, 3].

For evaluating speech rehabilitation outcomes, multidimensional assessment is recommended [4, 5]. This systematic review focuses on acoustic analysis, perceptual evaluation, and patient-reported outcomes (PROs) of the three substitute speech options. Acoustic voice analysis regularly includes pitch and amplitude measurements [6]. However, standard acoustic voice analyses are not always sufficient to measure substitute voices, because speech originating in the vibrating PE-segment, ES and TES, is known to contain more noise components and less regularity than laryngeal voice [7]. Perceptual evaluations of the speech rehabilitation methods also require a well-thought-out approach because of the deviances in regularity compared to laryngeal voices [8, 9]. Most convenient for such evaluations of substitute voices are overall impression of voice quality and impression of speech intelligibility [8, 9]. Results of speech rehabilitation from a patient's perspective are mostly evaluated by Quality of Life (QOL) questionnaires such as those of the European Organization for Research and Treatment of Cancer, Quality of Life Questionnaire (EORTC), the module for patients with head and neck cancer 35-item version (EORTC QLQ-H&N35)

and/or the EORTC QLQ-C30 questionnaire, which include questions about speech functioning [10, 11]. PROs, such as the Voice Handicap Index (VHI) and Voice-Related Quality of Life (V-RQOL) provide more detailed evaluations of speech rehabilitation results [12–14].

At present, a comprehensive literature review on the advantages and disadvantages of the current speech rehabilitation options has not been performed. Collecting the best evidence available on the three speech rehabilitation methods will likely help to build consensus about which speech rehabilitation after TL is optimal, and could aid in clinicians' decision-making, patients' counseling and reimbursement issues. In this systematic review, we focus on obtaining comparative acoustic, perceptual, and PROs for the three speech rehabilitation methods after TL. We aim to identify how outcomes of the various speech rehabilitation methods relate to those of normal laryngeal speech (healthy speakers), and what outcomes are favorable for each rehabilitation method.

2.2 Materials and methods

The literature on speech outcomes after total laryngectomy (TL) was reviewed by means of a systematic search strategy. This search strategy was conducted with specific attention to the primary and secondary outcomes of interest (Table 2.1). The most suitable primary and secondary outcomes were selected based on the literature. With the acoustic outcomes, we aimed to obtain objective information about the speech rehabilitation options. We aimed to obtain subjective information of the voices though perceptual ratings and PROs. We have chosen to indicate fundamental frequency (F0), Harmonics to Noise Ratio (HNR), and percentage of voicedness (%voiced) as primary acoustic outcomes. These outcomes are indicated by several authors to obtain information about the pitch, stability and the amount of noise components [7, 15–17]. Secondary acoustic outcomes of interest were jitter, shimmer, intensity, spectral tilt and maximum phonation time (MPT). These outcome variables are frequently used in the literature although some are known to be less reliable in substitute voicing [16, 17]. Primary perceptual outcomes of interest were overall impression of voice quality and intelligibility, derived from the IINFVo scale, where impression, intelligibility, noise, fluency and voicing is evaluated [18]. Secondary perceptual outcomes of interest were chosen from well-established perceptual assessment tools, such as the Grade Roughness Breathiness Asthenia Strain scale assessment (GRBAS [19], and other recommended perceptual parameters in TL-speech such as unintended additive noise, fluency, and voicing [8, 18]. Primary PROs were the widely used VHI [13] and V-RQOL[14]. As secondary PROs we included voice specific outcomes on the EORTC QLQ-H&N35 [11] and the EORTC QLQ-C30 [10], where general quality of life is evaluated including a specific subset of questions on communication. The literature search was performed by the medical information specialist. The search was conducted in PubMed, Embase (ovid), Scopus and PsychInfo. Terms searched for were "laryngectomy", "voice", "speech", "electrolarynx", "esophageal", "tracheoesophageal", "acoustics", "intelligibility", "voice quality", "quality of life" and their synonyms. The criteria for inclusion were that the written language was English, Dutch, German, Spanish or French. No filter for publication date was applied, and the search was performed in January 2016, with an update in December 2016.

Table 2.1: Research questions and outcomes of interest

Research questions		• Which comparative acoustic, perceptual, and PROs are available for TES, ES and ELS after TL • How do outcomes of TES, ES and ELS relate to those of normal laryngeal speech? • What outcomes are favorable in which rehabilitation method?
Primary Outcomes	**Acoustic**	**F0**: fundamental frequency, a result of the rate of vibration of the (neo) glottis [15] **HNR**: harmonics to noise ratio, ratio between the total energy of the periodic voice signal and the energy of noise components [15, 17, 20] **%voiced**: the percentage voicedness [9, 16, 17]
	Perceptual	**Voice quality**: impression of the overall voice quality [8, 9, 14] **Intelligibility**: impression of the intelligibility [8, 18]
	PROs	**VHI**: Voice Handicap Index [13] **V-RQOL**: voice-related quality of life [14]
Secondary Outcomes	**Acoustic**	**Jitter**: relative variability in the period-to-period frequency [20, 21] **Shimmer**: relative variability in the peak-to-peak amplitude [20, 21] **Intensity**: Loudness dB [8, 22] **Spectral tilt**: a comparison between low frequency energy (between 0 and 1 kHz) and high frequency energy (between 1 and 5 kHz) [21, 23] MPT: Maximum Phonation Time [22, 24]
	Perceptual	**GRBAS**: Grade Roughness Breathiness Asthenia Strain scale assessment [18, 19] **Unintended additive noise**: uncontrolled noises during speech [18] **Fluency**: the perceived smoothness of the sound production [18] **Voicing**: voicing is voiced or unvoiced where it is supposed to be voiced or unvoiced [18]
	PROs	**EORTC QLQ-H&N35**: European Organization for Research and Treatment of Cancer, Quality of Life Questionnaire module for patients with head and neck cancer 35-item version [11] **EORTC QLQ-C30**: European Organization for Research and Treatment of Cancer, Quality of Life Questionnaire C30 [10]

All study types were included. Publications were included when at least two types of speech methods were compared. In these cases the same procedures for acoustic and perceptual evaluations are warranted for the different speaker groups within one study. Speaker groups had to be $grn=7$. There had to be a comparison of two or more speaker groups within the study. At least one of the primary outcome measurements had to be reported. Studies were graded according to the criteria of risk on bias described by the Cochrane Handbook for Systematic Reviews of Interventions, shown in the appendix, i.e. A = low risk of bias, B = unclear risk of bias, and C = high risk of bias [25, 26]. Level A and level B rated studies were included. Articles were excluded when they mainly reported on device research, primary or secondary voice prosthesis placement, adverse effects, or on pulmonary rehabilitation. Articles only reporting on secondary outcomes or rated as a Level C article were excluded. Reference lists were checked to collect more data. Studies which were rated with low risk of bias (level A) were indicated as best evidence available. These level A rated studies are highlighted in the result section, as well as the significant outcomes reported by the included studies (level A and B). Only predefined outcome parameters were taken into account. Outliers were excluded in the overall results section and are elaborated on further in the discussion section. Outcomes of included studies were analyzed and, where possible, pooled. Data were tabulated and graphically represented in violin plots.

2.3 Results

The search and selection process is visualized in Figure 1. The first and second author screened 50 out of 2405 papers on title and abstract to meet the selection criteria for inclusion. The first author performed the remaining screening of titles and abstracts. Seventy papers were evaluated in full for relevance and validity by the first and second author. References of the articles that were retrieved in full were screened, which resulted in two additional articles. The first and second author both performed a critical appraisal of the design of the studies. The third author evaluated all non-English articles. A decision on the inclusion of the articles was made in consensus of the three raters.

The definitive selection included 28 publications. There were two papers that discussed the same study, but were written in two different languages [27, 28]. We included the English version [27]. Furthermore, there were two publications of the same author published in 2013 and 2015, the 2015 paper containing additional speakers and evaluations to the 2013 paper [29, 30]. Therefore, we have chosen to only report on the 2015 paper in this systematic review [30]. This left 26 papers for further evaluation (Table 2.2).

In Table 2.2 details of the selected studies are provided. The scope of the research, the number of included participants and risk of bias rating is shown. In total, only three of the 26 studies (12%) reached level A (low risk of bias), shown in bold [30–32]. The remaining articles reached level B (unclear risk of bias).

A total of 1,097 participants are included in the studies, only the groups of interest are taken into account. Groups of interest were ES (n=313), TES

(n=482), ELS (n=135), and a control group of healthy i.e. laryngeal speakers (n=167). Six studies only included male participants [37–41, 46]. One study only included female participants [33]. When gender is reported in studies, control groups are matched on or comparable to the sex of the TL groups. Fourteen studies [27, 32–34, 34, 37, 38, 40–42, 44, 45, 49, 52] reported acoustic outcomes; nine studies [30, 31, 36, 40, 41, 43, 46, 48, 54] reported perceptual outcomes; and eight studies [30, 39, 41, 43, 47, 49, 50, 53] reported PROs. Four studies [30, 40, 41, 43] reported a combination of outcome measures.

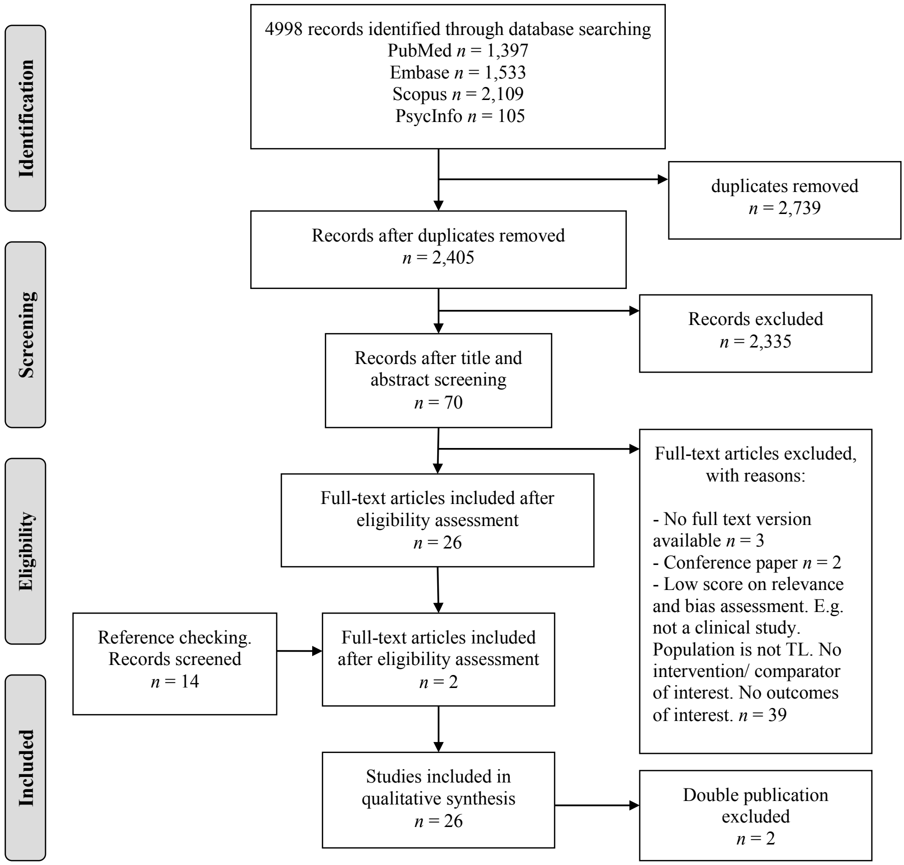

Figure 2.1: Flow diagram of study inclusion.

Table 2.2: Overview of specifications of the 26 included studies and their risk of bias assessment.

Author, year	Total N	ES n	TES n	ELS n	H n	Incl criteria	Gender	Age	Tumor location	Tumor stage	Treatment details	Time post-TL	Acoustic outcomes	Perceptual outcomes	PROs	Clear Tables and Figures	Reliability checks	Blinding	<30% loss-	Risk of Bias [26]
Arias et al. [27]	60	20	20	-	20	N	Y	Y	N	N	Y	N	Y	N	N	Y	N	-	Y	B
Bellandese et al. [33]	26	9	7	-	10	Y	Y	Y	N	N	Y	Y	Y	N	N	Y	Y	-	N	B
Blood [34]	30	10	10	-	10	N	N	Y	N	N	N	N	Y	N	N	Y	N	-	Y	B
Carello & Magnano [35]	14	7	7	-	-	N	Y	Y	N	N	N	N	Y	N	N	Y	N	-	Y	B
Crosetti et al. [31]	24	-	12	-	12	Y	Y	Y	Y	N	Y	Y	N	Y	N	Y	Y	Y	Y	A
De Maddalena et al. [36]	88	28	53	7	-	Y	Y	Y	N	N	N	N	N	Y	N	N	N	N	Y	B
DeBruyne et al. [37]	24	12	12	-	-	N	Y	Y	N	N	N	N	Y	N	N	N	N	-	Y	B
Deore et al. [38]	60	-	30	-	30	N	Y	N	Y	N	Y	N	Y	N	N	Y	N	-	Y	B
Eadie et al. [30]	36	2	23	11	-	Y	Y	Y	N	N	Y	Y	N	Y	Y	Y	Y	Y	Y	A
Evans et al. [39]	49	12	26	11	-	Y	Y	Y	N	N	Y	Y	N	N	Y	Y	-	N	Y	B
Finizia et al. [40]	22	-	12	-	10	Y	Y	Y	Y	Y	N	Y	Y	Y	N	N	Y	-	Y	B
Granda et al. [41]	18	10	8	-	-	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Y	N	-	Y	B
Kinishi & Amatsu [42]	35	5	20	-	10	N	Y	Y	Y	Y	N	Y	Y	N	N	Y	N	-	Y	B
Law et al. [43]	34	7	13	14	-	Y	Y	Y	N	N	N	Y	N	Y	Y	Y	Y	N	Y	B
Maccallum et al. [44]	20	10	-	-	10	N	Y	Y	N	N	N	N	Y	N	N	Y	N	-	Y	B
Merol et al. [45]	59	30	29	-	-	Y	Y	Y	N	N	Y	N	Y	N	N	N	-	-	Y	B
Miralles & Cervera [46]	40	10	20	-	10	Y	Y	Y	N	N	N	Y	N	Y	N	N	N	N	Y	B
Mourkabel et al. [47]	75	15	42	18	-	Y	Y	Y	N	Y	Y	Y	N	N	Y	Y	-	-	Y	B
Ng et al. [48]	42	15	12	15	-	Y	Y	Y	N	N	N	N	N	Y	N	N	Y	N	Y	B
Robbins et al. [49]	45	15	15	-	15	N	Y	Y	N	N	Y	N	Y	N	N	Y	N	-	Y	B
Rosso et al. [50]	48	13	20	15	-	Y	Y	Y	N	N	N	N	N	N	Y	Y	-	-	Y	B
Salturk et al. [51]	96	24	57	15	-	Y	N	Y	N	N	N	N	N	N	Y	Y	N	-	Y	B
Shim et al. [32]	40	20	-	-	20	Y	Y	Y	Y	N	Y	Y	Y	N	N	Y	Y	-	Y	A
Siric et al. [52]	20	10	10	-	-	N	Y	Y	N	N	Y	N	Y	N	N	Y	N	-	Y	B
Tiple et al. [53]	49	17	14	18	-	Y	Y	Y	N	N	N	Y	N	N	Y	N	N	N	Y	B
Williams & Watson [54]	43	12	10	11	10	N	N	Y	N	N	N	Y	N	Y	N	N	Y	N	Y	B

Answers are indicated by: Y = Yes, the study fulfilled the criterion; N = No, the study did not fulfill the criterion; - = insufficient information provided or the study did not address the particular outcome. In bold the publications rated with level A (low risk of bias) in the risk of bias analysis. Abbreviations: ES = esophageal speakers; TES = tracheoesophageal speakers; H = healthy speakers; TL = total laryngectomy; PROs = patient reported outcomes

Table 2.3: Comparative acoustic outcomes for speaker groups

	F0 vowel	F0 speech	HNR	MPT	Jitter	Shimmer	Intensity	Spectral tilt
ES>TES	Granda et al. [41] Merol et al. [45]	-	-	-	-	Siric et al. [52]	-	-
ES>H	Shim et al. [32]	-	-	-	-	-	-	
TES>ES	Arias et al. [27]* Bellandese et al. [33]* Blood [34]* Carello & Magnano [35] Kinishi & Amatsu [42] Robbins [49] Siric [52]*	Bellandese et al. [33] Robbins et al. [49] Blood [34]*	Arias et al. [27] Bellandese et al. [33] Carello & Magnano [35] Granda et al. [41] Siric et al. [52]	Carello & Magnano [35] Carello &Magnano [35] Granda et al. [41] Merol et al. [45] Robbins et al. [49] Siric et al. [52]*	Arias et al. [27] Carello & Magnano [35] Kinishi & Amatsu [42] Robbins et al. [49] Siric et al. [52]	Arias et al. [27] Robbins et al. [49]	Blood [34] Granda et al. [41] Siric et al. [52]*	DeBruyne et al. [37]
TES>H	-	-	-	-	-	-	-	-
H>ES	Arias et al. [27] Bellandese et al. [33] Blood [34] Kinishi & Amatsu [42]	Bellandese et al. [33] Blood [34] Robbins et al. [49]	Arias et al. [27]* Bellandese et al. [33] Maccallum et al. [44] **Shim et al.** [32]*	Robbins et al. [49]	Arias et al. [27]* Kinishi & Amatsu [42] Maccallum et al. [44] Robbins et al. [49] **Shim et al.** [32]*	Arias et al. [27]* Maccallum et al. [44] Robbins et al. [49] **Shim et al.** [32]*	Blood [34]* Robbins et al. [49]	**Shim et al.** [32]*
H>TES	Arias et al. [27]* Bellandese et al. [33] Blood [34] Deore et al. [38]* Finizia et al. [40] Kinishi & Amatsu [42] Robbins et al. [49]	Bellandese et al. [33] Blood. [34] Finizia et al. [40] Robbins et al. [49]	Arias et al. [27]* Bellandese et al. [33] Deore et al. [38]*	Deore et al. [38]* Finizia et al. [40] Robbins et al. [49]	Arias et al. [27]* Deore et al. [38]* Kinishi & Amatsu [42] Robbins et al. [49]	Arias et al. [27] Deore et al. [38] Robbins [49]	Blood [34] Robbins [49]	-

> Indicating a better mean group outcome. Level of significance was held at $p \leq .05$. Studies presented in bold had a level A risk of bias. ES esophageal speakers, TES tracheoesophageal speakers, ELS electrolarynx speakers, MPT maximum phonation time, V-RQOL voice-related quality of life

Ten studies [27, 34, 35, 37, 38, 42, 44, 49, 52, 54] did not mention any inclusion criteria. Eleven studies [27, 33–36, 38, 42, 44, 46, 46, 50, 54] failed to describe their method of selection and recruitment of participants. Two studies [40, 45] provided a detailed description of the selection process. The remaining articles [30, 32, 37, 39, 43, 47–49] mentioned the selection process briefly. Most patients were recruited via a clinical setting or via support groups.

Treatment details of the included groups were only provided in about half of the studies [27, 30–33, 38, 39, 41, 45, 47, 49, 52]. This variable was indicated as present when treatment details were provided. Nevertheless, surgical details, e.g. use of flaps during the surgery were not provided in any of the included studies.

2.3.1 Acoustic outcomes

In Table 2.3 acoustic outcomes for the included studies that reported on the primary (e.g. F0, HNR, MPT) and secondary outcomes (e.g. jitter, shimmer, intensity, spectral tilt) are presented. Comparative results for the different speaker groups are shown. None of the studies performed acoustical analysis on ELS, therefore ELS is not discussed in this section.

Fundamental frequency

Thirteen papers [27, 32–35, 37, 38, 40–42, 45, 49, 52] (n=443) reported fundamental frequency (F0) outcomes, including the level A categorized study of Shim et al. [32] (Figure 2.). Measurements are presented for evaluations in sustained vowels and in running speech. No distinction between male and female speakers is made. Most studies did not make this distinction since the sound source, the PE-segement, is similar in both groups. Not all F0 outcomes could be taken in account because in some studies the reporting was only range of F0 or in boxplots [37, 40]. Higher F0 values are designated as better [1]. The total range of F0 values for all groups of speakers in vowels and running speech is 64 - 179 Hz, which is reported in twelve studies [27, 32–35, 38, 40–42, 45, 49, 52]. The mean F0 value of 227 Hz for TES [52] and the mean F0 value of 246 Hz for ES [41] are considered outliers and were therefore excluded. The level A rated study of Shim et al. [32] showed non-significant higher mean F0 values for ES compared to healthy speakers, resp. 131 Hz and 124 Hz. Higher mean F0 values are found for healthy speakers compared to ES and TES (n=7) [27, 33, 34, 38, 40, 42, 49]. In two studies this difference was significant [27, 34]. For the speech rehabilitation methods higher F0 values are seen in the group of TES as compared to ES (n=7) [27, 33–35, 42, 49, 49]. In four studies this difference was significant [27, 33, 34, 52]

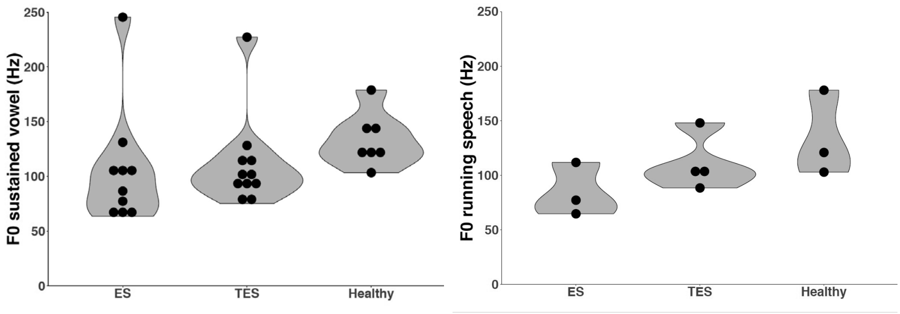

Figure 2.2: Left violin plot displaying the distribution of the mean F0 outcomes for sustained vowels in 12 studies [27, 32–35, 38, 40–42, 45, 49, 52] ES n=136, TES n=168, H n=115. Right violin plot displaying the distribution of the mean F0 outcomes for running speech in four studies [33, 34, 40, 49]. ES n=34, TES n=44, H n=35. Abbreviations: ES = esophageal speakers; TES = tracheoesophageal speakers; ELS = electrolarynx speakers; H = healthy speakers; F0 = fundamental frequency; Hz = Hertz.

Harmonics to noise ratio

Eight studies [27, 32, 33, 35, 38, 41, 44, 52] (n=240) included Harmonics to Noise Ratio (HNR) outcomes, including one level A study of Shim et al. [32], and seven level B studies [27, 33, 35, 38, 41, 44, 52]. Two studies measured Noise to Harmonics Ratio, which was recalculated to HNR [32, 35]. Higher HNR is reflecting better voice quality. The level A rated study of Shim et al. [32] showed better HNR outcomes for healthy speakers compared to ES. Healthy speech was rated as superior to substitute speech in five studies [27, 32, 33, 38, 44]. Comparison between ES and TES showed superior values for TES (Table 2.3; n=5) [27, 33, 35, 41, 52]. However, none of the studies comparing ES and TES found a significant difference between these substitute voices. This is reflected in the violin plots in Figure 3, depicting the comparable HNR outcomes for TES and ES.

Maximum phonation time

Eight level B rated studies [35, 37, 38, 40, 41, 45, 49, 52] (n=262) evaluated Maximum Phonation Time (MPT) by phonation of a vowel /a/ or /i/ (Table 2.3, Figure 3). Longer MPT is indicated as being better. Range of MPT values for all groups of speakers is 0.84sec - 23.87sec (n=8) [35, 37, 38, 40, 41, 45, 49, 52]. In the group of healthy speakers values for MPT are highest [38, 40, 49]. A significant longer MPT for healthy speakers is found in comparison with TES [38]. Within the comparison of the speech rehabilitation groups six studies [35, 37, 41, 45, 49, 52] found that TES has a longer MPT than ES, significant in one study [52].

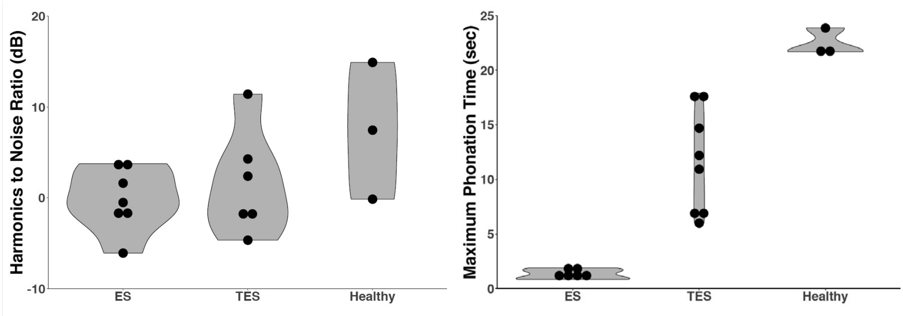

Figure 2.3: Left violin plot displaying the distribution of the mean HNR outcomes in eight studies [27, 32, 33, 35, 38, 41, 44, 52]. ES n=86, TES n=82, H n=90. Right violin plot displaying the distribution of the mean MPT outcomes in eight studies [35, 37, 38, 38, 40, 41, 45, 49, 52]. ES n=84, TES n=123, H n=55. Abbreviations: ES = esophageal speakers; TES = tracheoesophageal speakers; H=Healty speaker; ELS = electrolarynx speakers; HNR = harmonics to noise ratio; MPT = maximum phonation time.

Jitter

Comparison of jitter values between groups were made in eight studies [27, 32, 35, 38, 42, 44, 49, 52] (n=294). One study was categorized as level A [32] and seven studies as level B [27, 35, 38, 42, 44, 49, 52]. Low jitter value is related to better voice quality [9, 20]. Lowest values for jitter are found in the group of healthy speakers (n=6 [27, 32, 38, 42, 44, 49], Figure 4). A significant difference between healthy speakers and the groups of substitute voice speakers is found in four studies, one of which is listed a level A study [32], and three as level B [27, 38, 44] (Table 2.3). For the groups of substitute voice speakers lower jitter values were found for TES compared to ES, but these outcomes were not significant (n=5 [27, 35, 42, 49, 52], Table 2.3).

Shimmer

Six studies [27, 32, 38, 44, 49, 52] (n=245) reported on shimmer values, of which one study was categorized as level A [32], five as level B [27, 38, 44, 49, 52]. Outcomes for shimmer in percentage were compared. For two studies [49, 52], shimmer outcomes were presented in shimmer decibel. For this systematic review, these data were recalculated to percentages [49, 52]. A low shimmer value is related to a better voice quality [9, 20]. In the level A study of Shim et al. [32] significantly better shimmer outcomes are reported for healthy speakers compared to TES (Table 2.3). Shimmer values for all groups ranged from 0.4% - 18.4%. The value of 53.6% shimmer for ES, is considered to be an outlier [44]. It seems that based on shimmer, healthy speech can be rated as superior compared to ES and TES (n=5 [27, 32, 38, 44, 49], Table 2.3, Figure 4). No

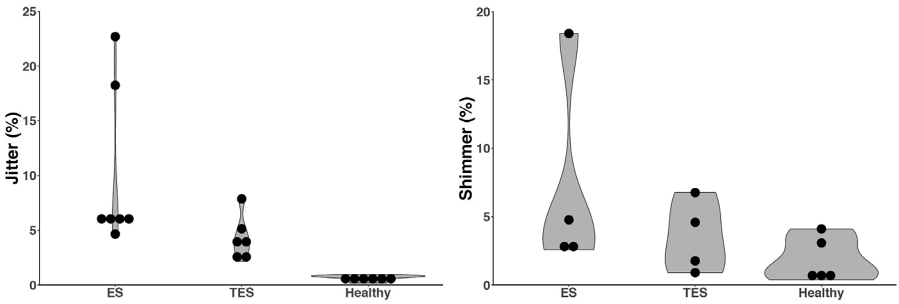

Figure 2.4: Left violin plot displaying the distribution of the mean jitter outcomes of eight studies [27, 32, 35, 38, 42, 44, 49, 52] ES n=87, TES n=102, H n=105. Right violin plot displaying the distribution of the mean shimmer outcomes of five studies ES n=65, TES n=65, H n=95 [27, 32, 38, 44, 49]. Abbreviations: ES = esophageal speakers; TES = tracheoesophageal speakers; H=Healthy Speaker; ELS = electrolarynx speakers; HNR = harmonics to noise ratio; MPT = maximum phonation time.

definite trend is seen between the speech rehabilitation methods ES and TES (n=3, Table 2.3) [27, 49, 52].

Intensity

Four level B categorized studies [34, 41, 49, 52] (n=113) reported intensity scores (Table 2.3). Intensity in decibel is not an absolute value, and therefore comparing the mean outcome is irrelevant. A higher intensity is indicated as being better [1]. Highest intensity scores are found for healthy speakers [34, 49]. Between substitute voices four studies [34, 41, 49, 52] found higher intensity scores for TES than for ES. In only one of these studies the higher intensity for TES compared to ES was significant [52].

Spectral tilt

Two studies [32, 37] (n=64) calculated the ratio between the energy above 4 kHz and the energy in the lower frequencies. One study was categorized as level A [32], the other as level B [37]. A larger ratio is correlating with better voice quality [55]. The level A rated study of Shim et al. [32] reported a larger ratio in healthy speakers than in ES. In the other study [37] the spectral tilt ratio was found to be larger for TES than for ES. It is not possible to draw overall conclusion from this due to the small number of studies reporting this outcome measure.

Table 2.4: Comparative perceptual and patient-reported outcomes for speaker groups

	Perceptual Voice quality	Perceptual Intelligibility	PROs VHI	PROs V-RQOL
ES>TES	-	-	Salturk et al. [51]* Tiple et al. [53]	-
ES>ELS	Ng et al. [48]	Williams&Watson [54]*	Salturk et al. [51]* Tiple et al. [53]	Moukarbel et al. [47]
ES > H	-	-	-	-
TES>ES	Law et al. [43] Ng et al. [48] Williams&Watson [54]*	Law et al. [43] Ng et al. [48] Williams&Watson [54]*	-	Moukarbel et al. [47]
TES>ELS	Eadie et al. [30]* Ng et al. [48] Miralles&Cervera [46] Williams&Watson [54]*	Eadie et al. [30]* Ng et al. [48] Williams&Watson [54]*	Eadie et al. [30]	Moukarbel et al. [47]*
TES>H	-	-	-	-
ELS>ES	Law et al. [43]	Law et al. [43] Ng et al. [48]	-	-
ELS>TES	Law et al. [43]	Law et al. [43]	Tiple et al. [53]	-
ELS>H	-	-	-	-
H>ES	Williams&Watson [54]*	Williams&Watson [54]*	-	-
H>TES	Finizia et al. [40] Williams&Watson [54]*	Finizia et al. [40] Williams&Watson [54]* Crosetti [31]*	-	-
H>ELS	Williams&Watson [54]*	Williams&Watson [54]*	-	-

> Indicating a better mean group outcome. Level of significance was held at $p \leq .05$. Studies presented in bold had a level A risk of bias. ES esophageal speakers, TES tracheoesophageal speakers, ELS electrolarynx speakers, MPT maximum phonation time, V-RQOL voice-related quality of life

2.3.2 Perceptual outcomes

In Table 4, comparative perceptual results for the different speaker groups are shown. Studies that reported on the primary outcomes "voice quality" and "intelligibility" are presented. Initially formulated outcome variables, which were not reported in the included studies, cannot be discussed. This concerns the percentage of voicedness, "Grade Roughness Breathiness Asthenia Strain scale assessment" (GRBAS), unattended additive noise, fluency and voicing.

Voice quality

Voice quality was perceptually evaluated in five studies [30, 40, 43, 48, 54] (n=177). One of these studies was categorized as level A [30], four level B [40, 43, 48, 54]. Across these studies, different evaluation methods were used. In the level A study by Eadie et al. [30] speech acceptability ratings were obtained for ES, TES and ELS measured by a visual analog scale (VAS). The audio recordings were evaluated by 48 listeners, and speakers judged their own speech acceptability. In another study [43] the evaluators were instructed to rate the severity of the speech impairment on an 11-point scale with equal-appearing interval, from no speech impairment to a severe speech impairment. This study made a distinction between younger and older listeners, and in the present review the results for both groups of listeners were pooled [43]. One study [40] performed perceptual evaluation of TES compared to healthy speakers on a VAS 0-100. Another study [48] used a one to seven equal-interval scale for rating. A score of one indicated severe hoarseness, while a score of seven indicated a clear voice quality. A scale of one to seven was also used in another study [54], in which evaluators rated videotaped speaking fragments of ES, TES and ELS for voice quality [54].

The level A rated study of Eadie et al. [30] found that TES are rated with a significant more acceptable voice quality compared to ELS (Table 4). When speakers had to judge their own voice recordings no differences were found in speech acceptability between TES and ELS [30]. Two other studies [46, 48] also found better voice quality outcomes for TES compared to ELS, though these differences were non-significant [46, 48]. Three studies [43, 48, 54] found a better voice quality for TES compared to ES, which was significant in one of these [54]. In one study [43], ELS was rated as having a better voice quality compared to ES and TES. Another study indicated ES as having a better voice quality compared to ELS [48]. When a comparison between substitute voice speakers and healthy speakers is made, the group of healthy speakers is rated to have the best voice quality [40, 54]. One study underlined this with significance [54].

Intelligibility

In eight studies [30, 31, 36, 40, 43, 46, 48, 54] (n=329) intelligibility was assessed, and two of these were level A studies [30, 31], and six level B studies [36, 40, 43, 46, 48, 54]. Different methods were used to measure intelligibility outcomes. Two studies [48, 54] used a self-developed seven-equal interval scale, while one study [40] used a VAS of 0-100. The VAS and interval scales were used to rate the speakers' intelligibility. Two studies [30, 43] used the Sentence Intelligibility Test (SIT) to perform a per protocol analysis. Other studies developed their own assessment procedure. One study [36] evaluated TES with a group including ES as well as ELS under different conditions. Outcomes are presented with the percentage of correct heard words [36]. Another study [31] performed a telephone listening task where words and sentences of TES and healthy speakers were transcribed. One more study [46] evaluated audio samples on intelligibility in ES and TES by phoneme confusion matrices. In this study no overall intelligibility scores were given [46].

Comparative group outcomes are shown in Table 2.4. For the speech rehabilitation methods TES is mostly rated to have a better intelligibility compared to ES and ELS. In two studies [30, 54]. including the level A rated study of Eadie et al. [30], this is confirmed with significance [30, 54]. In one study, ES is rated to have a significantly better intelligibility compared to ELS [54]. The intelligibility of healthy speakers is rated as superior compared to the speech rehabilitation methods. Two studies [31, 54], including the level A rated publication of Crosetti et al. [31], indicated that this difference was significant.

In one study [36], a comparison was made between the group of TES and a group of ES combined with ELS, and therefore, this study is not presented in Table 2.4. TES scored significant better on intelligibility when communicating with strangers in a situation where there was no face-to-face contact [36]. One study [46] found that for TES, fricative consonants had the highest number of confusions. Whereas in ES, nasals caused the highest number of confusions [46]. There was no significant difference with regards to intelligibility found between ES and TES [46].

2.3.3 Patient-reported outcome

In Table 2.4 patient-reported outcomes (PROs) for the included studies are presented. Comparative results for the different speaker groups are shown. Predefined outcome variables, which were formulated in the research question but not reported in the included studies, will not be discussed. This concerns the EORTC QLQ-H&N35, and EORTCQLQ-C30. None of the studies compared PROs on voice quality of healthy speakers with alaryngeal speakers. Therefore, only comparative outcomes of the speech rehabilitation methods can be shown.

Voice Handicap Index

The Voice Handicap Index (VHI) was used in six studies [30, 39, 41, 50, 51, 53] (n=296). Four level B studies [39, 41, 50, 53] used the full version of the VHI. Two studies, including the level A rated study of Eadie et al. [30], used the ten item version of the VHI [30, 51]. In two studies [41, 50] no mean outcomes per speaker group were reported, only number of speakers per severity level were reported. One study [39] compared TES to a group of speakers using other, non-surgical voice restoration methods (EL, ES, mouthing, and writing). In the present review, this comparison of TES with a heterogeneous group of speakers was not taken into account (TES n=35 compared to non-surgical voice restoration n=27) [39].

One study [51], reports significant better vocal functioning for ES compared to the groups of TES and ELS [51] (Table 4). In the level A rated study of Eadie et al. [30] no differences were found between TES and ELS. Although the scores of TES were slightly better, these differences were not significant (Table 4) [30]. Additionally, three level B rated studies [41, 50, 53] did not find any significant differences between groups. In one study [41] ES and TES both reported a slight or moderate perceived voice handicap. In another study [50] ES, TES and ELS reported a moderately perceived voice handicap, and no group differences were found. Furthermore, the study of Tiple et al. [53], did not find any significant differences between ES, TES and ELS. According to this study, however having no communication method available at all leads to a significantly worse VHI score compared to having ES as communication option [53].

Voice-Related Quality of Life

One level B rated study [47] (n=75) used the Voice-Related Quality of Life (V-RQOL) within the groups of ES, TES and ELS [47]. ES reported a better V-RQOL compared to ELS. TES reported a better V-RQOL than ES and ELS. The V-RQOL of TES was significantly better compared to ELS [47].

2.3.4 Summary of results

In Table 2.5 a summary of the significant differences between the speech rehabilitation is provided. Comparative studies of the three rehabilitation methods themselves show that TES is rated as superior to ES for the acoustic outcome measures F0, MPT and intensity [27, 33, 34, 52], whereas no acoustic data is available for ELS in the included studies. According to the applied perceptual evaluations, TES is rated as superior to ES and ELS, with regards to both voice quality and intelligibility. ES is superior to ELS for the perceptual outcome measure intelligibility. In PRO studies, none of the speech rehabilitation methods showed evidently better outcomes. One study reported significant better outcomes for TES compared to ES [47], but another study showed the opposite [51]. A level A rated study reports a similarly moderate degree of perceived voice handicap in TES and ELS [30].

Table 2.5: Summary of table 2.4, studies that found a significant difference between speech methods per outcome measure

	TES>ES	TES>ELS	ES>ELS	ES>TES
Fundamental frequency	Arias et al. [27] Bellandese et al. [33] Blood [34] Siric et al. [52]	—	—	—
MPT	Siric et al. [52]	—	—	—
Intensity	Siric et al. [52]	—	—	—
Perceptual voice quality	Williams&Watson [54]	Eadie et al. [30] Williams&Watson [54]	—	—
Perceptual intelligibility	Williams&Watson [54]	Eadie et al. [30] Williams&Watson [54]	Williams&Watson [54]	—
PROs	—	Moukarbel [47]	Salturk et al. [51]	Salturk et al. [51]

$>$ Indicating a better mean group outcome. Level of significance was held at $p \leq .05$. Studies presented in bold had a level A risk of bias. ES esophageal speakers, TES tracheoesophageal speakers, ELS electrolarynx speakers, MPT maximum phonation time, V-RQOL voice-related quality of life

2.4 Discussion

This systematic review underlines that the three main TL speech rehabilitation methods are acoustically and perceptually deviant from healthy speech. In PROs no comparison is made between the substitute speech rehabilitation groups and healthy speakers. Significantly better outcomes are reported for TES compared to ES for the acoustic parameters, fundamental frequency, maximum phonation time and intensity. Perceptually, TES is rated with a significantly better voice quality and intelligibility than ES and ELS. None of the speech rehabilitation groups reported evidently better outcomes in patient-reported outcomes.

These outcomes need to be interpreted with caution. Only three of the 26 included studies are rated with a low risk of bias (level A). Most outcomes, thus, stem from level B rated studies. The included studies contain small numbers of patients, and inferential statistics is not always performed. In most studies the methodology of the acoustic measurements was not specified, leading to possibly incorrect outcomes. We found several extreme outliers in F0 and shimmer, that we had to exclude because of this [41, 44, 52]. Difficulties in reliable measuring intensity values are acknowledged, no absolute values are reported but only outcomes within studies.

This systematic review shows once more that there is an urge for standardized measurement tools for evaluations of substitute voice speakers. Auditory-perceptual evaluations are often considered as being the gold standard for voice and speech evaluation. However, the great dispersion between raters has to be acknowledged. Researchers have proposed rating tools for standardized evaluations [7, 16, 18]. Nevertheless, these are not yet generally adopted. An interesting new approach is the development of automatic assessment tools, which are designed to provide objective outcomes, with some promising results recently being reported [56, 57]. Even though not all present automatic assessment tools seem suitable for analyzing substitute voices, in our opinion this is the most promising way to obtain objective voice outcomes.

The number of PRO studies that could be included in this review is limited. The EORTC QLQ-H&N35 and EORTCQLQ-C30 were defined as relevant outcome measure but not reported in the included studies. We did not find studies which specifically report the outcomes on the speech domain of this questionnaires for the different speaker groups. The VHI and V-RQOL are generally applied to evaluate vocal functioning after TL. The Communication and Participation Item Bank (CPIB) is a recently developed questionnaire [30], which is why it was not initially defined as an outcome of interest. However, the level A rated study of Eadie et al. [30] showed strong correlations between the VHI-10 and CPIB short form scores. In this study, the speech rehabilitation groups were asked to judge their own voice quality and intelligibility. These outcomes also strongly correlated with the CPIB short form scores. Therefore, the CPIB short form can be seen as a useful tool to obtain patients' opinion on vocal functioning which fits within the widely applied International Classification of Functioning (ICF) [58].

TES has favorable outcomes on the acoustic variables F0, MPT and intensity compared to ES. Both speech methods are generated within the same sound source, the PE-segment. The most likely explanation for the more favorable acoustic voice outcomes in TES is that this type of speech is pulmonary driven. It is feasible that with the pulmonary airflow, the tidal volume (roughly 5-600 ml) of TES, a more stable and better controlled airflow is created. The higher pressure could lead to controlled hypertonicity or a movement of the PE-segment to a more cranial position. This can be an explanation for the higher F0 values which are found in TES. For ES only a minimal volume of air is available, about 60-80 ml, which is roughly 2% of the lung capacity, and controlling the pressure is not really possible [1]. This limited airflow and volume lead to a shorter phonation time for ES and, presumably, to a lower F0 and lower intensity.

No limitation in publication date was applied. Several studies published in the 80's and 90's were included. During the 80's ES was known as the gold standard for speech rehabilitation, and TES was just introduced. It is likely that ES was educated fairly well in this decade. Esophageal speakers may have achieved more satisfactory outcomes in the earlier publication period than present.

We assume that the result of the speech rehabilitation efforts plays a role in self-reported outcomes. Especially ES patients often require a prolonged and intensive rehabilitation period and success is not guaranteed. Therefore, ES speakers could be more satisfied and proud of their accomplishments than TES speakers, who acquire their speech more rapidly.

For the studies included in this systematic review, it is very likely that recruitment bias exists. In most studies, recruiting and selection of participants is not described. For the acoustic and perceptual outcomes, it can be assumed that only speakers with a fairly good level of speech were included. Some authors report this bias by mentioning that they only included excellent speakers [27, 33, 34, 48, 52]. One study reported that all patients in the group of ELS failed to achieve intelligible ES and five failed in TES [51]. Also, studies reported that they had to exclude participants because of lack of speech performance or had to exclude audio files from acoustic analysis due to lack of periodicity [32, 37].

Obviously, there are more aspects influencing the acoustic, perceptual and PROs of speech rehabilitation after TL. In Table 2.6 several of these aspects related to the speech rehabilitation methods are listed. Valuable information that can explain functional outcomes after TL is missing in the reported studies. Most studies do not mention treatment details or time since TL. Furthermore, information about offered speech rehabilitation methods and aggregate practice time with the speech language pathologist is lacking. Knowledge of caretakers, and differences in health care and insurance systems play a role in the speech rehabilitation options that can be offered. Also, patients' personal factors should be taken into account when offering speech rehabilitation. Medical problems such as neurological disorders, causing a lack of dexterity or trainability, can hamper any rehabilitation technique and influence the choice. Societal participation, which includes family life and employment status also plays a role in patients' preference for a speech rehabilitation method.

Table 2.6: Aspects of the three speech rehabilitation methods with regards to required equipment, costs and dependence on healthcare system [1, 48, 53, 59–62]

	Esophageal speech (ES)	Tracheoesophageal Speech (TES)	Electrolarynx speech (ELS)
Mechanical or prosthetic device required	No	Yes	Yes
Hand occupied during voicing	No	Yes/no, some patients are able to use an automatic speaking valve	Yes
Dependence on speech language pathologist (SLP)	Yes, nowadays fewer SLP's have knowledge of providing ES therapy	Yes, knowledge of voice prosthesis equipment and TES rehabilitation is required	Yes
Duration of the therapeutic process to functional speech	Training time, mostly concerns several months	Useful speech is mostly achieved within five training sessions	Useful speech is mostly achieved within five training sessions
Financial implications	No material costs. More therapeutic costs during often prolonged training period.	Material costs, higher than ES and ELS. Potential reimbursement issues. Lower therapeutic costs that ES, comparable to ELS.	Material costs, lower than TES. Lower therapeutic costs that ES, comparable to TES.
Overall success achieving useable speech	Low success rate	High success rate	High success rate

Besides voice quality, physical capacity, emotional well-being and social functioning are also affecting general quality of life in TL patients [63, 64]. Poor general condition is negatively associated with successful voice rehabilitation [65]. Additionally, the extent of the surgery pays a role, e.g. in case of a pharyngolaryngectomy even more functional speech problems and reduced quality of life is reported [66].

2.5 Conclusions

This systematic review consists of 26 studies reporting on multidimensional voice outcomes after total laryngectomy. Only three of these studies could be rated with a low risk of bias. This number is insufficient to draw firm conclusions. Most studies were rated with a unclear risk of bias because of flaws in patient selection and methodology.

For acoustic outcomes, tracheoesophageal speech (TES) seems to be more comparable to healthy speech. Significantly better outcomes for fundamental frequency, maximum phonation time and intensity are found in TES more than in ES. TES seems to be most pleasant and comprehensible in the perceptual evaluations, followed by esophageal speech. Speaking with an electrolarynx was found to be least pleasant and comprehensible. For the PROs, all speaker groups report a degree of voice handicap. However, none of the speech rehabilitation methods were clearly indicated as achieving more satisfactory outcomes in self-reported vocal functioning.

2.5.1 Compliance with ethical standards

Disclosure of potential conflicts of interest: The Netherlands Cancer Institute receives a research grant from Atos Medical (Malmö, Sweden), which contributes to the existing infrastructure for quality of life research of the Department of Head and Neck Oncology and Surgery. The authors have no other funding, financial relationships, or conflicts of interest to disclose. *Ethical approval*: This article does not contain any studies with human participants performed by any of the authors.

References

[1] E. C. Ward and C. J. van As-Brooks, *Head and neck cancer: treatment, rehabilitation, and outcomes.* Plural Publishing, 2014.

[2] P. Farrand and F. Duncan, "Generic health-related quality of life amongst patients employing different voice restoration methods following total laryngectomy," *Psychology, Health & Medicine*, vol. 12, pp. 255–265, May 2007.

[3] P. Farrand and R. Endacott, "Speech determines quality of life following total laryngectomy: the emperors new voice?," in *Handbook of Disease Burdens and Quality of Life Measures.*, pp. 1989–2001, Springer, 2010.

[4] I. M. Verdonck-de Leeuw, R. Rinkel, and C. Leemans, "Evaluating the impact of cancer of the head and neck," in *Head and neck cancer: treatment, rehabilitation, and outcomes* (E. C. Ward and C. J. van As-Brooks, eds.), pp. 27–56, Plural Publishing, 2007.

[5] P. H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. Van De Heyning, M. Remacle, and V. Woisard, "A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques," *European Archives of Oto-Rhino-Laryngology*, vol. 258, pp. 77–82, Feb. 2001.

[6] M. Brockmann-Bauser and M. J. Drinnan, “Routine acoustic voice analysis: time to think again?,” *Current Opinion in Otolaryngology & Head and Neck Surgery*, vol. 19, pp. 165–170, June 2011.

[7] M. Moerman, J.-P. Martens, and P. Dejonckere, “Multidimensional assessment of strongly irregular voices such as in substitution voicing and spasmodic dysphonia: A compilation of own research,” *Logopedics Phoniatrics Vocology*, vol. 40, pp. 24–29, Jan. 2015.

[8] T. Most, Y. Tobin, and R. C. Mimran, “Acoustic and perceptual characteristics of esophageal and tracheoesophageal speech production,” *Journal of Communication Disorders*, vol. 33, no. 2, pp. 165–181, 2000.

[9] M. Moerman, G. Pieters, J.-P. Martens, M.-J. Van der Borgt, and P. Dejonckere, “Objective evaluation of the quality of substitution voices,” *European Archives of Oto-Rhino-Laryngology and Head & Neck*, vol. 261, no. 10, pp. 541–547, 2004.

[10] N. K. Aaronson, S. Ahmedzai, B. Bergman, M. Bullinger, A. Cull, N. J. Duez, A. Filiberti, H. Flechtner, S. B. Fleishman, J. C. de Haes, *et al.*, “The european organization for research and treatment of cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology,” *Journal of the National Cancer Institute*, vol. 85, no. 5, pp. 365–376, 1993.

[11] K. Bjordal, E. Hammerlid, M. Ahlner-Elmqvist, A. de Graeff, M. Boysen, J. F. Evensen, A. Biörklund, J. R. J. de Leeuw, P. M. Fayers, M. Jannert, T. Westin, and S. Kaasa, “Quality of life in head and neck cancer patients: validation of the european organization for research and treatment of cancer quality of life questionnaire-H&N35,” *Journal of Clinical Oncology*, vol. 17, pp. 1008–1008, Mar. 1999.

[12] T. L. Eadie, K. M. Yorkston, E. R. Klasner, B. J. Dudgeon, J. C. Deitz, C. R. Baylor, R. M. Miller, and D. Amtmann, “Measuring communicative participation: A review of self-report instruments in speech-language pathology,” *American Journal of Speech-Language Pathology*, vol. 15, pp. 307–320, Nov. 2006.

[13] B. H. Jacobson, A. Johnson, C. Grywalski, A. Silbergleit, G. Jacobson, M. S. Benninger, and C. W. Newman, “The voice handicap index (VHI): development and validation,” *American Journal of Speech-Language Pathology*, vol. 6, pp. 66–70, Aug. 1997.

[14] N. D. Hogikyan and G. Sethuraman, “Validation of an instrument to measure voice-related quality of life (V-RQOL),” *Journal of Voice*, vol. 13, pp. 557–569, Dec. 1999.

[15] R. J. Baken and R. F. Orlikoff, *Clinical measurement of speech and voice.* Singular Publishing Group, San Diego, 2000.

[16] P. H. Dejonckere, M. B. J. Moerman, J. P. Martens, J. Schoentgen, and C. Manfredi, "Voicing quantification is more relevant than period perturbation in substitution voices: an advanced acoustical study," *European Archives of Oto-Rhino-Laryngology*, vol. 269, pp. 1205–1212, Apr. 2012.

[17] C. J. van As-Brooks, F. J. Koopmans-van Beinum, L. C. W. Pols, and F. J. M. Hilgers, "Acoustic signal typing for evaluation of voice quality in tracheoesophageal speech," *Journal of Voice*, vol. 20, pp. 355–368, Sept. 2006.

[18] M. Moerman, J.-P. Martens, L. Crevier-Buchman, E. Haan, S. Grand, C. Tessier, V. Woisard, and P. Dejonckere, "The INFVo perceptual rating scale for substitution voicing: development and reliability," *European Archives of Oto-Rhino-Laryngology*, vol. 263, pp. 435–439, May 2006.

[19] M. Hirano, "Psycho-acoustic evaluation of voice: GRBAS scale," *Clinical Examination of Voice*, vol. 5, pp. 83–84, 1981.

[20] P. Boersma, "Stemmen meten met Praat," *Stem-, Spraak- en Taalpathologie*, vol. 12, no. 4, pp. 237–251, 2004.

[21] I. R. Titze, *Workshop on acoustic voice analysis: Summary statement.* National Center for Voice and Speech, 1995.

[22] J. Robbins, H. B. Fisher, E. C. Blom, and M. I. Singer, "A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production," *Journal of Speech and Hearing Disorders*, vol. 49, pp. 202–210, May 1984.

[23] M. L. Ng, H. Liu, Q. Zhao, and P. K. Lam, "Long-term average spectral characteristics of Cantonese alaryngeal speech," *Auris Nasus Larynx*, vol. 36, pp. 571–577, Oct. 2009.

[24] R. H. Pindzola and B. H. Cain, "Duration and frequency characteristics of tracheoesophageal speech," *Annals of Otology, Rhinology & Laryngology*, vol. 98, no. 12, pp. 960–964, 1989.

[25] J. P. T. Higgins and S. Green, *Cochrane handbook for systematic reviews of interventions*, vol. 5 of *Cochrane book series.* John Wiley & Sons, 2008.

[26] J. P. T. Higgins and S. Green, *Cochrane handbook for systematic reviews of interventions*, vol. 4 of *Cochrane book series.* John Wiley & Sons, 2011.

[27] M. R. Arias, J. L. Ramón, M. Campos, and J. J. Cervantes, "Acoustic analysis of the voice in phonatory fistuloplasty after total laryngectomy," *Otolaryngology—Head and Neck Surgery*, vol. 122, no. 5, pp. 743–747, 2000.

[28] M. Rosique, J. Ramon, M. Campos, and J. Cervantes, "Acoustic voice analysis in phonatory fistuloplasty after total laryngectomy," *Acta Otorrinolaringologica Española*, vol. 50, no. 2, pp. 129–133, 1999.

[29] T. L. Eadie, A. M. B. Day, D. E. Sawin, K. Lamvik, and P. C. Doyle, "Auditory-perceptual speech outcomes and quality of life after total laryngectomy," *Otolaryngology–Head and Neck Surgery*, vol. 148, pp. 82–88, Jan. 2013.

[30] T. L. Eadie, D. Otero, S. Cox, J. Johnson, C. R. Baylor, K. M. Yorkston, and P. C. Doyle, "The relationship between communicative participation and postlaryngectomy speech outcomes," *Head & Neck*, vol. 38, no. S1, pp. E1955–E1961, 2016.

[31] E. Crosetti, M. Fantini, G. Arrigoni, L. Salonia, A. Lombardo, A. Atzori, V. Panetta, A. Schindler, A. Bertolin, G. Rizzotto, and G. Succo, "Telephonic voice intelligibility after laryngeal cancer treatment: is therapeutic approach significant?," *European Archives of Oto-Rhino-Laryngology*, vol. 274, pp. 337–346, Jan. 2017.

[32] H.-J. Shim, H. R. Jang, H. B. Shin, and D.-H. Ko, "Cepstral, spectral and time-based analysis of voices of esophageal speakers," *Folia Phoniatrica et Logopaedica*, vol. 67, no. 2, pp. 90–96, 2015.

[33] M. H. Bellandese, J. W. Lerman, and H. R. Gilbert, "An acoustic analysis of excellent female esophageal, tracheoesophageal, and laryngeal speakers," *Journal of Speech, Language, and Hearing Research*, vol. 44, pp. 1315–1320, Dec. 2001.

[34] G. W. Blood, "Fundamental frequency and intensity measurements in laryngeal and alaryngeal speakers," *Journal of Communication Disorders*, vol. 17, pp. 319–324, Oct. 1984.

[35] M. Carello and M. Magnano, "A first comparative study of oesophageal and voice prosthesis speech production," *EURASIP Journal on Advances in Signal Processing*, vol. 2009, p. 6, Dec. 2009.

[36] H. De Maddalena, H. Pfrang, R. Schohe, and H.-P. Zenner, "Sprachverständlichkeit und psychosoziale anpassung bei verschiedenen stimmrehabilitationsmethoden nach laryngektomie," *Laryngo-Rhino-Otologie*, vol. 70, no. 10, pp. 562–567, 1991. doi:10.1055/s-2007-998098.

[37] F. Debruyne, P. Delaere, J. Wouters, and P. Uwents, "Acoustic analysis of tracheo-oesophageal versus oesophageal speech," *The Journal of Laryngology & Otology*, vol. 108, pp. 325–328, Apr. 1994.

[38] N. Deore, S. Datta, R. Dwivedi, R. Palav, R. Shah, S. Sayed, M. Jagde, and R. Kazi, "Acoustic analysis of tracheo-oesophageal voice in male total

laryngectomy patients," *The Annals of The Royal College of Surgeons of England*, vol. 93, pp. 523–527, Oct. 2011.

[39] E. Evans, P. Carding, and M. Drinnan, "The voice handicap index with post-laryngectomy male voices: Research report," *International Journal of Language & Communication Disorders*, vol. 44, pp. 575–586, Jan. 2009.

[40] C. Finizia, H. Dotevall, E. Lundström, and J. Lindström, "Acoustic and perceptual evaluation of voice and speech quality: A study of patients with laryngeal cancer treated with laryngectomy vs irradiation," *Archives of Otolaryngology–Head & Neck Surgery*, vol. 125, pp. 157–163, Feb. 1999.

[41] C. Granda Membiela, M. Fernández Gutiérrez, S. Mamolar Andrés, L. Santamarina Rabanal, P. Sirgo Rodríguez, and C. Álvarez Marcos, "Laryngectomized voice rehabilitation: handicap, perception and acoustic analysis," *Rev Logop Foniatr Audiol*, vol. 36, no. 3, pp. 127–134, 2016.

[42] M. Kinishi and M. Amatsu, "Pitch perturbation measures of voice production of laryngectomees after the Amatsu tracheoesophageal shunt operation," *Auris Nasus Larynx*, vol. 13, pp. 53–62, Jan. 1986.

[43] I. K.-Y. Law, E. P.-M. Ma, and E. M.-L. Yiu, "Speech intelligibility, acceptability, and communication-related quality of life in chinese alaryngeal speakers," *Archives of Otolaryngology–Head & Neck Surgery*, vol. 135, p. 704, July 2009.

[44] J. K. MacCallum, L. Cai, L. Zhou, Y. Zhang, and J. J. Jiang, "Acoustic analysis of aperiodic voice: perturbation and nonlinear dynamic properties in esophageal phonation," *Journal of Voice*, vol. 23, pp. 283–290, May 2009.

[45] J. Merol, F. Swierkosz, O. Urwald, T. Nasser, and M. Legros, "Acoustic comparison of esophageal versus tracheoesophageal speech," *Revue de Laryngologie-Otologie-Rhinologie*, vol. 120, no. 4, pp. 249–252, 1999.

[46] J. L. Miralles and T. Cervera, "Voice intelligibility in patients who have undergone laryngectomies," *Journal of Speech, Language, and Hearing Research*, vol. 38, pp. 564–571, June 1995.

[47] R. V. Moukarbel, P. C. Doyle, J. H. Yoo, J. H. Franklin, A. M. Day, and K. Fung, "Voice-related quality of life (v-rqol) outcomes in laryngectomees," *Head & Neck*, vol. 33, no. 1, pp. 31–36, 2011.

[48] M. L. Ng, C.-L. I. Kwok, and S.-F. W. Chow, "Speech performance of adult cantonese-speaking laryngectomees using different types of alaryngeal phonation," *Journal of Voice*, vol. 11, pp. 338–344, Sept. 1997.

[49] J. Robbins, H. B. Fisher, E. C. Blom, and M. I. Singer, "A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production," *Journal of Speech and Hearing Disorders*, vol. 49, pp. 202–210, May 1984.

[50] M. Rosso, L. Siric, R. Ticac, R. Starcevic, I. Segec, and N. Kraljik, "Perceptual evaluation of alaryngeal speech," *Collegium Antropologicum*, vol. 36, no. (Suppl 2), pp. 115–118, 2012.

[51] Z. Saltürk, A. Arslanoğlu, E. Özdemir, G. Yıldırım, I. Aydoğdu, T. L. Kumral, G. Berkiten, Y. Atar, and Y. Uyar, "How do voice restoration methods affect the psychological status of patients after total laryngectomy?," *Hals-Nasen-Ohren*, vol. 64, pp. 163–168, Mar. 2016.

[52] L. Širić, D. Šoš, M. Rosso, and S. Stevanović, "Objective assessment of tracheoesophageal and esophageal speech using acoustic analysis of voice," *Collegium antropologicum*, vol. 36, no. 2, pp. 111–114, 2012.

[53] C. Tiple, S. Matu, F. V. Dinescu, R. Muresan, R. Soflau, T. Drugan, M. Giurgiu, A. Stan, D. David, and M. Chirila, "Voice-related quality of life results in laryngectomies with today's speech options and expectations from the next generation of vocal assistive technologies," in *5th IEEE Conference on E-Health and Bioengineering (EHB)*, pp. 1–4, 2015. doi:10.1109/EHB.2015.7391472.

[54] S. E. Williams and J. B. Watson, "Speaking proficiency variations according to method of alaryngeal voicing," *The Laryngoscope*, vol. 97, no. 6, pp. 737–739, 1987.

[55] Y. Maryn, N. Roy, M. De Bodt, P. Van Cauwenberge, and P. Corthals, "Acoustic measurement of overall voice quality: A meta-analysis," *The Journal of the Acoustical Society of America*, vol. 126, pp. 2619–2634, Nov. 2009.

[56] R. P. Clapham, J.-P. Martens, R. J. van Son, F. J. Hilgers, M. M. van den Brekel, and C. Middag, "Computing scores of voice quality and speech intelligibility in tracheoesophageal speech for speech stimuli of varying lengths," *Computer Speech & Language*, vol. 37, pp. 1–10, 2016.

[57] Y. Maryn, P. Corthals, P. Van Cauwenberge, N. Roy, and M. De Bodt, "Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels," *Journal of Voice*, vol. 24, pp. 540–555, Sept. 2010.

[58] World Health Organization, *International classification of functioning, disability, and health: children & youth Version: ICF-CY.* World Health Organization, 2007.

[59] S. Kresić, M. Veselinović, G. Mumović, and S. M. Mitrović, "Possible factors of success in teaching esophageal speech," *Medicinski Pregled*, vol. 68, no. 1-2, pp. 5–9, 2015.

[60] M. J. McAuliffe, E. C. Ward, L. Bassett, and K. Perkins, "Functional speech outcomes after laryngectomy and pharyngolaryngectomy," *Archives of Otolaryngology–Head & Neck Surgery*, vol. 126, pp. 705–709, June 2000.

[61] S. Moon, F. Raffa, R. Ojo, M. A. Landera, D. T. Weed, Z. Sargi, and D. Lundy, "Changing trends of speech outcomes after total laryngectomy in the 21st century: A single-center study: speech rehabilitation postlaryngectomy," *The Laryngoscope*, vol. 124, pp. 2508–2512, Nov. 2014.

[62] C. G. Tang and C. F. Sinclair, "Voice restoration after total laryngectomy," *Otolaryngologic Clinics of North America*, vol. 48, pp. 687–702, Aug. 2015.

[63] A. Pereira da Silva, T. Feliciano, S. Vaz Freitas, S. Esteves, and C. Almeida e Sousa, "Quality of life in patients submitted to total laryngectomy," *Journal of Voice*, vol. 29, pp. 382–388, May 2015.

[64] M. Schuster, J. Lohscheller, P. Kummer, U. Hoppe, U. Eysholdt, and F. Rosanowski, "Quality of life in laryngectomees after prosthetic voice restoration," *Folia Phoniatrica et Logopaedica*, vol. 55, no. 5, pp. 211–219, 2003.

[65] S. Singer, M. Merbach, A. Dietz, and R. Schwarz, "Psychosocial determinants of successful voice rehabilitation after laryngectomy," *Journal of the Chinese Medical Association*, vol. 70, pp. 407–423, Oct. 2007.

[66] S. Mahalingam, R. Srinivasan, and P. Spielmann, "Quality-of-life and functional outcomes following pharyngolaryngectomy: a systematic review of literature," *Clinical Otolaryngology*, vol. 41, pp. 25–43, Feb. 2016.

CHAPTER 3

Multidimensional evaluation of voice outcomes following total laryngectomy: a prospective multicenter cohort study

Abstract

Purpose: The purpose of this study is to assess the general course of acoustic, patient rated, and clinician rated voice outcomes from pre- up to twelve months post-laryngectomy.

Methods: Patients admitted to a total laryngectomy in five participating hospitals in Australia and The Netherlands were included. Assessments took place at pre-, 3 months, 6 months, and twelve months post-surgery. Voice outcomes are evaluated with the Acoustic Voice Quality Index (AVQI), perceptual scales, and patient reported outcome measures including VHI-10 and EQ-5D-5L. Statistical analyses include descriptive statistics, t-tests (pre- to 6 months post-surgery), Linear Mixed Effect models.

Results: The study included 43 participants. A significant worsening of AVQI is seen from pre- to post-surgery evaluated with t-test ($p<.001$). The Lin-

[0]Appeared as: van Sluis, K.E., van Son, R.J.J.H., van der Molen, L., McGuinness, A.J., Palme, C.E., Novakovic, D., Stone, D., Natsis, L., Charters, E., Jones, K., Dirven, R., van den Brekel, M.W.M. "Multidimensional evaluation of voice outcomes following total laryngectomy: a prospective multicenter cohort study." *European Archives of Oto-Rhino-Laryngoly*, 1-14, 2020.

ear Mixed Effect model confirmed Time as a significant factor in predicting AVQI score ($p \leq .001$), as well as perceptual rated voice quality by the clinician (p=.015) and patient-reported perceptual rated voice quality (p=.002). No statistical significance was found in VHI-10 scores over time.

Conclusions: Successful tracheoesophageal speech was achieved in most participants, some had to rely on augmentative alternative communication methods. Patient reported outcomes indicate acceptance of the condition and sufficient coping on the long term. However, acoustic rated voice quality is abnormal at all post-surgery time-points. AVQI proved to be a useful instrument to evaluate tracheoesophageal speech. There is a need for validation and determination of cut-off values for VHI-10 and AVQI for use in tracheoesophageal speech.

3.1 Introduction

One of the most important rehabilitation goals after total laryngectomy is voice rehabilitation. To compensate for the loss of voice, patients ideally rehabilitate speech with a voice prosthesis, so called tracheoesophageal speech [1–3]. If this is not possible, alternative communication methods include esophageal speech, electrolarynx speech, or augmentative alternative communication. Successful tracheoesophageal speech after laryngectomy is not guaranteed as outcomes in intelligibility, voice quality, and experienced voice handicap varying vary strongly between laryngectomized patients.

To evaluate voice outcomes, it is recommended to use multi-dimensional analysis which combines objective and subjective outcome measures [4]. Voice recordings of connected speech and sustained phonation can be used to objectively measure voice outcomes with acoustic analysis, focusing on pitch, perturbation, and harmonics-to-noise ratio. The Acoustic Voice Quality Index (AVQI) is a widely used measure reflecting a number of acoustic outcomes in one score [5, 6]. Subjective measures, on the other hand, include clinician and patient rated perceptual evaluation of voice and speech, and Patient Reported Outcome Measures (PROMs) assessing Quality of Life (QoL) and speech related QoL.

Little is known about the course of voice outcomes in the first year after surgery [2]. Present studies, prospectively assessing the course of QoL and reported voice problems, demonstrated that, in the long term, health related QoL and speech related QoL improve post-surgery compared to pre-laryngectomy [7–12]. Before laryngectomy QoL is often lower compared to reference standard due to initial organ preservation treatment or by the tumour itself [7, 8]. Immediately after laryngectomy QoL scores drop even further. The following year after surgery some patients recover back to baseline whilst some do not recover [7, 8]. For longitudinal QoL studies there is a significant selection bias, as patients whose health problems prevent their participation through the duration of the study, are often excluded from study analysis, which may result in over optimistic QoL outcomes [7, 13–15]. Studies reporting acoustic voice outcome after total laryngectomy often compare different groups of voice restoration methods, and most report only on sufficient or even excellent speakers, potentially leading to selection bias [2, 16–18]. Despite this, several studies have demonstrated that poorer speech related QoL is associated with lower health related QoL scores after total laryngectomy [19, 20].

Prospective multidimensional evaluation over all groups of substitute-voice-speakers after laryngectomy has not yet been described in the literature. This study aims to assess the change of acoustic, patient rated, and clinician rated voice outcomes from pre- up to twelve months post-surgery. These outcomes could potentially play a role for both patients and clinicians to assist them in counselling and decision-making regarding treatment and rehabilitation.

3.2 Methods

3.2.1 Study design

A prospective multicenter design was conducted over five hospitals. Data was collected between April 2015 till May 2019 in the following institutes: The Netherlands Cancer Institute, The Netherlands, Amsterdam; Chris O'Brien Lifehouse, Sydney, Australia; Westmead Hospital, Sydney, Australia; Liverpool Hospital, Sydney, Australia; Royal Prince Alfred Hospital, Sydney, Australia. Ethical clearance was obtained for the Netherlands Cancer Institute (number IRBd18005 / N18VOQ) as well for the Australian sites (Protocol no X16-0287 & HREC/16/RPAH/367).

Patients eligible for total laryngectomy were approached to participate. Inclusion criteria were: over 18 years of age, curative intent laryngectomy, physically and cognitively able and willing to perform assessments. Informed consent was obtained from all participants. When participants during the course of the study were palliated or died, follow-up assessments were cancelled.

Data were collected at four time-points for each participant: prior to total laryngectomy (T0), three months (T1), six months (T2), and twelve months (T3) post-surgery. Study assessments included perceptual evaluation, voice recordings and patient reported outcome measures (PROMs). Data collection was performed by an experienced Speech Language Pathologist (SLP). Voice recordings included reading aloud a text, phonation of the vowel /a/ at normal pitch, as well as low, high, soft and loud.

Visual Analogue Scales (VAS) were used to perceptual rate voice quality, resulting in a score of 0 to 10, with 0 representing worst and 10 indicating the best voice quality. Perceptual scores were provided by the clinician as well as the participant, resulting in the variables Perc. Voice SLP and Perc. Voice Pt. The use of this VAS perceptual scales are derived from a dedicated perceptual rating scale for substitute voices [21].

PROMs consisted EQ-5D-5L and Voice Handicap Index 10 item version

(VHI-10). The EQ-5D-5L is a validated patient-report questionnaire that assesses a patient's current health related QoL [22]. It consists of 5 dimensions: mobility, self-care, daily activities, pain/discomfort and anxiety/depression. The final continuous outcome ranges from 0 to 1, a higher score indicates better health related QoL [22]. Scores were interpreted with the Dutch country specific reference values [23]. The VHI-10 assesses experienced voice handicap [24]. VHI-10 includes ten questions covering three sub-themes: functional, physical and emotional. The total VHI-10 continuous outcome is a score ranging from 0 to 40, a higher score indicates a greater handicap. Scores above 11 are considered as abnormal [20, 24–26]. The VHI-10 is a widely used and validated questionnaire, although not specifically validated for use after total laryngectomy.

3.2.2 Demographics and oncological history

Demographics and oncological history were collected during the first and second assessment. Demographic variables included sex and age at time of surgery. Oncological history included tumour site, T-stage, and N-stage as defined by the pathologist post-surgery, timing of (chemo)radiotherapy, and primary, functional or salvage indication for total laryngectomy. Surgery specific data included neck dissection, neopharyngeal reconstruction, tongue base resection,myotomy of the upper esophageal sphincter, primary voice prosthesis placement, and secondary voice prosthesis placement. In the finalizing phase of the study the variables were checked with information retrieved from the local data desk.

3.2.3 Acoustic analysis

Segmentation, acoustic analysis, and obtaining AVQI scores is performed using Praat [27]. The main outcome measure Acoustic Voice Quality Index (AVQI) requires recordings of a sustained vowel /a/ and a read aloud text [5, 6]. Sustained /a/ sounds of at least three seconds were used. If no single realization of 3 seconds was available, realizations were concatenated. From the read aloud text or read aloud sentences four seconds of connected speech was used. If these included long pauses, these were removed. The AVQI algorithm includes the cepstral peak prominence, harmonics-to-noise ratio, shimmer local, shimmer local dB, as well as the slope and tilt of the regression line through the long-term average spectrum. When incorporated into Praat, the analysis script estimates an AVQI score, which ranges from 0 to 10. A lower score indicates a better voice quality, >2.95 is the cut-off point, scores above are indicated as distorted. Participants who were unable to produce voice post-surgery but did perform the assessment were rated with an AVQI score of 10.

3.2.4 Statistical analysis

The data is analysed with help of IBM SPSS software to perform descriptive statistics [28] and R [29] for inferential statistics and modelling. No sample size calculation was performed since numbers of inclusion were based on the available patients admitted to TL.

Study sample characteristics were tabulated and visualized. Primary outcomes were VHI-10, AVQI, Perc. Voice SLP, and Perc. Voice Pt. Paired t-tests between T0 and T2 were performed with statistical significance level set at $p \leq .05$. To investigate treatment variables, three oncologic treatment variables were transformed to dichotomous variables, including a) primary surgical treatment vs. salvage surgical treatment, b) primary closure vs. major reconstruction of the neopharynx, and c) a History of CRT vs. RT. The variable c) History of CRT vs. RT proved to be redundant and was dropped. Definitions of the definite chosen variables are shown in Table 1 .

Table 3.1: Transformed oncological treatment factors into dichotomous variables

Variable	**Includes participants with**
Treatment	
Primary surgical treatment	Total laryngectomy as a primary cancer treatment Total laryngectomy as a treatment for a secondary primary tumour
Salvage surgical treatment	Total laryngectomy as a salvage treatment in case of recurrent disease Total laryngectomy as a treatment for a dysfunctional larynx
Reconstruction	
Primary closure Major reconstruction	Primary closure of the neopharynx Major reconstruction of the neopharynx with the use of flaps including a Pectoralis Major-flap, free flap or gastric pull up

Correlations between primary outcome measures are investigated using linear mixed effect models with (pseudo) R2 andChi-square ANOVA on $Y \sim X + (1|\text{Subject}) + (1|T)$ against $Y \sim 1 + (1|\text{Subject}) + (1|T)$. Scatter plots are made for visualization (Appendix I, II). Because of multiple testing we used Bonferroni correction and adapted alpha to $\leq .01$

To estimate the importance of the factors studied for outcomes in VHI-10, AVQI, and perceptual rated voice quality overtime, linear mixed effect models were created (Appendix III, IV, V). The model analyses the relationship between AVQI,VHI-10, and Perc. Voice SLP on the one hand and the fixed effects Time (T0, T1, T2, T3), Treatment (primary surgery vs.salvage), and

Reconstruction (Primary closure vs. Major reconstruction) on the other hand.

3.3 Results

3.3.1 Study sample

Inclusion, follow-up, and availability of data at the assessments are shown in the flow chart in Figure 1. Over all sites there was a total of 72 possible candidates who underwent total laryngectomy in the study time frame, of whom 43 were included in the study. Thirty-four from the Netherlands, nine from the Australian sites. Reasons for exclusion were: decline to participate (n=8), live out of area (n=7), missed by the clinician (n=11), total laryngectomy in combination with total glossectomy (n=1), no medical information and follow-up assessments available (n=2).

Patient characteristics are shown in Table 2. The majority of the included participants were male (n=33;77%), mean age was 64 years old at time of surgery (range 43-84). For 19 participants (44%) the total laryngectomy was the primary surgical treatment, in 24 cases 56% total laryngectomy was a salvage treatment. In 16 participants (37%) of the cases primary closure of the neopharynx was performed, major reconstruction was needed in 27 (63%) of the cases. The Australian patients (n=9) did not differ substantially from the Dutch. Within the Australian group all patients were male, for 33% (n=3) total laryngectomy was the primary treatment, 78% (n=7) had a major reconstruction of the neopharynx.

Before surgery 16 participants (37%) had a tracheostomy, which influences their communicative abilities. Although a high number of participants received a voice prosthesis, satisfactory voice rehabilitation with tracheoesophageal speech was not accomplished in all cases. Methods of communication are tabulated in Table 3. In total, 93% of the participants received a voice prosthesis, 79% (n=34) received primary puncture, 14% (n=6) secondary puncture. Seven percent (n=3) did not receive a voice prosthesis. Verbal communication with tracheoesophageal speech was documented in 17 out of 27 participants at T1, 22 out of 25 participants at T2, and 20 out of 22 participants T3.

Of the total group, 30% (n=13) died within the first year after surgery and were excluded from the analysis. Two thirds of this group (n=9) did not achieve tracheoesophageal speech, they had to depend on augmentative alternative communication. Most of this group (n=7) did receive a voice prosthesis but could not use this due to postoperative complications such fistulas, only two participants in this group did not receive a voice prosthesis due to medical issues.

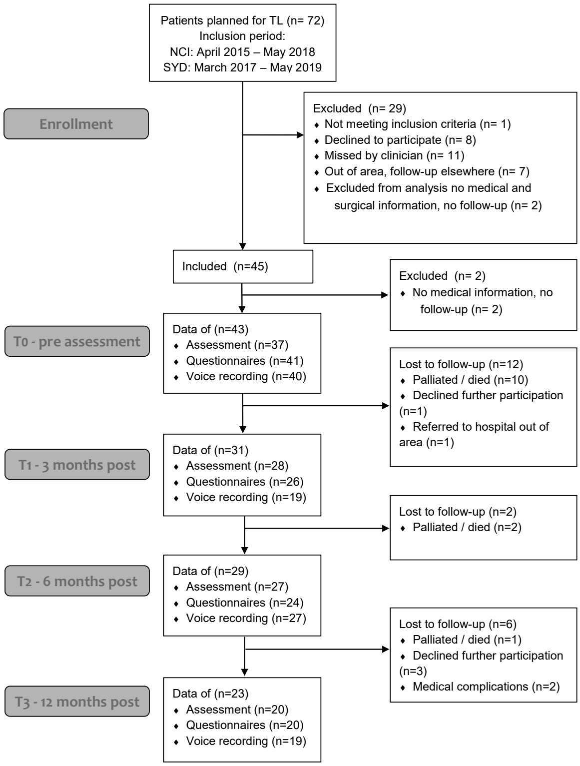

Figure 3.1: Flow-chart of study inclusion and follow-up of participants.

3.3.2 General course of self-reported outcomes and acoustic voice quality

Mean scores for EQ-5D-5L, VHI-10, AVQI, Perc. Voice SLP, and Perc. Voice Pt for the total group and the defined sub-groups for the time points T0 (pre-surgery), T1 (3 months post), T2 (6 months post), and T3 (12 months post) are shown in Table 4. Primary outcome measures show high variation, which is demonstrated by the large standard deviation provided in Table 4. Figure 2 shows the course over time, scaled 0-10, a higher score indicating better outcome. After surgery a worsening of all voice related values is seen, which gradually improves over time. Statistical significance was reached for AVQI ($p<.001$) for difference of T0 to T2, for the other outcome measures no statis-

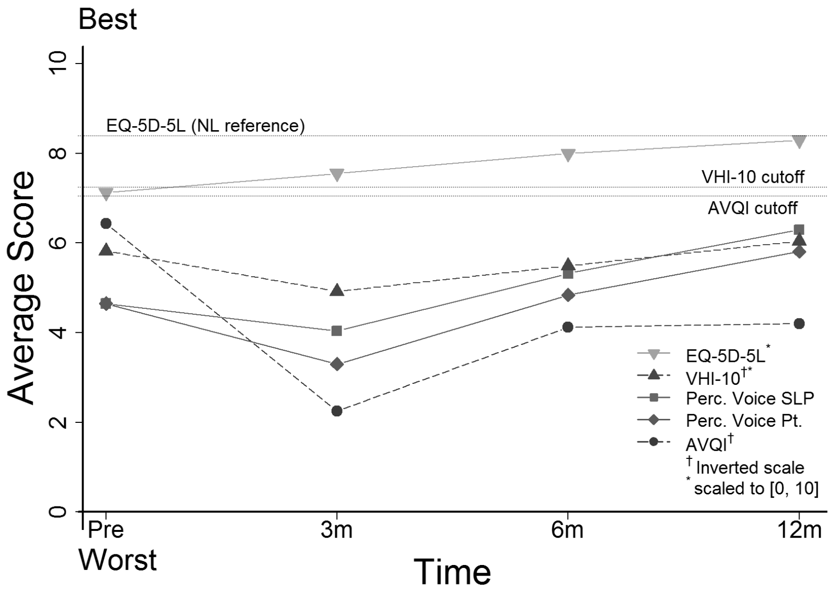

Figure 3.2: Graph visualizing mean scores for the total group for EQ-5D-5L, VHI-10, AVQI scores, Perc. Voice SLP, and Perc. Voice Pt. at each time point. EQ-5D-5L and VHI-10 scaled 0-10, AVQI and VHI-10 inverted. For easier interpretation we inverted and scaled all outcome measures 0-10.
Abbreviations: EQ-5D-5L: scores are obtained with the EQ-5D-5L and range from 0 to 1, AVQI: Acoustic Voice Quality Index (range 0-10); Perc. Voice SLP: Perceptual rated voice quality by the SLP (visual analogue scale 0-10); Perc. Voice Pt: Perceptual rated voice quality by the participant (visual analogue scale 0-10).

tically significant difference was found with paired t-test.

Worst mean EQ-5D-5L values are reported at T0, mean 0.712 (SD 0.203) (Table 4). Over time a gradually improvement of mean EQ-5D-5L values are seen. At T3 the mean EQ-5D-5L value is 0.830 (SD 0.164), which is equivalent to the reference value of 0.839 (SD 0.179) determined for the Dutch general population aged 60-70 years [23]. Before surgery 70% of the participants report a score lower than this reference value, Figure 3 shows that there is an improvement in participant reported scoring with only 48% of participant scores being outside normal limits.

Mean values for VHI-10 were at all assessment time points above 11, which is indicated as disordered [26]. Before surgery participants report a mean VHI-10 score of 16.7 (SD 10.6), this worsens at T1 with a mean score of 20.3 (SD 10.0), and over time gradually improves back to baseline level at T3 with a score of 15.8 (SD 12.0). Figure 3 visualizes the percentage of participants reporting a score above 11, which is before surgery 71%, At T2 87%, declining to 50% at

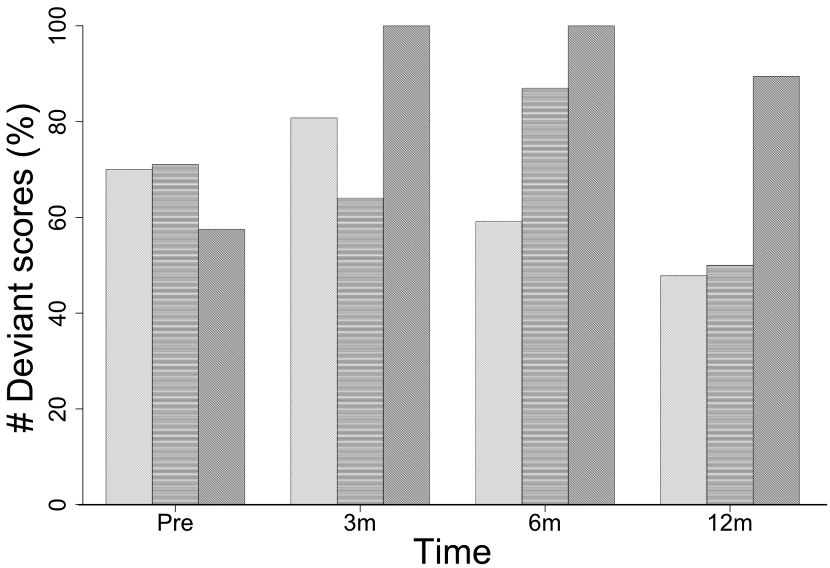

Figure 3.3: Graph visualizing participants (%) with unsatisfactory (abnormal) scores over time for EQ-5D-5L, VHI-10, and AVQI.

T3. When investigating individual course, a response shift is seen after surgery; Some participants expressed they were happy to be alive and satisfied with the fact that they can express themselves verbally, they indicate quality of the sound of the voice as less important, whilst during pre-surgery assessment their VHI-10 was clearly impaired.

The mean AVQI score rising from 3.57 (SD1.69) at T0 to 8.07 (SD 2.77) at T2 indicates a clear deterioration of acoustic voice quality after surgery (Figure 2). At all assessment time points participants voices are dysphonic, with a mean AVQI score above 2.95. AVQI scores remain impaired and never reach baseline level again. Figure 3 shows that 58% of the participants have an AVQI score indicating impairment at T0, this increases to 100% at T1 and T2. At T3, numbers are still increased to 90% of the participants.

A similar course of perceptual outcome evaluations by the clinician and the participant are found, T0 scores are: Perc. Voice SLP score 4.65 (SD 2.58), and Perc. Voice Pt score 4.65 (SD 2.58), deterioration is seen at T1, gradually improving over time. At the time-points T1-T3 mean Perc. Voice SLP scores are consistently about 0.5 points higher compared to the perceptual evaluation of the participant.

3.3.3 Correlations between outcome measurements

When outcomes for the multidimensional assessment methods assessed over time are pooled, strong correlations are found between the dimensions of voice related outcomes. Correlations between the multidimensional voice related outcomes are statistically significant (see appendix I). No statistically significant correlation is found between AVQI and EQ-5D-5L (p=.228).

Correlations between the outcome measurement instruments for the post-surgery time-points (T1-T3) show statistically significant correlations between voice related outcomes (see Appendix II). Statistical significance is lost for voice related outcomes with QoL, seen in the correlation between VHI-10 and EQ-5D-5L (p=.021), and AVQI and EQ-5D-5L (p=.467). AVQI and VHI-10 (p=.017) still correlate strongly but this is not statistically significant with statistical level set at $p \leq .01$. With investigation of the correlation between VHI-10 and AVQI for the post-surgery time points (T1-T3) an AVQI cut-off score of 6 would be indicative for a VHI-10 score $>$11, indicating an unsatisfactory voice after total laryngectomy.

3.3.4 Predictors of voice outcome

We created linear mixed effects models to explore the effect of Treatment, Reconstruction, and Time studied for the main voice outcomes VHI-10, AVQI, Perc. Voice SLP, and Perc. Voice Pt. Time is indicated as a significant factor in predicting AVQI score ($p \leq .001$), Perc. Voice SLP (p=.015), Perc. Voice Pt. (p=.002) but not for VHI-10 (p=.368). Modelling predicted outcomes for the groups for a) primary surgical treatment vs. salvage surgical treatment and b) primary closure vs. major reconstruction did not reach statistical significance.

3.4 Discussion

In our study, quality of life, measured with EQ-5D-5L, is lowest before surgery. It is known that levels of anxiety and self-care are severely impacted before as well as up to 14 days post-surgery [7, 30, 31]. We observe improvement at the three months post-surgery assessment, it is likely that patients have begun to adjust to their condition compared to 14 days post-surgery. In the long term, at 12 months post-surgery, mean score for the total group, is comparable to reference values for the age group above 60 years old [23]. This positive result might be influenced by drop-out of patients who were excluded due to mortality, nevertheless, it indicates that the remaining patients are fairly well adjusted to their condition. This general course of worsening after surgery and gradual improvement over time, corresponds to findings of earlier studies assessing the course of QoL [7, 8].

Mean values for VHI-10, were at all assessment time points above 11, which is indicated as having a voice handicap [25, 26]. This is in line with earlier studies showing patient-reported voice problems as a result of tumour presence, tracheostomy and earlier organ sparing oncologic treatment, as well as after total laryngectomy [2, 32, 33]. It is acknowledged that the VHI and VHI-10 are not specifically validated for use after total laryngectomy. A study of Moerman et al. has introduced a corrected VHI score (30 item version) specifically to use after TL, which copes with unanswered items [34]. This is useful, since not all questions apply after TL. Future studies could develop this corrected score for the VHI-10, validate the instrument for use after TL, and determine a cut-off score.

The acoustic voice outcomes, measured with AVQI, are impaired at all timepoints. However, we found a significant deterioration after total laryngectomy. Both t-test (pre- to 6 months post-surgery) and Linear Mixed Effect modelling showed statistical significance (both $p<.001$ resp.) Earlier research showed a strong correlation between AVQI and perceptual rated voice quality [35]. This study again shows strong correlation between AVQI and perceptual rated voice quality, as well as between AVQI and VHI-10, indicating that these tools measure the same construct. With the confirmation of the AVQI correlating to perceptual outcomes, as well as detecting differences over time, there is justification for AVQI use in tracheoesophageal speech samples [4, 36]. In this study, an AVQI score of ≥ 6 correlates with a VHI-10 score >11. This cut-off should be validated in a larger study.

We find a statistically significant effect of time in perceptual outcome evaluations of voice quality by the clinician as well as the participant (LME model). There is a clear deterioration in perceptual rated voice quality and intelligibility after surgery followed by a gradual improvement over 12 months.

No effect is found for the investigated oncologic treatment variables a) primary surgical treatment vs. salvage surgical treatment and b) primary closure vs. major reconstruction. It is known that oncological history of CRT negatively influences complication rates including fistula, and stricture [37, 38], but we found no influence on QoL or voice outcomes. Earlier literature showed inferior voice quality in patients with total laryngectomy who received a major reconstruction of the neopharynx [39]. Previously, Jacobi et al. also reported optimal voice characteristics in tubed flap reconstructions [40]. This shows that the voice after flap reconstruction can be comparable as after primary closure. However, we could not confirm that, the low number of patients did not allow us to look at specific reconstruction techniques.

3.4.1 Strengths and limitations

To our knowledge, this is the first study prospectively assessing a combination of acoustic, patient rated, and clinician rated voice outcomes from pre- up to twelve months post-surgery. The prospective character of the study aims to overcome a selection bias of including only excellent speakers. The unique approach with assessing acoustic, self-reported and perceptual outcomes over time provides information about the course of voice outcome and QoL. With the combination of instruments which are used, effectiveness and responsiveness of the instruments for changes over time are evaluated. By conducting this study in five hospitals in two countries, a variety of patients, languages, and treatment strategies are involved. We evaluated effects of medical detail on voice outcome, and although the number of participants of our study led to no significant results in medical history factors, this framework is useful for ongoing work.

This study has some limitations. Due to the small sample size, multiple assessments, and the variety of outcome measures we were forced to perform the LME modelling on summarized dichotomized variables. With restructuring variables into dichotomous variables information about details in the surgery are lost, e.g. Major Reconstruction is used as a summarized variable which originally included details on type and extent of (flap) reconstruction. Although all evaluation tools are widely used, they are not validated for use after TL. By conducting this study as a prospective cohort study, we aimed to overcome selection bias; nevertheless, a number of participants were not included, assessments were missed due to logistic reasons and medical complications, and participant mortality were excluded from the study. Therefore, outcomes are collected from patients who are alive and willing to fulfil study related procedures, which may lead to overestimation of the outcomes. We anticipated on evaluating different voice methods, e.g. esophageal speech and electrolarynx speech. In this cohort, however, no esophageal speakers were present and only two participants used electrolarynx speech. Therefore, no sub-group analysis between voice methods could be performed.

3.4.2 Recommendations for clinical practice and future research

Thirty per cent (n=13) of participants did not complete the study due to mortality. Sadly, nine participants did not reach acceptable (TE-)speech and had to depend on augmentative alternative communication methods such as typing, writing and mouthing in the palliative phase of their life. For clinical practice, it is recommended to inform patients about the possibility to end-up without sufficient tracheoesophageal speech, especially when prognosis is poor.

The instruments in this study have shown to be useful to detect a difference over time from pre- to one year post-surgery. Former studies that have evaluated voice outcomes after total laryngectomy utilise a wide variety of measurement tools and time points after surgery [2, 4]. AVQI, VHI-10, EQ-5D-5L, and VAS scales for perceptual ratings, used in our study, proved to be sensitive to detect differences over time from pre- to post-surgery. Sensitivity is lacking when differences between treatment groups and over time post-surgery have to be detected. Continued efforts are needed to establish the optimal tools, and validate these instruments for research and clinical practice in this population.

Improvement for patient-reported voice functioning and QoL at twelve months post-surgery was found, whilst AVQI score remains altered (Figure 3). This could be interpreted as a response shift with a change of internal standards, values, and meaning of QoL [41]. The response shift could be explained by the ability of human beings to adapt to life events. Investigating this response shift specific for the head and neck cancer group is an important issue for further research. To develop a full picture of what speech related QoL means for individuals before and after a total laryngectomy we suggest to perform studies with a combination of acoustic, patient rated, and clinician rated methods, to explore how speech related QoL is related to these measures.

3.5 Conclusion

Outcomes show that voice-related outcomes are already impaired before surgery, all worsen after surgery with a gradual improvement from six up to twelve months post-surgery. A response shift is seen in VHI-10, were acoustic measured voice quality worsen, reported voice handicap indicates acceptance of the condition and sufficient coping in the long term.

The study leads to recommendations for clinical practice; before total laryngectomy, patients should be counselled on the expected course of voice problems after surgery, with a focus on the long-term acceptable outcomes which are reached in TE-speakers. The discrepancy between reported voice handicap and objective acoustic rated voice quality, clearly demonstrates that a patient's adjustment to post-laryngectomy dysphonia does not solely rely on their acoustically measured voice quality. As such, clinicians should utilise a range of measures – both acoustic (instrumental) and patient or clinician reported, to comprehensively analyse a patient's vocal ability. Lastly, patients should be prepared for the possibility that they might not accomplish acceptable tracheoesophageal speech during their post-treatment phase, especially when medical complications occur, or oncologic treatment fails. This may be more common in the salvage procedures.

The findings of this study have implications for future research. A specific AVQI cut-off value for tracheoesophageal speech should be determined, as well as assessing the discriminative power of this instrument in this type of speech. Validation of the VHI-10 specifically for use after total laryngectomy is needed. We demonstrate a change in response of patient-reported outcomes after total laryngectomy in the relation to acoustic outcomes. Patient reported outcome measures reflect the way patients accept their condition and cope with their permanent altered speech. This is likely to vary depending on their access to support (medical, nursing and allied health, funding and equipment, support of family and friends). Future research in vocal functioning after total laryngectomy should expand beyond vocal impairment, evaluating psychosocial consequences and participation restrictions. Simultaneously, investigating the effect of medical history, including oncologic treatment factors on voice outcome, can ultimately lead to personalized pre-surgery counselling.

Acknowledgements

The authors acknowledge Atos medical AB (Malmö, Sweden) for their research grant, which contributes to the existing infrastructure for quality of life research of the Department of Head and Neck Oncology and Surgery at The Netherlands Cancer Institute. Anne Kornman, Merel Latenstein, Nadya van Gent, Katrina Moore, Jodi Gibson, Amy Freeman-Sanderson, Elizabeth Walker, Anne Taranto, and Tia Croft are acknowledged for the support they provided to implement the study procedures as well as their contribution towards the data collection.

Appendix I

Correlations between the outcome measurement instruments for all time points (T0-T3) show statistically significant correlations between AVQI and VHI-10 (p<.001), Perc. Voice SLP and VHI-10 (p<.001), Perc. Voice pt and VHI-10 (p<.001), Perc. Voice pt and AVQI (p<.001), Perc. Voice pt and Perc. Voice SLP (p<.001), Perc. Voice SLP and AVQI (p<.001) and between VHI-10 and EQ-5D-5L (p=.003). No statistically significant correlation is found between AVQI and EQ-5D-5L (p=.228).

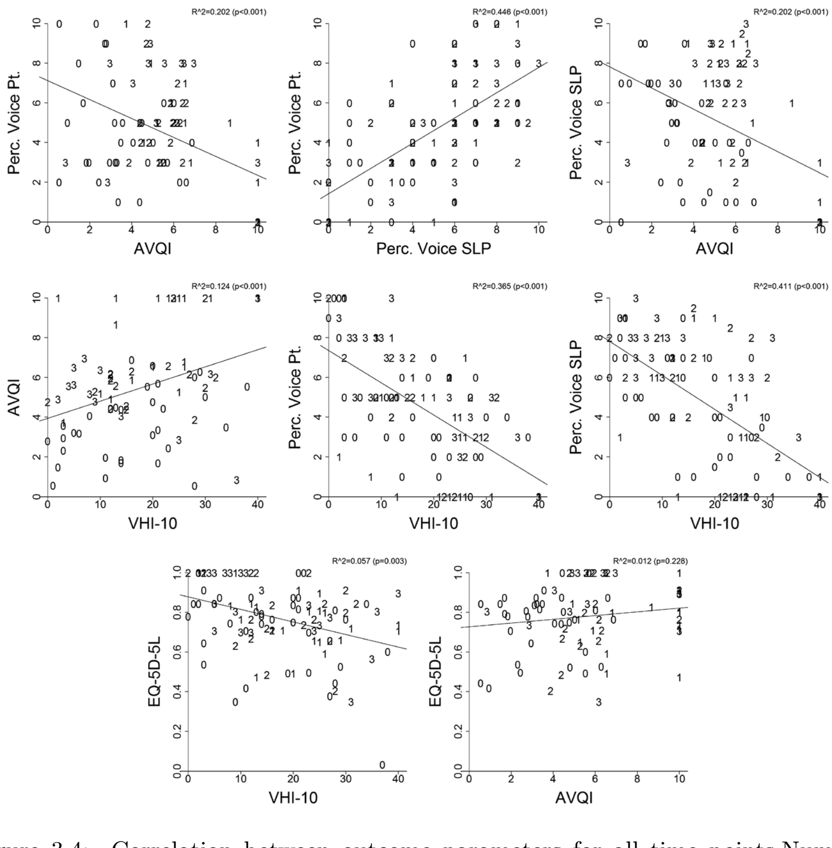

Figure 3.4: Correlation between outcome parameters for all time points.Numbers indicate T (T0-T3). Y ~ X + (1|Subject) R^2: pseudo R2 (explained variance fixed factors); p: > Chi square (ANOVA vs Y ~ 1 + (1|Subject)). Abbreviations: 0= pre-laryngectomy; 1= 3 months post-laryngectomy; 2= 6 months post-laryngectomy; 3= 12 months post-laryngectomy.

Appendix II

Correlations between the outcome measurement instruments for post-operative time points (T1-T3) show statistically significant correlations between Perc. Voice SLP and VHI-10 ($p<.001$), Perc. Voice pt and VHI-10 ($p<.001$), Perc. Voice pt and AVQI ($p<.001$), Perc. Voice pt and Perc. Voice SLP ($p<.001$), and Perc. Voice SLP and AVQI ($p<.001$). No statistically significant correlation is found between AVQI and EQ-5D-5L ($p=.228$), AVQI and VHI-10 ($p<.017$), and between VHI-10 and EQ-5D-5L ($p=.021$).

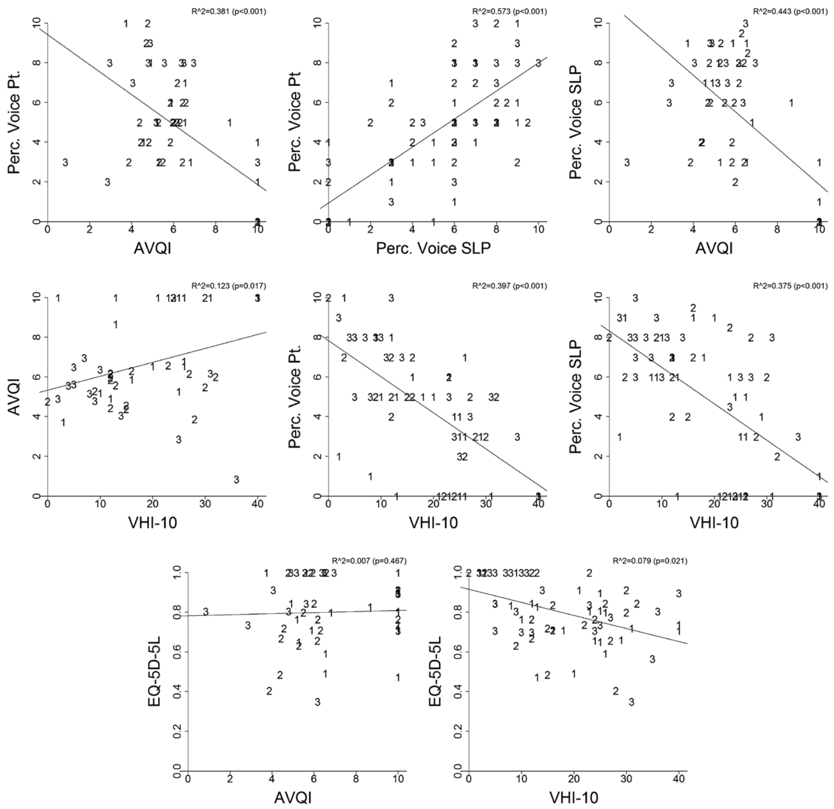

Figure 3.5: Correlation between outcome scores for the post-laryngectomy timepoints. Numbers indicate T (T1-T3). Y X + (1|Subject) R^2: pseudo R2 (explained variance fixed factors); p: > Chi square (ANOVA vs Y 1 + (1|Subject)). Abbreviations: 0= pre-laryngectomy; 1= 3 months post-laryngectomy; 2= 6 months post-laryngectomy; 3= 12 months post-laryngectomy.

Appendix III

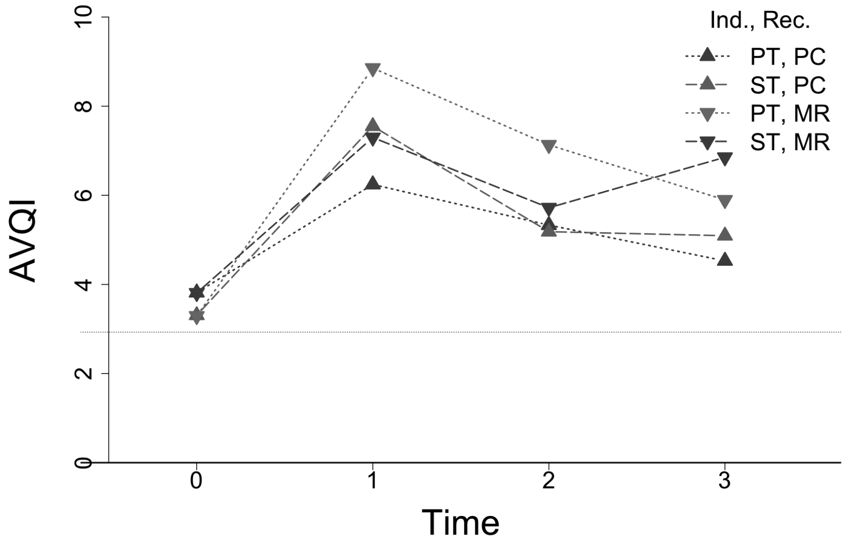

Figure 3.6: AVQI as a function of time point and treatment. AVQI Time * Ind. * Rec. Pseudo R2: 0.440 (R2m fixed factors), 0.519 (R2m complete model). ANOVA: 1 vs Time: p < 6.28e-12. Abbreviations; Ind.: Indication, Rec.: Reconstruction, PT: Primary surgical Treatment, ST: Salvage surgical Treatment, PC: Primary closure, MR: Major reconstruction. Time points: 0 = pre, 1 = 3m, 2 = 6m, 3 = 12m.

Appendix IV

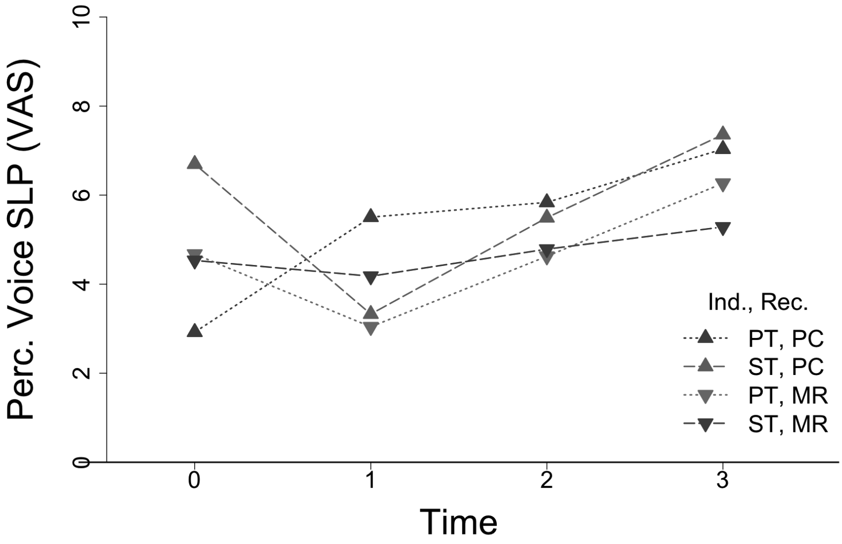

Figure 3.7: Perceptual Voice evaluation as a function of time point, reconstruction, and treatment indication. SLP Voice Time * Ind. * Rec. Pseudo R2: 0.147 (R2m fixed factors), 0.434 (R2m complete model). ANOVA: 1 vs Time: $p < 0.015$. Abbreviations; Ind.: Indication, Rec.: Reconstruction, PT: Primary surgical Treatment, ST: Salvage surgical Treatment, PC: Primary closure, MR: Major reconstruction. Time points: 0 = pre, 1 = 3m, 2 = 6m, 3 = 12m.

Appendix V

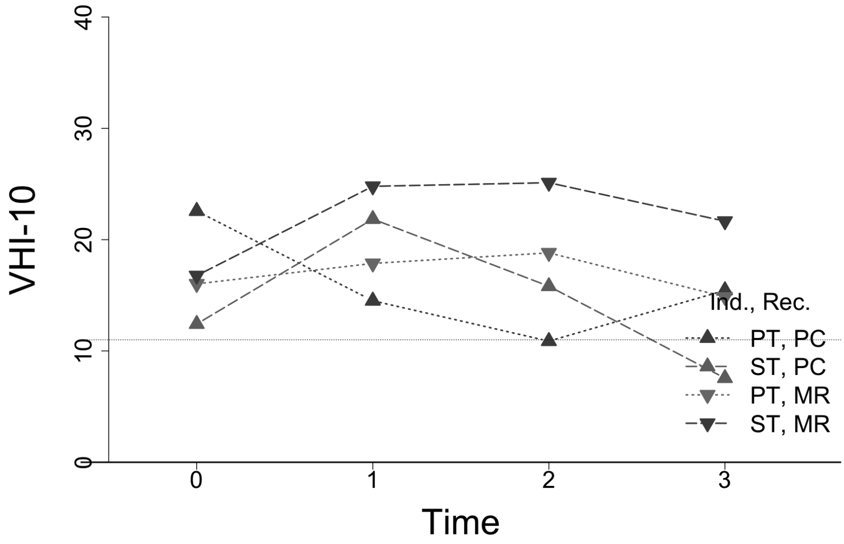

Figure 3.8: VRQOL score as a function of time point and treatment. VRQOL score Time * Ind. * Rec. Pseudo R2: 0.168 (R2m fixed factors), 0.281 (R2m complete model). ANOVA: 1 vs Time: $p < 0.368$. Abbreviations; Ind.: Indication, Rec.: Reconstruction, PT: Primary surgical Treatment, ST: Salvage surgical Treatment, PC: Primary closure, MR: Major reconstruction. Time points: 0 = pre, 1 = 3m, 2 = 6m, 3 = 12m.

References

[1] G. Sharpe, V. C. Costa, W. Doubé, J. Sita, C. McCarthy, and P. Carding, "Communication changes with laryngectomy and impact on quality of life: a review," *Quality of Life Research*, vol. 28, no. 4, pp. 863–877, 2019.

[2] K. E. van Sluis, L. van der Molen, R. J. J. H. van Son, F. J. M. Hilgers, P. A. Bhairosing, and M. W. M. van den Brekel, "Objective and subjective voice outcomes after total laryngectomy: a systematic review," *European Archives of Oto-Rhino-Laryngology*, vol. 275, no. 1, pp. 11–26, 2018.

[3] S. Singer, M. Merbach, A. Dietz, and R. Schwarz, "Psychosocial determinants of successful voice rehabilitation after laryngectomy," *Journal of the Chinese Medical Association*, vol. 70, no. 10, pp. 407–423, 2007.

[4] A. Hurren and N. Miller, "Voice outcomes post total laryngectomy," *Current Opinion in Otolaryngology & Head and Neck Surgery*, vol. 25, no. 3, pp. 205–210, 2017.

[5] B. Barsties and Y. Maryn, "External validation of the acoustic voice quality index version 03.01 with extended representativity," *Annals of Otology, Rhinology, Laryngology*, vol. 125, no. 7, pp. 571–583, 2016.

[6] Y. Maryn, M. De Bodt, B. Barsties, and N. Roy, "The value of the acoustic voice quality index as a measure of dysphonia severity in subjects speaking different languages," *European Archives of Oto-Rhino-Laryngology*, vol. 271, no. 6, pp. 1609–1619, 2014.

[7] S. Singer, H. Danker, O. Guntinas–Lichius, J. Oeken, F. Pabst, J. Schock, H. Vogel, E. F. Meister, C. Wulke, and A. Dietz, "Quality of life before and after total laryngectomy: results of a multicenter prospective cohort study," *Head & Neck*, vol. 36, no. 3, pp. 359–368, 2014.

[8] E. Armstrong, K. Isman, P. Dooley, D. Brine, N. Riley, R. Dentice, S. King, and F. Khanbhai, "An investigation into the quality of life of individuals after laryngectomy," *Head and Neck*, vol. 23, no. 1, pp. 16–24, 2001.

[9] T. D. Woodard, A. Oplatek, and G. J. Petruzzelli, "Life after total laryngectomy: a measure of long-term survival, function, and quality of life," *Archives of Otolaryngology-Head & Neck Surgery*, vol. 133, no. 6, pp. 526–532, 2007.

[10] F. W. Deleyiannis, J. Weymuller, E. A., M. D. Coltrera, and N. Futran, "Quality of life after laryngectomy: are functional disabilities important?," *Head & Neck*, vol. 21, no. 4, pp. 319–324, 1999.

[11] C. Finizia, E. Hammerlid, T. Westin, and J. Lindstrom, "Quality of life and voice in patients with laryngeal carcinoma: a posttreatment comparison of

laryngectomy (salvage surgery) versus radiotherapy," *The Laryngoscope*, vol. 108, no. 10, pp. 1566–1573, 1998.

[12] M. A. List, C. A. Ritter-Sterr, T. M. Baker, L. A. Colangelo, G. Matz, B. R. Pauloski, and J. A. Logemann, "Longitudinal assessment of quality of life in laryngeal cancer patients," *Head & Neck*, vol. 18, no. 1, pp. 1–10, 1996.

[13] K. Bjordal, M. Ahlner-Elmqvist, E. Hammerlid, M. Boysen, J. F. Evensen, A. Biörklund, M. Jannert, T. Westin, and S. Kaasa, "A prospective study of quality of life in head and neck cancer patients. Part II: Longitudinal data," *The Laryngoscope*, vol. 111, no. 8, pp. 1440–1452, 2001.

[14] E. Hammerlid, C. Mercke, M. Sullivan, and T. Westin, "A prospective quality of life study of patients with laryngeal carcinoma by tumor stage and different radiation therapy schedules," *The Laryngoscope*, vol. 108, no. 5, pp. 747–759, 1998.

[15] F. Fang, Y. Liu, Y. Tang, C. Wang, and S. Ko, "Quality of life as a survival predictor for patients with advanced head and neck carcinoma treated with radiotherapy," *Cancer*, vol. 100, no. 2, pp. 425–432, 2004.

[16] M. R. Arias, J. L. Ramón, M. Campos, and J. J. Cervantes, "Acoustic analysis of the voice in phonatory fistuloplasty after total laryngectomy," *Otolaryngology - Head and Neck Surgery*, vol. 122, no. 5, pp. 743–747, 2000.

[17] M. H. Bellandese, J. W. Lerman, and H. R. Gilbert, "An acoustic analysis of excellent female esophageal, tracheoesophageal, and laryngeal speakers," *Journal of Speech, Language, and Hearing Research*, vol. 44, no. 6, pp. 1315–1320, 2001.

[18] L. Širić, D. Šoš, M. Rosso, and S. Stevanović, "Objective assessment of tracheoesophageal and esophageal speech using acoustic analysis of voice," *Collegium Antropologicum*, vol. 36, no. 2, pp. 111–114, 2012.

[19] M. Schuster, H. Toy, J. Lohscheller, U. Eysholdt, and F. Rosanowski, "Quality of life and voice handicap of laryngectomees using tracheoesophageal substitute voice," *Laryngo-Rhino-Otologie*, vol. 84, no. 2, pp. 101–107, 2005.

[20] E. Lundström, B. Hammarberg, and E. Munck-Wikland, "Voice handicap and health-related quality of life in laryngectomees: assessments with the use of vhi and eortc questionnaires," *Folia Phoniatrica et Logopaedica*, vol. 61, no. 2, pp. 83–92, 2009.

[21] M. Moerman, J.-P. Martens, L. Crevier-Buchman, E. de Haan, S. Grand, C. Tessier, V. Woisard, and P. Dejonckere, "The infvo perceptual rating scale for substitution voicing: development and reliability," *European Archives of Oto-Rhino-Laryngology*, vol. 263, no. 5, pp. 435–439, 2006.

[22] M. van Reenen and B. Janssen, "EQ-5D-5L user guide: basic information on how to use the EQ-5D-5L instrument," *Rotterdam: EuroQol Research Foundation*, 2015.

[23] M. M. Versteegh, K. M. Vermeulen, S. M. Evers, G. A. de Wit, R. Prenger, and E. A. Stolk, "Dutch tariff for the five-level version of EQ-5D," *Value in Health*, vol. 19, no. 4, pp. 343–352, 2016.

[24] B. H. Jacobson, A. Johnson, C. Grywalski, A. Silbergleit, G. Jacobson, M. S. Benninger, and C. W. Newman, "The voice handicap index (VHI) development and validation," *American Journal of Speech-Language Pathology*, vol. 6, no. 3, pp. 66–70, 1997.

[25] C. A. Rosen, A. S. Lee, J. Osborne, T. Zullo, and T. Murry, "Development and validation of the Voice Handicap Index-10," *The Laryngoscope*, vol. 114, no. 9, pp. 1549–1556, 2004.

[26] R. E. Arffa, P. Krishna, J. Gartner-Schmidt, and C. A. Rosen, "Normative values for the Voice Handicap Index-10," *Journal of Voice*, vol. 26, no. 4, pp. 462–465, 2012.

[27] P. P. G. Boersma and D. J. M. Weenink, *Praat: Doing Phonetics by Computer: Version 3.4.* [computer program] Instituut voor Fonetische Wetenschappen, 1996.

[28] IBM Corp, "IBM SPSS statistics for Windows, version 22," *IBM Corp, Armonk*, 2013.

[29] R Core Team, "R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012," 2013. ISBN: 3-900051-07-0.

[30] C. I. F. Almonacid, A. J. Ramos, and M.-A. Rodríguez-Borrego, "Level of anxiety versus self-care in the preoperative and postoperative periods of total laryngectomy patients," *Revista Latino-Americana de Enfermagem*, vol. 24, p. e2707, 2016.

[31] S. Singer, O. Krauß, J. Keszte, G. Siegl, K. Papsdorf, E. Severi, J. Hauss, S. Briest, A. Dietz, and E. Brähler, "Predictors of emotional distress in patients with head and neck cancer," *Head & Neck*, vol. 34, no. 2, pp. 180–187, 2012.

[32] I. Jacobi, L. van der Molen, H. Huiskens, M. A. Van Rossum, and F. J. Hilgers, "Voice and speech outcomes of chemoradiation for advanced head and neck cancer: a systematic review," *European Archives of Oto-Rhino-Laryngology*, vol. 267, no. 10, pp. 1495–1505, 2010.

[33] T. L. Eadie, D. Otero, S. Cox, J. Johnson, C. R. Baylor, K. M. Yorkston, and P. Doyle, "The relationship between communicative participation and postlaryngectomy speech outcomes," *Head & Neck*, vol. 38, no. S1, pp. E1955–E1961, 2016.

[34] M. Moerman, J.-P. Martens, and P. Dejonckere, "Application of the voice handicap index in 45 patients with substitution voicing after total laryngectomy," *European Archives of Oto-Rhino-Laryngology*, vol. 261, no. 8, pp. 423–428, 2004.

[35] K. E. van Sluis, M. W. M. van den Brekel, F. J. M. Hilgers, and R. J. J. H. van Son, "Long-term stability of tracheoesophageal voices.," in *Proceedings of Interspeech 2016, San Francisco*, pp. 102–106, 2016.

[36] B. Barsties and M. De Bodt, "Assessment of voice quality: current state-of-the-art," *Auris Nasus Larynx*, vol. 42, no. 3, pp. 183–188, 2015.

[37] A. A. Forastiere, N. Ismaila, J. S. Lewin, C. A. Nathan, D. J. Adelstein, A. Eisbruch, G. Fass, S. G. Fisher, S. A. Laurie, and Q.-T. Le, "Use of larynx-preservation strategies in the treatment of laryngeal cancer: American society of clinical oncology clinical practice guideline update," *Journal of Clinical Oncology*, vol. 36, no. 11, pp. 1143–1169, 2018.

[38] D. A. Silverman, S. V. Puram, J. W. Rocco, M. O. Old, and S. Y. Kang, "Salvage laryngectomy following organ-preservation therapy–an evidence-based review," *Oral Oncology*, vol. 88, pp. 137–144, 2019.

[39] D. G. Deschler, M. W. Herr, J. R. Kmiecik, R. Sethi, and G. Bunting, "Tracheoesophageal voice after total laryngopharyngectomy reconstruction: Jejunum versus radial forearm free flap," *The Laryngoscope*, vol. 125, no. 12, pp. 2715–2721, 2015.

[40] I. Jacobi, A. J. Timmermans, F. J. M. Hilgers, and M. W. M. van den Brekel, "Voice quality and surgical detail in post-laryngectomy tracheoesophageal speakers," *European Archives of Oto-Rhino-Laryngoly*, vol. 273, no. 9, pp. 2669–2679, 2016.

[41] L. McClimans, J. Bickenbach, M. Westerman, L. Carlson, D. Wasserman, and C. Schwartz, "Philosophical perspectives on response shift," *Quality of Life Research*, vol. 22, no. 7, pp. 1871–1878, 2013.

CHAPTER 4

The acoustic contrast between the Dutch consonants /t/ and /d/ is reduced in tracheoesophageal speech

Abstract

The purpose of this study is to describe the acoustic changes of the consonants /t/ and /d/ in Dutch speaking individuals before and after total laryngectomy. The speech of seventeen participants was recorded before and after treatment. Eighteen tokens from a read-aloud text were obtained with /t/ or /d/ in initial position. Prevoicing, burst duration and duration of the vowel following the consonant were analyzed. The results show that the acoustic contrast of the /t/ and /d/ in initial position is reduced after treatment. Post-operatively, the presence of prevoicing of /d/ decreases, and burst duration increases. In the post-operative situation, therefore, /d/ becomes more similar to /t/ acoustically.

4.1 Introduction

A total laryngectomy implicates surgical removal of the larynx. This procedure is mostly carried out because of advanced laryngeal cancer. With the removal of the larynx the natural voice is lost and individuals have to regain speech

[0] Appeared as: van Sluis, K., Kapitein, M., van Son, R., Boersma, P. "The acoustic contrast between the Dutch consonants /t/ and /d/ is reduced in tracheoesophageal speech." *Proceedings of International Congress of Phonetic Sciences ICPhS 2019, 5-9 August, Melbourne, Australia*, 2019.

with the help of a substitute voice. The preferred substitute voice method in the Netherlands is tracheo-esophageal speech, with help of a voice prosthesis. Voice is generated by vibrations of the pharyngo-esophageal segment. Tracheo-esophageal speech is usually perceived as rough, irregular and reduced in dynamic range [1]. Several studies have described the acoustic characteristics of tracheo-esophageal speech [2–4]. Weak periodicity and a high noise component are characteristics of tracheo-esophageal speech [5]. Tracheo-esophageal speakers have higher values for jitter and shimmer compared to laryngeal voices [1]. This is a result of using the PE-segment as voicing source, as the tissue is less suited to sustain stable vibrations compared to the vocal folds. Jitter and shimmer will not be discussed in this paper. The focus is on fundamental frequency (F0), harmonics-to-noise ratio (HNR), percentage voiced (%V), and maximum voicing duration (MVD).

The main difficulties for tracheo-esophageal speakers regarding intelligibility are in the production of word-initial consonants, rather than consonants in word-final position [6]. Tracheo-esophageal speakers also encounter difficulties distinguishing between voiced and voiceless consonants, voicing vowels, maintaining pitch, phrasing, and producing /h/ and /k/, the consonants that are produced in areas most affected by the operation [6].

According to Jongmans et al. [6], listeners often confuse laryngectomees' initial plosives. In Dutch, word-initial voiced plosives are produced with prevoicing (a negative voice onset time), while voiceless plosives are produced without aspiration (zero voice onset time) [7]. Van Alphen [7] suggests that presence of prevoicing is the most reliable cue to voicing distinction for listeners. In the study of Jongmans et al. [6] prevoicing is not taken into account. Jongmans et al. [6]suggest that in tracheo-esophageal speech, the speaker's effort to create a voicing contrast affects the duration of the burst of the plosive as well as the duration of the following vowel. The aim of this research was to investigate how the acoustic features (prevoicing, burst duration and vowel duration) of initial plosives /t/ and /d/ change in Dutch speaking individuals before and after total laryngectomy.

4.2 Methods

4.2.1 Participants

Seventeen subjects who underwent total laryngectomy were included in the study (2 female, mean age 65, range 47-78). Participants were included in this study when a pre- and post-treatment recording was present. Subjects in palliative setting were excluded. All participants did receive speech language therapy sessions, until satisfactory speech was reached. Post-treatment recordings were made after on average 6 months after total laryngectomy (range 3-35). This study has been approved by the Institutes Research Board (registration number IRBd18005).

4.2.2 Data collection

Recordings are made during regular appointments at the speech language pathologist. Over time, three different microphones were used during consultations; HS5 Samson Aerobic Headset Mini Jack/XLR STAGE 55/CONCERT 77/AKG Version, Shure SM10A-CN headset with a Blue Icicle USB microphone preamplifier and Samson Qv10e micro-phone. Participants were instructed to pronounce a sustained /a/ as long as possible and to read aloud a Dutch text of neutral content (*Tachtig dappere fietsers* 'eighty brave cyclists'). In this text, 18 token words occurred (9 starting with /t/, 9 starting with /d/). For two participants, only the first and last two sentences of this text were recorded, since they had difficulty reading the text in its entirety resulting in 11 tokens.

4.2.3 Acoustic analysis

Speech recording data were processed in Praat [8]. Voicing characteristics before and after treatment are analyzed on sustained /a/, to check whether voice contrast analysis is possible. All /a/ recordings were cut to a length of 2.0 seconds of the most stable part (as done in van As-Brooks et al. [3]), to ensure optimized and similar data for all participants. The measured acoustic parameters are: F0 (median/SD), percentage voiced (%V), Maximum Voicing Duration (MVD), and Harmonics-to-Noise Ratio (HNR). Pitch was determined with the standard settings of the To Pitch command using the cross-correlation method with the exception of: pitch floor 40 Hz, ceiling 200 Hz; voicing threshold 0.40. The median F0 was determined as the median value of all the voiced frames. %V was determined by counting the number of voiced and unvoiced frames as determined by the Pitch algorithm. MVD was taken as the longest voiced segment after an automatic annotation with the To TextGrid (vuv): 0.2, 0.1 command. HNR was calculated adapted from van As-Brooks et al. [3] as the smoothed maximum found after calculating To Harmonicity (cc) with a minimum pitch of 40Hz, silence threshold 0, and 4.5 periods per window (1 period per window if no HNR was found). For each t/d token, the begin and end of the release burst were marked (at nearest zero-crossing). The length of the release burst was measured. The presence of an F0 value inside a 35ms window before the start of the release was considered evidence of prevoicing. A window of 35ms was chosen to preclude an influence of the preceding word. Vowel duration was calculated from the end of the release burst to the vowel offset.

4.2.4 Statistical analysis

All data is collected and processed for statistical analysis in R [9]. A Kolmogorov-Smirnov test showed the data was not normally distributed. Therefore, the Wilcoxon Matched Pairs Signed Rank test (WMPSR) is performed to analyse differences between pre- and post-total laryngectomy voice recordings. Statistical significance is corrected for false discovery rate [10]. To estimate the importance of the factors studied for the acoustic realizations of the alveolar plosives, linear mixed effect models were created [11, 12], The model analyses the relationship between prevoicing, burst duration, and vowel duration on the one hand and the fixed effects time (pre- / post-treatment) and phoneme identity (/t/ /d/) on the other hand, including the interaction term. The by-subject and by-word intercepts and random slopes of the effects of time and phoneme identity were used as random effects. P values were estimated by approximating the Student t-test statistics of the coefficients by Z-test statistics, following Barr et al. [11]. An R Markdown script and the original /t/ /d/ data have been made available on http://www.fon.hum.uva.nl/rob/ICPhS19/.

4.3 Results

In Table 4.1 the acoustic characteristics of the sustained vowel /a/ recorded from the participants before and after surgery are presented. In three participants after treatment no median F0 could be determined. After treatment acoustic features are changed. Median F0 values of tracheo-esophageal speech have a mean of 61Hz, coming from a pre-operative value of mean 138Hz. The mean percentage voiced frames drops from 93% to 55%. Harmonics-to-noise ratio (HNR) and maximum voicing duration (MVD) decreases. A statistically significant difference between pre- and post-treatment values was found for F0-median, %V, HNR and MVD ($p \leq .05$, WMPSR).

Table 4.1: Range, mean, standard deviation and statistical testing of the acoustic values pre- and post-operative of sustained vowel /a/ n=17. Abbreviations: F0: fundamental frequency, %V: percentage voiced, HNR: harmonics-to-noise ratio, MVD: maximum voicing duration. *$p \leq .05$ tested with WMPSR test ˆ (n=14)

Acoustic Parameter		**Range**	**Mean**	**SD**	***p*-value**
F_0-median	Pre	79-277	138ˆ	56	
(Hz)	Post	17-132	61	40	.002*
F_0-SD	Pre	1-27	8	10	
(Hz)	Post	1-42	11	13	.391
%V	Pre	15-100	93	21	
	Post	0-100	55	38	.003*
HNR	Pre	4-20	15	5	
(dB)	Post	0-8	3	4	<.001*
MVD	Pre	.3-2	1.9	.4	
(sec)	Post	0-2	1.2	.8	.013*

Table 4.2 shows the results with acoustic features of the consonant /t/

for the pre- and post-operative speech conditions. No statistically significant difference was seen for the parameters prevoicing and burst duration between pre- and post-operative condition. An increase in vowel duration following the initial consonant /t/ was seen for the post-operative speech condition (p=.017).

Table 4.2: Mean and standard deviation, minimum and maximum value and statistical testing of the acoustic values for the /t/ pre- and post-operative n=17. *$p \leq$.05 tested with WMPSR test

Acoustic Parameter		**Mean**	**SD**	**Min**	**Max**	***p*-value**
Prevoicing	Pre	1.3	3.7	0	11.1	
(%)	Post	1.3	5.4	0	22.2	1.00
Burst dur.	Pre	37.2	10.9	25.2	64.2	
(ms)	Post	38	5.5	29.7	47.7	.747
Vowel dur	Pre	93.9	16.6	73.6	138.1	
(ms)	Post	107.5	32	68.2	174	.017*

Table 4.3 shows the results with acoustic features of the consonant /d/ and outcomes of statistical testing for the pre- and post-operative speech conditions. The presence of prevoicing for the consonant /d/ decreases significantly in the post-operative speech condition (p=.003). Burst duration of /d/ increases significantly in the post-operative speech condition ($p<$.001). No difference was found for the vowel duration for the vowel following the initial /d/ (p=.120).

Table 4.3: Mean and standard deviation, minimum and maximum value and statistical testing of the Acoustic Values for the /d/ pre- and post-operative n=17. *$p \leq$.05 tested with WMPSR test

Acoustic Parameter		**Mean**	**SD**	**Min**	**Max**	***p*-value**
Prevoicing	Pre	57.5	36.7	0	100	
(%)	Post	14.4	23.1	0	77.8	.003*
Burst dur.	Pre	24.5	5.9	15.9	37.4	
(ms)	Post	34.9	6.5	23.6	47.3	<.001*
Vowel dur	Pre	101.5	12.6	77.3	129.9	
(ms)	Post	109.1	21.2	79.8	150.1	.120

To obtain a rough estimate of the sizes of the effects of time (pre/post-treatment) and phoneme identity (t/d) on the acoustic measurements, an analysis of linear mixed effect models was performed using maximal random factors (see Table 4.4, Figure 4.1) [11]. Models using random phoneme slopes did not converge and only random slopes for time were used. Interactions between phoneme identity and time were significant for prevoicing and burst duration.

Table 4.4: Linear mixed effect models analysis Y $\sim$ phoneme*time + (1 + time|speaker) + (1 + time|word). p determined assuming t follows Z-test (Normal) statistics. †: $t = 2.43$, $p = 0.015$

	Y: Prevoicing (%)	**Burst dur. (ms)**	**Vowel dur. (ms)**
Intercept	1.31	37.21	93.23
Phoneme	56.21	-12.66	8.3
Time	0.77	0.99	12.79†
Phoneme:time	-44.87	9.44	-4.91
t(Phoneme:time)	-8.24	3.28	-1.07
p	1.7 10-16	0.001	0.28

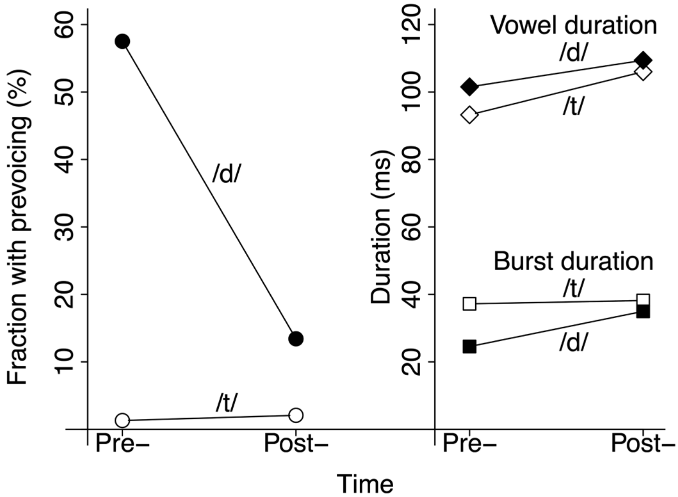

Figure 4.1: Visualization of the linear mixed effect models based on the estimates of Table 4.4

4.4 Discussion

Several studies have compared differences between healthy laryngeal speech and speech after total laryngectomy [1]. After total laryngectomy the pharyngo-esophageal segment functions as the new voice source. The resulting tracheo-esophageal speech is has a lower F0 and high noise components [1, 3]. In this study, post-treatment values of F0, %V, HNR, and MVD worsened compared to the pre-operative values. The voices of the participants included in this research are deviant from healthy laryngeal voices. In the pre-operative situation, the voices are distorted due to either tumor presence on the vocal cords, earlier treatment with (chemo) radiotherapy and/or presence of a trachea cannula. These acoustic changes in the vowel /a/ suggests an effect on acoustic features of consonants as we studied specific in the /t/ and /d/ tokens. This study has shown that the acoustic contrast of the /t/ and /d/ in word initial position reduces after total laryngectomy (Tables 4.2 and 4.3). Prevoicing and burst duration values for /t/ do not change significantly after surgery. For the /d/, presence of prevoicing decreases and Burst duration increases post-operatively. Values for /d/ become more like /t/ in the post-operative speech condition. Vowel duration following the initial consonant /t/ was higher in the post-operative condition compared to pre-operative condition (Tables 4.2 and 4.4). The duration of the vowel following the initial consonant /d/ also increased but this increase was not significant. An explanation might be an overall slowing of the speech rate. With help of linear mixed effect models, the combined effects of phoneme identity and treatment were analyzed (Table 4.4, Figure 4.1). There was a strong interaction of phoneme identity and time (pre- / post-treatment), that confirmed the conclusion that the effect of treatment was limited to the initial /d/. And thus, that the difference between /t/ and /d/ became smaller after treatment. For tracheo-esophageal speakers, the reduced acoustic contrast between /t/ and /d/ in initial position could lead to intelligibility issues. In Dutch /t/ and /d/ in phonology are marked as phonemes. Earlier research on tracheo-esophageal speech has shown that an intended /d/ was more often misheard as /t/ than that an intended /t/ was misheard as /d/ ([6], Table 4.3 confusion matrix). Our study provides evidence that prevoicing and burst duration changes for /d/ might explain at least part of these intelligibility issues.

The current study has some strengths and limitations. An advantage of the current research is that our analysis is performed on running speech. Recording running speech leads to most natural speaking conditions. Another strength of the study is the pre- and post-treatment within subject design. To our knowledge, there have not been studies that compared pre- and post-operative voice characteristics within the same individual. In this approach, changes in voice and speech can be spotted at an individual level. Limitations of the current study include the use of different microphones and that the used text was not phonetically balanced. Only /t/ and /d/ in initial position were frequently present in the text which left out other phonemes with voicing distinction for analysis. Therefore, our data did not contain enough tokens containing other plosives (/b/ and /p/; /g/ and /k/) to investigate the effect of treatment on other places of articulation. The aerodynamics of voicing and the size of the air cavity before the constriction, suggest that these effects might be different for /g/ and /k/, for /b/ and /p/ than for /t/ and /d/.

4.5 Conclusions

The acoustic features of initial consonants /t/ and /d/ do move closer together, with /d/ becoming more like /t/ in tracheo-esophageal speech. This could explain results from earlier research that showed asymmetric confusion between

/d/ and /t/ from these speakers. Further research on larger sets of running speech is recommended to create better understanding of the intelligibility issues. Total laryngectomy patients are a vulnerable group with communication deficits and need for speech rehabilitation. Special attention towards intelligibility issues is recommended.

4.6 Acknowledgements

The Netherlands Cancer Institute receives a research grant from Atos Medical (Malmö, Sweden), which contributes to the existing infrastructure for quality of life research of the Department of Head and Neck Oncology and Surgery. The authors have no other funding, financial relationships, or conflicts of interest to disclose. The authors would like to express their gratitude to the Speech Language Pathologists of The Netherlands Cancer Institute for their contribution to the data collection.

References

[1] K. E. van Sluis, L. van der Molen, R. J. van Son, F. J. Hilgers, P. A. Bhairosing, and M. W. van den Brekel, "Objective and subjective voice outcomes after total laryngectomy: a systematic review," *European Archives of Oto-Rhino-Laryngology*, pp. 1–16, 2018.

[2] T. Drugman, M. Rijckaert, C. Janssens, and M. Remacle, "Tracheoesophageal speech: A dedicated objective acoustic assessment," *Computer Speech & Language*, vol. 30, no. 1, pp. 16–31, 2015.

[3] C. J. van As-Brooks, F. J. Koopmans-van Beinum, L. C. W. Pols, and F. J. M. Hilgers, "Acoustic signal typing for evaluation of voice quality in tracheoesophageal speech," *Journal of Voice*, vol. 20, no. 3, pp. 355–368, 2006.

[4] T. Most, Y. Tobin, and R. C. Mimran, "Acoustic and perceptual characteristics of esophageal and tracheoesophageal speech production," *Journal of Communication Disorders*, vol. 33, no. 2, pp. 165–181, 2000.

[5] C. D. Van Gogh, J. M. Festen, I. M. Verdonck-de Leeuw, A. J. Parker, L. Traissac, A. D. Cheesman, and H. F. Mahieu, "Acoustical analysis of tracheoesophageal voice," *Speech Communication*, vol. 47, no. 1-2, pp. 160–168, 2005.

[6] P. Jongmans, F. Hilgers, L. Pols, and C. van As-Brooks, "The intelligibility of tracheoesophageal speech, with an emphasis on the voiced-voiceless distinction," *Logopedics Phoniatrics Vocology*, vol. 31, no. 4, pp. 172–181, 2006.

[7] P. M. Van Alphen and R. Smits, "Acoustical and perceptual analysis of the voicing distinction in dutch initial plosives: The role of prevoicing," *Journal of Phonetics*, vol. 32, no. 4, pp. 455–491, 2004.

[8] P. Boersma and D. Weenink, "Praat, doing phonetics by computer," 2017. [computer program].

[9] R Core Team, "R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012," 2015. ISBN: 3-900051-07-0.

[10] Y. Benjamini, D. Drai, G. Elmer, N. Kafkafi, and I. Golani, "Controlling the false discovery rate in behavior genetics research," *Behavioural Brain Research*, vol. 125, no. 1-2, pp. 279–284, 2001.

[11] D. J. Barr, R. Levy, C. Scheepers, and H. J. Tily, "Random effects structure for confirmatory hypothesis testing: Keep it maximal," *Journal of Memory and Language*, vol. 68, no. 3, pp. 255–278, 2013.

[12] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *arXiv preprint*, vol. arXiv:1406.5823, pp. 1–51, 2014.

CHAPTER 5

Long-term stability of tracheoesophageal voices

Abstract

Long-term voice outcomes of 13 tracheoesophageal speakers are assessed using speech samples that were recorded with at least 7 years in between. Intelligibility and voice quality are perceptually evaluated by 10 experienced speech and language pathologists. In addition, automatic speech evaluations are performed with tools from Ghent University. No significant group effect was found for changes in voice quality and intelligibility. The recordings showed a wide interspeaker variability. It is concluded that intelligibility and voice quality of tracheoesophageal voice is mostly stable over a period of 7 to 18 years.

5.1 Introduction

Total laryngectomy (TL) refers to removal of the entire larynx as a treatment for advanced stage laryngeal cancer [1]. During the surgical procedure the airway and digestive tract are separated. With removal of the larynx the natural voice is lost. Voice rehabilitation is one of the most important goals after total laryngectomy [2]. In the early 1980s, insertion of a tracheoesophageal voice prosthesis was introduced [3]. A voice prosthesis is a one-way valve that is inserted in a puncture tract created between the trachea and esophagus. Airflow

[0] Appeared as: van Sluis, K.E., van den Brekel, M.W.M., Hilgers, F.J.M., van Son, R.J.J.H. “Long-term stability of tracheoesophageal voices.” *Proceedings of Interspeech 2016, San Francisco*: 102-106, 2016.

from the lungs to the mouth is thus reestablished. Henceforth, the patient is able to produce pulmonary driven speech again. The new voice source is the pharyngoesophageal segment (PES).

Tracheoesophageal speech is considered the gold standard in restoring communicative functioning after TL [1]. It is considered as the most natural way of voice restoration according to intelligibility, pitch, and range [4]. Success rates for acquiring tracheoesophageal speech are reported up to 95% [5]. The reached endpoint in voice quality and speech intelligibility varies between patients [1]. Effective vibratory functioning of the PES is crucial in acquiring tracheoesophageal speech. Knowledge about long-term voice outcomes of tracheoesophageal speakers so far is scarce. There are some studies that include evaluations of tracheoesophageal speakers on the long-term, up to 18 years post-TL [6–10]. These papers, however, do not evaluate the groups of patients by follow-up time [6, 7, 9, 10]. Studies which consider long-term follow-up thus far only assess communication mode and quality of life [4].

In voice and speech assessment, a multidimensional approach is preferred. Acoustic, perceptual, aerodynamic, stroboscopic and self-assessment can be used to evaluate voice quality [11]. Substitute voices characteristically deviate from healthy speakers because of strong voice irregularities and require a well-thought-out approach [12]. As communication is mostly a perceptual matter, perceptual evaluations are considered the "gold standard". For substitute voices, judgments of experienced speech-language pathologists (SLP's) are considered as more consistent than judgments of naïve raters [9]. Various perceptual scales are applied in the literature to rate substitute voices. The IINFVo rating scale was specifically developed for substitute speech [13]. The five IINFVo scale parameters are: overall impression (I), impression of intelligibility (I), unintended additive noise (N), fluency (F) and voicing (Vo) [14]. Two of the rating scales, over-all impression and intelligibility, appear to be the most reliable [13–15] and are used in this study. The former refers to the acceptability or pleasantness of the voice (voice quality) and the latter refers to the clarity and understandability of words and sentences [14, 15]. During the last decade, automatic speech and voice analysis became feasible [16, 17]. Automatic analysis is promising in providing consistent ratings and for analyzing trends within a single speaker.

The present study aims to identify changes in tracheoesophageal speech over time by analyzing perceptual and automatic evaluations of voice recordings.

5.2 Speech and methods

5.2.1 Speakers and recordings

The Netherlands Cancer Institute has a long history of speech collection for TL research. Recordings from 13 TL-patients, who participated in studies between 1996 and 2014, are included in the present study (all male, median age at treatment 55 years, range 44-75, all gave informed consent). Inclusion was possible when voice recordings of the 149-word Dutch text with neutral content, "Tachtig dappere fietsers" [Eighty brave cyclists], were available from the same speaker with an interval of at least 7 years (T1 and T2, with 7-18 years in between).

All speakers had undergone laryngectomy and used a Provox voice prosthesis (Atos Medical. Hörby, Sweden). 27 recordings were made during the latter half of the 1990's (I), in 2007 (II), and in 2014 (III), in total 35 minutes of speech. No effort was made to ensure correct reading of the text so the actual words uttered vary somewhat. Speaker UCX did not complete the full text once and speaker K9S read a longer variant of the text once. For speaker KRH, there were recordings for all three periods (T1-T3). The recordings were made as part of different studies, each using different equipment (see Table 5.2). For this study, recordings were digitized and converted to 44.1 kHz sampling rate and 16-bit Signed Integer PCM encoding (RIFF/WAVE). No audio compression had been used on the recordings.

5.2.2 Perceptual evaluation

Recordings were evaluated by ten experienced SLPs (experts), including one of the authors (KvS). Experts did the evaluations at home in a self-paced online listening experiment. At the time, experts were not informed about the details of the speakers. All experts were female, mean age 29.9 year (range 22-49). Eight were native speakers of Dutch. Two were native German speakers, who acquired Dutch as a second language. All experts were certified Dutch SLP's. Evaluations were done using standard web browsers. There were two experiments. In experiment 1, the experts were asked to grade recordings of one single, long sentence as having better or worse speech intelligibility and voice quality. The experts used two slider rules as computerized visual-analog scales (VAS). In experiment 2, the same experts evaluated two pairs of short sentences from each speaker. The experts were asked to judge which version of the sentence in the pair was better and to what extent. The evaluation was again done using slider rules for speech intelligibility and voice quality. Experts could listen to the stimuli as often as they wanted. Stimuli were presented in pseudo-random order, different for each expert. The results were scored between 0-1000 (pseudo-continuous).

Table 5.1: Available patients/recordings, see text.

Period			T1	T2
I	1996-1999	[18]	8	-
II	2007	[19]	5	7
III	2014		-	6

In experiment 1, a single, 16-word, sentence was used which resulted in 26 stimuli (13 for T1 and 13 for T2). In experiment 2, two different short sentences were used, 7 and 8 words long. Each pair in Experiment 2 was presented in both orders, T1/T2 and T2/T1, and both sentences were used. For each speaker there were four pairs, two sentences in two orders. For one speaker, UCX, one sentence of the recording was missing for T2. The missing sentence was replaced by another sentence. The results of this mixed pair are omitted here. In total there were 56 stimulus pairs in experiment 2, 2x13 sentence recording pairs in two orderings and 4 additional T2/T3 stimulus pairs for speaker KRH. Both experiment 1 and 2 were preceded with 5 practice items that were drawn from other speakers not in the test set.

5.2.3 Automatic evaluation

The full 149 word recordings were automatically evaluated at the Department of Electronics and Information Systems, Ghent University with Automatic Speech analysis In Speech Therapy for Oncology (ASISTO) [20, 21]. Two applications for evaluating intelligibility were used, one using text aligned automatic speech recognition (ELIS), and one using alignment free recognition (ELISALF) [16, 20, 22]. A separate application evaluated voice quality based on the acoustic voice quality index, AVQI [16], which combines, e.g., shimmer and cepstral peak prominence. For comparability, the automatic intelligibility, 0-100 (0 worst), and AVQI, 0-8 (0 best), scores were scaled linearly to fit the perceptual evaluation results from experiment 1. No automatic evaluation was obtained for the variant readings of speakers K9S and UCX.

5.3 Results

Results of experiment 1 and the automatic evaluation scores were recalculated to pairwise differences between T2 and T1 (score at T2 minus score at T1). The four pairwise result scores of each speaker in Experiment 2 were averaged to a single preference score between [-500, 500] (after subtraction of 500). This procedure averages out any T1/T2 order bias. The averaging was done with the two remaining scores for the one speaker with a missing pair (UCX). The

Table 5.2: Recording sessions.

Period	Recorder	Microphone
1996-1999	Sony TCD-8†	AKG-c410
2007	Edirol Roland R1*	Sennheiser MD421
2014	Edirol Roland R09*	Samson Qv10e

†Digital Audio Tape (DAT) Deck. *Digital SD WAVE recorder use of different recording equipment and procedures can introduce a bias in the evaluations. A test comparing the T1 results between periods I and II and the T2 results between periods II and III (experiment 1) showed that averaged ratings between these periods were not different (Student t-test, p>.05). However, the small number of speakers makes the power of these tests low (c.f. Table 5.1). To determine for which of the 13 speakers the evaluations differed, a level of significance of $p \leq .004$ is used (Bonferroni correction). Statistical tests were performed in R [23].

5.3.1 Experiment 1

The variation in the perceptual scores in experiment 1 was high (see Figure 5.1). Only two speakers had statistically significant lower perceptual intelligibility scores for T2 than T1 ($p \leq .004$, not shown). The ELIS and ELISALF scores were strongly correlated with pooled perceptual intelligibility scores (R>.80, p<.001, n=24) and for T1 and T2 separately (R>0.78, p<.005, n=12 each). The perceptual voice quality scores differed for three speakers, two speakers had lower scores for T2 than T1 and one had higher scores ($p \leq .004$). For all other speakers, the differences were not statistically significant (p>.004). The AVQI scores were moderately correlated with pooled perceptual voice quality scores ($|R|$ >0.60, p<.002, n=24) and for T1 separately ($|R|$= 0.70, p<.02, n=12), but not for T2 ($|R|$=0.45, p>.05, n=12). Perceptual intelligibility and voice quality T2-T1 difference scores were strongly correlated (R=0.89, p<.001, n=13). ELIS and AVQI T2-T1 differences were also correlated (R=0.75, p<.01, n=11). The consistency of the evaluations was estimated by correlating the scores of individual experts against the average score of all the other experts (n=9). The correlations were between R=0.6 and R=0.9 for both Intelligibility and Voice Quality. Automatic scores were correlated with the average of all ten experts. The consistency of the scores for ELIS and ELISALF compared favorably against the intelligibility scores of individual experts: $R \geq 0.8$. Correlation of automatic AVQI scores was comparable to the least consistent expert: R=0.6.

5.3.2 Experiment 2

Eight speakers showed a statistical significant difference in intelligibility between T1 and T2 in experiment 2 and seven of them also showed a difference in voice quality (see Figure 5.2). The ELIS speaker difference scores were modestly correlated with the average pairwise perceptual ratings for intelligibility (R=0.61, $p \leq .05$). The correlation of the ELISALF difference scores with the perceptual ratings was even marginally lower (R=0.58, p>.05). The correlation between ELIS and ELISALF difference scores was statistically not significant (R=0.56, p>.05). Because of this, we focus on the ELIS scores for the remainder of this paper. Intelligibility and voice quality were strongly correlated (R=0.99, p<.001). AVQI scores were strongly correlated to voice quality and

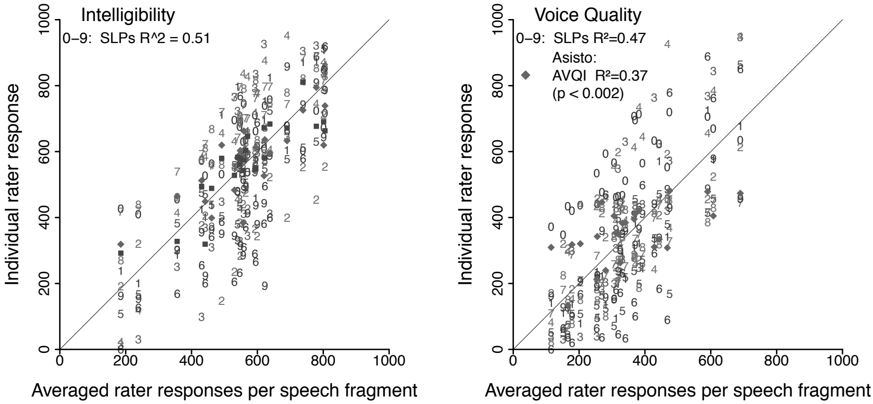

Figure 5.1: Variation in perceptual evaluations and automatic scores in experiment 1. Numbers: individual expert raters; filled symbols: automatic ASISTO scores. Correlations are with average expert responses.

thus also to intelligibility (R=0.87 and R=0.84, $p \leq .001$). This makes AVQI a better predictor of perceptual intelligibility in experiment 2 than the ELIS scores. Differences between periods I and II in Figure 5.2 were not significant for ELIS or AVQI ($p > .025$, Bonferroni correction).

Overall, five speakers had statistically significant worse intelligibility at T2 (T2-T1 $<$ 0), two speakers were better at T2 (T2-T1 $>$ 0), and four were neither better nor worse (T2-T1 0), see Figure 5.2. One speaker was scored with worse intelligibility at T2 and unchanged voice quality (WWL). In total, there are roughly as many speakers that showed a decline in intelligibility and voice quality at T2 as showed unchanged or improved intelligibility and voice quality. The ELIS scores tended to score the T2 as more intelligible than the T1 recordings. Currently, it is not clear how to interpret this difference with perceptual intelligibility scores. The AVQI scores were distributed more like the corresponding perceptual voice quality scores, in line with the high correlation between AVQI and voice quality scores. In this sample of 13 speakers, three distinct levels of change can be distinguished: better at T2, worse at T2, and no difference. When speakers from each of these levels are compared against speakers from other levels a statistical significant difference is found (p $<$ .001). Together, the automatic and perceptual results presented in Figure 5.2 indicate that there is no definite trend in the changes in intelligibility and voice quality after 7 years or more. There might be a somewhat bigger probability for a decline in intelligibility and voice quality than the reverse. However, it is clear that the differences between speakers in direction and extend of change over time are large.

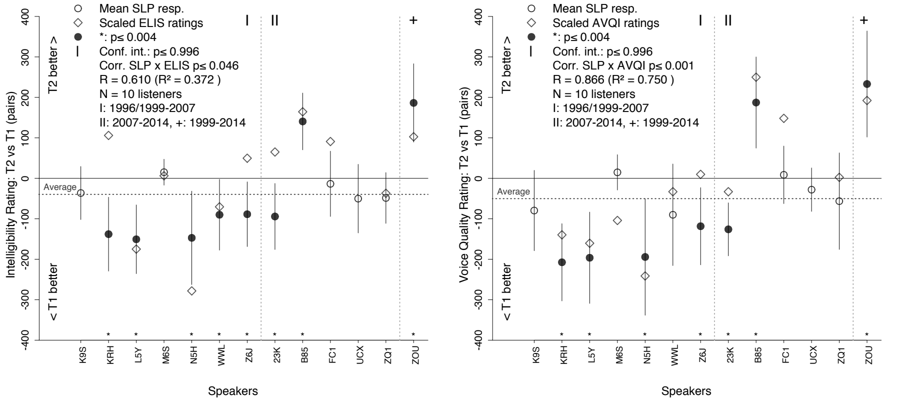

Figure 5.2: Pairwise comparisons, experts and ASISTO ratings. Left: Intelligibility and ELIS, Right: Voice quality and AVQI. Statistics based on Student t-test and Pearson's product-moment correlation.

5.3.3 Consistency between recordings

For one speaker, KRH, there were three evaluated recordings over a span of 18 years, one from each recording period. All three recordings were used to get a rough (n=1) estimate of the variability in evaluation outcomes (Table 5.3, Bonferroni correction $p \leq .01$). It appears that the experts can judge the speech samples quite consistently. Only the voice quality results for period I in experiment 1 differed from the other periods (I versus II and III, $p < .01$). None of the other evaluations differed between periods ($p > .01$). Pairwise comparisons showed significant differences in experiment 2 ($p < 0004$, underlined), except for voice quality between periods II-III. The automatic scores for this speaker, ELIS and AVQI, were rather stable over this time course (Table 5.3, Experiment 1), but the difference scores were variable (Table 5.3, Experiment 2).

5.4 Discussion

Long-term stability of voice quality and intelligibility in tracheoesophageal speakers, to our knowledge, has not yet been described. This study presents a unique dataset, in which perceptual and automatic voice assessment complement each other. It must be noted, though, that speech samples of only a small group of speakers was available. Differences in surgical techniques and treatment modalities are not included in this study because of the small sample size. The voice recordings were made in three time periods, and different audio recording equipment was used each time (Table 5.2). Our analysis did not reveal any systematic differences between time periods that could be attributed to these equipment differences. For future research, we are collecting recordings in a consistent setting.

The anatomical and physiological changes in voice production, which TL patients are facing, are immense. In tracheoesophageal speech, voice is produced by the PES that originally does not have a function in sound production. Some tracheoesophageal speakers present a fairly good voice, whilst others are rated as more deviant in voice quality and intelligibility. The differences between recordings vary. On average, a slight decrease over time is seen in perceptually rated voice quality and intelligibility (Figure 5.2). This might indicate an effect of aging.

The perceptual evaluations tend to be scattered between the expert raters (Figure 5.1). In the literature it is stated that expert raters such as SLP's provide more reliable outcomes than naïve listeners. To assess the consistency of the raters, for one speaker three recordings were evaluated. It appears that the experts can judge the speech quite consistently (Table 5.3). Using pairwise comparisons, as in experiment 2, is more sensitive to differences. Pairwise comparison results in more consistent ratings than rating individual samples, as in experiment 1.

Changes in voice quality and intelligibility are dependable within individual speakers. When voice quality is rated as good by perceptual evaluation, intelligibility tends to be as well. The strong correlation (R=0.99, p<.001 in experiment 2) between these outcome measures confirms this dependency. The fact that independent automatic measures, AVQI and ELIS, are also correlated shows that this correlation is part of the speech signal itself. These (high) correlations indicate that intelligibility problems with TE substitute voices might emerge from a lower perceptual voice quality.

The AVQI was developed for analyzing a combination of sustained vowels and running speech samples [17]. There were no sustained vowel recordings for some of our speakers. Therefore, AVQI analysis was partially performed, i.e., on running speech only. Our results show that this procedure already provided sufficient information (c.f. [20]). The AVQI scores correlate strongly with voice

Table 5.3: Results in Experiment 1 and 2 for speaker KRH. Intell. : Intelligibility, VQ: Voice Quality. *: p<.01 with other periods. __: p<.004. See text

	Experiment 1			Exp. 2	
Period	I	II	III	I-II	II-III
Intell.	801	739	731	**-138**	**-66**
ELIS	620	726	581	106	-145
VQ	*690	443	461	**-208**	-97
AVQI	474	334	409	-140	75

quality scores, and therefore also with intelligibility. Since perceptual voice quality and intelligibility are strongly correlated it is shown that for these speakers, AVQI provides consistent information on both perceived voice quality and intelligibility. The AVQI was an even better predictor of perceived intelligibility than the automatic ELIS scores. Ideally an automatic speech analysis program, which detects differences over time, is needed. The ELIS evaluation tool is used to evaluate individual speech samples. Comparisons between (T2 – T1) samples are made afterwards. For the future it would be recommended to develop an automatic assessment tool that can directly evaluate differences between speech samples.

5.5 Conclusions

Voice quality and intelligibility of tracheoesophageal speakers is more or less stable over a period of 7 to 18 years. There might be a slight decrease in the quality of the tracheoesophageal speech in some speakers, but, if at all present, this could not be consistently ascertained. Voice quality and intelligibility are correlated when rated perceptually by experts as well as when evaluated automatically. To get more insight in the long-term changes of speech quality it is recommended to systematically collect data of a large group of tracheoesophageal speakers over a longer period of time. Tools for automatic evaluation of speech quality are very promising for analyzing trends within individual speakers.

5.6 Acknowledgements

The Department of Head and Neck Oncology and Surgery of the Netherlands Cancer Institute received an unrestricted research grant from Atos Medical (Hörby, Sweden). The authors have no other funding, financial relationships, or conflicts of interest to disclose.

References

[1] E. C. Ward and C. J. van As-Brooks, *Head and neck cancer: treatment, rehabilitation, and outcomes.* Plural Publishing, 2014.

[2] P. Farrand and R. Endacott, "Speech determines quality of life following total laryngectomy: The emperors new voice?," in *Handbook of Disease Burdens and Quality of Life Measures*, pp. 1989–2001, Springer, 2010.

[3] E. D. Blom, M. I. Singer, and R. C. Hamaker, "Tracheostoma valve for postlaryngectomy voice rehabilitation," *Annals of Otology, Rhinology & Laryngology*, vol. 91, no. 6, pp. 576–578, 1982.

[4] S. Singer, M. Merbach, A. Dietz, and R. Schwarz, "Psychosocial determinants of successful voice rehabilitation after laryngectomy," *Journal of the Chinese Medical Association*, vol. 70, no. 10, pp. 407–423, 2007.

[5] B. M. R. Op de Coul, F. J. M. Hilgers, A. J. M. Balm, I. B. Tan, F. J. A. Van den Hoogen, and H. Van Tinteren, "A decade of postlaryngectomy vocal rehabilitation in 318 patients: a single institution's experience with consistent application of provox indwelling voice prostheses," *Archives of Otolaryngology–Head & Neck Surgery*, vol. 126, no. 11, pp. 1320–1328, 2000.

[6] E. Lundström and B. Hammarberg, "Speech and voice after laryngectomy: perceptual and acoustical analyses of tracheoesophageal speech related to voice handicap index," *Folia Phoniatrica et Logopaedica*, vol. 63, no. 2, pp. 98–108, 2011.

[7] M. J. McAuliffe, E. C. Ward, L. Bassett, and K. Perkins, "Functional speech outcomes after laryngectomy and pharyngolaryngectomy," *Archives of Otolaryngology–Head & Neck Surgery*, vol. 126, no. 6, pp. 705–709, 2000.

[8] W. M. Mendenhall, C. G. Morris, S. P. Stringer, R. J. Amdur, R. W. Hinerman, D. B. Villaret, and K. T. Robbins, "Voice rehabilitation after total laryngectomy and postoperative radiation therapy," *Journal of Clinical Oncology*, vol. 20, no. 10, pp. 2500–2505, 2002.

[9] C. J. van As, F. J. Koopmans-van Beinum, L. C. W. Pols, and F. J. M. Hilgers, "Perceptual evaluation of tracheoesophageal speech by naive and experienced judges through the use of semantic differential scales," *Journal of Speech, Language, and Hearing Research*, vol. 46, no. 4, pp. 947–959, 2003.

[10] M. F. Ramírez, F. G. Doménech, S. B. Durbán, M. C. Llatas, E. E. Ferriol, and R. L. Martínez, "Surgical voice restoration after total laryngectomy: long-term results," *European Archives of Oto-Rhino-Laryngology*, vol. 258, no. 9, pp. 463–466, 2001.

[11] P. H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. Van De Heyning, M. Remacle, and V. Woisard, "A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques," *European Archives of Oto-rhino-laryngology*, vol. 258, no. 2, pp. 77–82, 2001.

[12] M. Moerman, J.-P. Martens, and P. Dejonckere, "Multidimensional assessment of strongly irregular voices such as in substitution voicing and spasmodic dysphonia: A compilation of own research," *Logopedics Phoniatrics Vocology*, vol. 40, no. 1, pp. 24–29, 2015.

[13] M. B. J. Moerman, J.-P. Martens, M. J. Van der Borgt, M. Peleman, M. Gillis, and P. H. Dejonckere, "Perceptual evaluation of substitution voices: development and evaluation of the (I) INFVo rating scale," *European Archives of Oto-Rhino-Laryngology and Head & Neck*, vol. 263, no. 2, pp. 183–187, 2006.

[14] M. Moerman, G. Pieters, J.-P. Martens, M.-J. Van der Borgt, and P. Dejonckere, "Objective evaluation of the quality of substitution voices," *European Archives of Oto-Rhino-Laryngology and Head & Neck*, vol. 261, no. 10, pp. 541–547, 2004.

[15] T. Most, Y. Tobin, and R. C. Mimran, "Acoustic and perceptual characteristics of esophageal and tracheoesophageal speech production," *Journal of Communication Disorders*, vol. 33, no. 2, pp. 165–181, 2000.

[16] R. Clapham, C. Middag, F. Hilgers, J.-P. Martens, M. Van Den Brekel, and R. Van Son, "Developing automatic articulation, phonation and accent assessment techniques for speakers treated for advanced head and neck cancer," *Speech Communication*, vol. 59, pp. 44–54, 2014.

[17] Y. Maryn, M. De Bodt, and N. Roy, "The acoustic voice quality index: toward improved treatment outcomes assessment in voice disorders," *Journal of Communication Disorders*, vol. 43, no. 3, pp. 161–174, 2010.

[18] C. J. van As-Brooks, F. J. Koopmans-van Beinum, L. C. Pols, and F. J. Hilgers, "Acoustic signal typing for evaluation of voice quality in tracheoesophageal speech," *Journal of Voice*, vol. 20, no. 3, pp. 355–368, 2006.

[19] F. J. M. Hilgers, A. H. Ackerstaff, M. van Rossum, I. Jacobi, A. J. M. Balm, I. B. Tan, and M. W. M. van den Brekel, "Clinical phase I/feasibility study of the next generation indwelling Provox voice prosthesis (Provox Vega)," *Acta Oto-Laryngologica*, vol. 130, no. 4, pp. 511–519, 2010.

[20] R. P. Clapham, J.-P. Martens, R. J. J. H. van Son, F. J. M. Hilgers, M. W. M. van den Brekel, and C. Middag, "Computing scores of voice quality and speech intelligibility in tracheoesophageal speech for speech stimuli of varying lengths," *Computer Speech & Language*, vol. 37, pp. 1–10, 2016.

[21] Ghent University, "Asisto." https://asisto.elis.ugent.be/, 2016.

[22] C. Middag, R. Clapham, R. Van Son, and J.-P. Martens, "Robust automatic intelligibility assessment techniques evaluated on speakers treated for head and neck cancer," *Computer Speech & Language*, vol. 28, no. 2, pp. 467–482, 2014.

[23] R Core Team, "R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012," 2014. ISBN: 3-900051-07-0.

CHAPTER 6

Interaction of functional and participation issues on quality of life after total laryngectomy

Abstract

Objective: Total laryngectomy (TL) leads to lifelong physical changes which can lead to functional and participation issues. To assess the relationship between self-reported quality of life and functional and participation issues, a large international online questionnaire was used.

Method: A questionnaire was sent out to 8119 recipients of whom 1705 (21%) responded. The questionnaire consisted of 26 questions regarding demographic information, product use of the respondents, experienced overall health and independence, and functional and participation issues. Respondents were grouped based on sex, age, time since TL, educational level, and country of residence. Questions were grouped in one measure of reported quality of life (r-QoL) and seven issue themes ("esthetic issues", "experienced limitations in daily activities", "avoiding social activities", "communication issues", "experienced vulnerability due to environmental factors", "pulmonary issues", and "sleep issues") to assess the underlying relations.

[0]Appeared as: Leemans, M, van Sluis, K.E., van Son, R.J.J.H., van den Brekel, M.W.M. "Interaction of functional and participation issues on quality of life after total laryngectomy." *Laryngoscope Investigative Otolaryngology* 5,3: 453-460, 2020.

Results: This study showed that more functional and participation issues and a lower r-QoL are reported in the group of younger respondents (<60 years), women, and respondents who have had the TL procedure less than 2 years ago. The issue themes "experienced limitations in daily activities" and "avoiding social activities" are related to r-QoL. Most participants report "pulmonary issues," and these issues have a strong correlation with most other themes.

Conclusion: The ability to participate in meaningful and social activities is a major factor in r-QoL. Due to the frequency and strong correlations of pulmonary issues with other issue themes, pulmonary issues might be an underlying cause of many other issues.

6.1 Introduction

A total laryngectomy (TL) leads to lifelong changes in physical, psychological and social functioning, severely impacting the patients' normal daily functioning and quality of life. First of all, due to the removal of the larynx, one of the immediate consequences is that the patient loses his ability of laryngeal speech. The fact that patients now breathe through a tracheostoma instead of their nose or mouth can lead to pulmonary problems, such as coughing, dyspnea, excessive mucus production, forced expectoration, and stoma cleaning [1–6]. A total laryngectomy can also lead to other functional problems such as difficulties in swallowing, olfaction, sleeping difficulties, fatigue, and pain in throat and neck [4, 6, 7].

The functional consequences of the TL procedure also impact the patient's self-image, which can lead to psychological problems such as anxiety and depression [4, 5, 8, 9], and reduced sexuality [7, 10–12]. Lastly, due to the visible scarring and difficulties with communication it can lead to worsened social interactions and participation [4, 5, 7, 13].

To help total laryngectomized patients rehabilitate their lost functions, in particular their pulmonary condition and speech, multiple medical devices have been developed in the last few decades. The most important medical devices include the voice prosthesis, and the Heat and Moisture Exchanger (HME) with various fixation methods. It has been proven that HMEs reduce pulmonary issues, such as coughing and mucus production, and improve other related functional issues such as voice and sleeping [14–18]. To restore speech and voice, tracheoesophageal speech generated with the help of a voice prosthesis has become the 'gold standard' in the many parts of the world [19].

To ensure an optimal rehabilitation and quality of life for each patient, an understanding of possible consequences of the TL procedure and the correlation of self-reported participation and functional issues can provide a unique insight. Atos Medical AB (Malmö, Sweden), with help of ReD Associates (Copenhagen, Denmark), a strategy consultancy company, recently send out a questionnaire to Atos Medical clients from different countries to investigate the hidden needs and complaints of total laryngectomized patients. We were able to use the obtained responses for our analysis. Thus, with the aforementioned physical and social consequences of the TL procedure in mind, we developed the following research question: What relation exists between demographic characteristics and reported Quality of Life (r-QoL) ratings of total laryngectomized individuals and their self-reported participation and functional issues?

6.2 Materials and methods

6.2.1 Questionnaire

An online questionnaire was developed by Atos Medical AB (Malmö, Sweden) and ReD Associates (Copenhagen, Denmark) with input from the Netherlands Cancer Institute. The scope of the questionnaire was to assess the impact of TL on daily life and examine the use of medical devices, experienced functional and participation issues, and possible hidden needs regarding medical devices for the rehabilitation after TL. A pilot version of the questionnaire was send via email by Atos Medical AB to a cohort of 250 TL clients in the United States of America, with a response rate of 12%. Based on the pilot, adjustments were made. The final version of the online questionnaire consisted of 26 main questions regarding demographic information (not obligatory) and product use of the respondents, experienced overall health and independence, and experienced functional and participation issues. The final questionnaire was sent out via email by Atos Medical AB to 8119 clients in nine countries; the United Kingdom, United States, Germany, France, Sweden, the Netherlands, Brazil, Italy, and Spain. All approached Atos Medical clients were treated with TL and older than 18 years. One reminder-email was sent out and the questionnaire was available online for three weeks. The data were collected by ReD Associates and made available to the Netherlands Cancer Institute.

6.2.2 Statistical analysis

The responses of the questionnaire were analyzed anonymously by the Netherlands Cancer Institute using the statistical package *R* (version 3.5.1.). Respondents from the pilot study (n=29) were excluded from the analysis, resulting in a cohort of n=1705. As primary outcome measure the relation between the reported quality of life (r-QoL) ratings and the reported participation and functional issues was tested. As secondary outcome measure the relations between demographics and participation and functional issues were tested. A linear model was selected with the "step" function (setting direction “both”) in R using the Bayesian Information Criterion [20]. The relative importance of the themes was determined with the “calc.relimp” function (setting type “first”) [21].

6.2.3 Grouping of respondents

For the analysis of the primary and secondary outcomes, the respondents were grouped. The grouping was based on:

- Sex: male vs. female
- Age: <60 years of age vs. 60+ years of age. In the questionnaire, the respondents were asked to indicate their age through a choice between five decade age brackets. The age brackets were simplified in our analyses to just two age groups, roughly representing the 'pre-retirement' age group and 'post-retirement' age group.
- Time since TL: <2 years since TL vs. 2+ years since TL. This grouping was based on clinical experts consensus that the most initial rehabilitation issues of the TL procedure are resolved within two years.
- Educational level: respondents without tertiary education (defined as an educational degree after High School) vs. respondents with tertiary education.
- Country of residence: country specific analyses were only performed for countries with a response rate of >5%. The countries Sweden, Brazil and Spain were therefore excluded in this specific analysis.

6.2.4 Grouping of questions

Because this questionnaire was not based on validated QoL scales or validated questionnaires, we performed a clustering of semantically related questions into general issue themes. Because some questions could belong to multiple themes, the semantic clustering of questions was based on the experience of the clinical experts (K. E. v S. and M. vd B.), discussed in multiple consensus meetings.

The self-reported ratings of overall health and independence (both rating scales from 0 to 10) were combined to one sum measure: the reported r-QoL rating (scale from 0 to 20). In this article, the term r-QoL is used to refer to this combined measure of the following two specific scale questions:

- "How would you rate your overall health from 0 to 10? 0 means worst imaginable health state, 10 means best imaginable health state"
- "On a scale from 0 to 10, how independent do you feel in completing the activities you want? 0 means the least imaginable independence in completing the activities you want, 10 means the most imaginable independence in completing the activities you want"

The internal correlation between these two scale questions is textitR^2=.366 (percentage of variance explained).

Clustering of related self-reported participation and functional issues questions was performed into the following general issue themes (see Appendix A for an overview of the grouped questionnaire questions per theme):

- Esthetic issues (5 yes-no questions)
- Experienced limitations in daily activities (9 yes-no questions)
- Avoiding social activities (2 yes-no questions)
- Communication issues (4 yes-no questions)
- Experienced vulnerability due to environmental factors (7 yes-no questions)
- Pulmonary issues (14 yes-no questions)
- Sleep issues (3 yes-no questions)

6.2.5 Average number of reported experienced issues per theme

The average number of reported issues per patient per theme (as a percentage of the maximum number of questions of that theme), and the influence of the grouping of respondents on the number of reported experienced issues is presented. This approach was chosen in order to make the themes mutually comparable on the basis of seriousness, since the average weight per issue is not linear and the number and content of issue questions per theme were different.

Table 6.1: Demographic characteristics of respondents. Total number of respondents n=1705. Respondents of which we obtained the complete demographic information n=1624 (including sex, age and employment status)

Characteristic		No.	(%)
Countries	United Kingdom	159	(9)
	United States	785	(46)
	Germany	98	(6)
	France	342	(20)
	The Netherlands	184	(11)
	Italy	79	(5)
	Sweden	43	(3)
	Brazil	10	(<1)
	Spain	4	(<1)
Sex†	Male	1361	(80)
	Female	263	(15)
Age [y]†	<60	296	(17)
	60+	1328	(78)
Time since TL [y]	<2	336	(20)
	2-5	568	(33)
	5-10	392	(23)
	> 10	294	(17)
	No answer	115	(7)
Employment status†	Retired	1124	(66)
	Full-time	183	(11)
	Part-time	108	(6)
	occasionally	0	(0)
	Unpaid work	62	(4)
	Seeking work	33	(2)
Education	No High School	113	(7)
	High School	508	(30)
	Occupational	411	(24)
	University	617	(36)
	No answer	56	(3)
Level of tertiary education (defined as an educational degree after High School)	Total	1028	(62)
	United Kingdom	105	(67)
	United States	519	(68)
	Germany	50	(54)
	France	215	(67)
	The Netherlands	76	(43)

†A small number of respondents (n=81) did not consent to provide this (personal) demographical information. TL = total laryngectomy

6.3 Results

6.3.1 Respondents

The demographic characteristics of the respondents are shown in Table 6.1. In total, 1705 clients completed the questionnaire (response rate of 21%), of which the majority were from United States, France, The Netherlands, United Kingdom, Germany and Italy. A minority of the respondents were from Sweden, Brazil and Spain. Of 1624 respondents, we obtained the complete demographic information, including the more personal information, such as sex, age, and employment status (n=81 respondents did not give consent to provide this information). Type of voice rehabilitation was not in the scope of this questionnaire. Most respondents are male (male-to-female ratio 5:1) with a median age in the 60-69 bracket, and have had their TL procedure in the last five years (median 5 years ago). The distribution of age between countries is very comparable. The education level of the respondents, however, varies between countries, with 68% of respondents having tertiary education in the US, versus only 43% of respondents in The Netherlands and Italy.

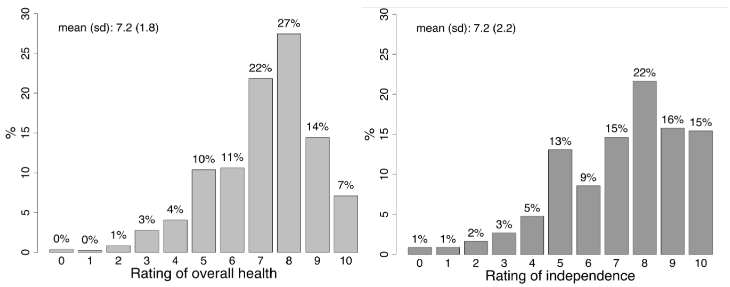

Figure 6.1: Distribution of the self-reported rating of overall health and independence (the score of 0 is the worst imaginable situation, the score of 10 is the best imaginable situation). Due to rounding off, the sum doesn't add to 100%.

6.3.2 r-QoL rating

Figure 6.1 shows the distribution of the two separate rating scales (scale 0-10) of which this combined r-QoL rating consists; a score of 7 (out of 10) or higher was given by more than 50% of the respondents for their overall health and independence. Age and time since TL procedure have a significant influence (a negative relation p=.004, and a positive relation p<.001, respectively) on the overall health rating, and time since laryngectomy has a significant influence on the independence rating (negative relation $p \leq .001$).

Since the overall health and independence ratings have a high internal correlation and were combined to one sum measure, from now on they will be represented as one outcome measure r-QoL. The average combined r-QoL rating is 14.4 (scale 0-20). It appears that respondents under 60 years of age and who have had their TL procedure less than two years ago at the time of the questionnaire rate their r-QoL much lower. Sex, educational level and country of residence did not have a significant influence (p>.05) on the r-QoL rating.

6.3.3 Average number of reported experienced issues per theme

Figure 6.2 shows the average percentage of reported issues per patient per theme (as a percentage of the number of questions in that theme), and the influence of the grouping of respondents on the number of reported issues. The figures do not represent the percentage of respondents that experiences these issues, but the percentage of specific issues within the theme that an average respondent will have experienced or encountered. Because the results are averaged over a large group of respondents, it gives a sensitive comparison method. See Appendix A for an overview of the grouped questionnaire questions per theme and their response rate. Educational level and country of residence did not have a significant influence (p>.05) on the experienced issues per theme and were therefore excluded from the presentation of the results. Sex, age and time since TL procedure do all have an influence on the number of experienced issues of almost all themes. In general, younger respondents (<60 yrs), especially women, who have had the TL procedure less than two years ago, are uniquely disadvantaged in terms of reported participation and functional issues.

6.3.4 Correlations between r-QoL rating and themes

The correlations between the r-QoL and different themes, representing overarching issues, can be found in Table 6.2 and Figure 6.3. To illustrate, as seen in Table 2 approximately 29% of the variance R^2 in the r-QoL ratings can be explained by the reported issues in the themes 'experienced limitations in daily activities' (71% of the 29%) and 'avoiding social activities' (29% of the 29%).

However, the theme 'pulmonary issues' seems to play an important role and has a significant correlation to most other themes: approximately 41% of the variance (textitR^2) can be explained by the reported issues in the themes 'experienced limitations in daily activities' (22% of the 41%) , 'avoiding social activities' (10% of the 41%), 'communication issues' (16% of the 41%), 'experienced vulnerability due to environmental factors' (20% of the 41%), and 'sleep issues' (32% of the 41%).

The significance of the pulmonary issues can be illustrated by the fact that of the specific included questions within this 'pulmonary issues' theme, for example 89% of the respondents report they have to clean out mucus from their stoma or HME several times a day (see Appendix A, 7.5.), and 47% of

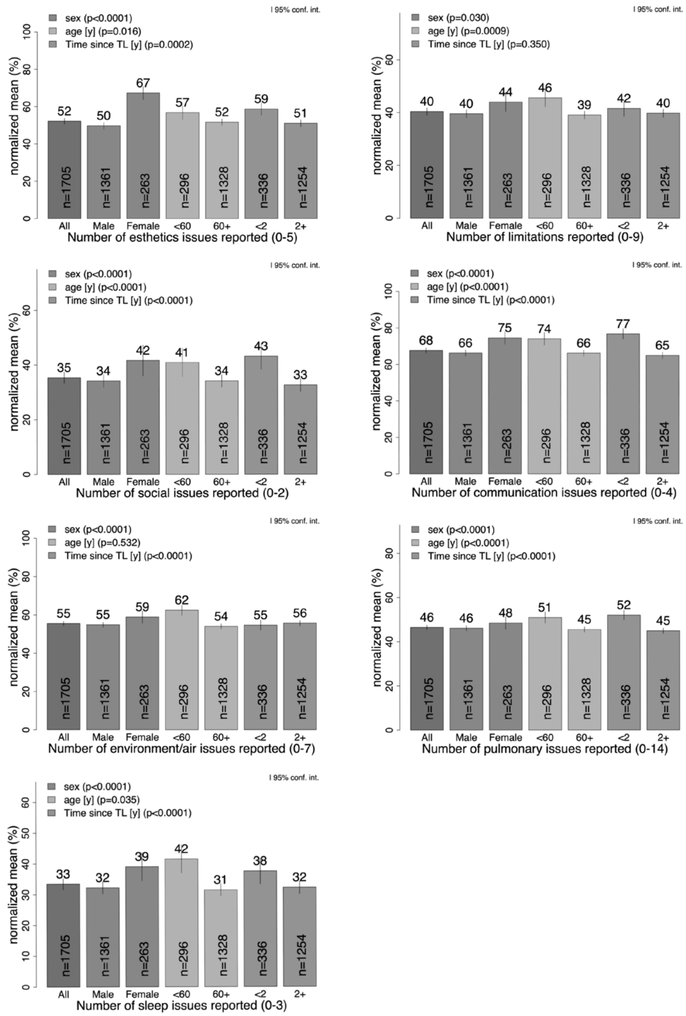

Figure 6.2: The influence of the grouping of respondents on the average percentage of reported issues per respondent per theme. r-QoL = reported Quality of Life, TL = total laryngectomy. Note: The specific grouping of respondents does not add up to the total number of respondents: missing data in the 'time since TL' (n=115 respondents gave 'no answer'), 'sex' and 'age' (n=81 respondents did not give consent to provide this personal information).

the respondents experience frequent coughing during the day (see Appendix A, 7.1.).

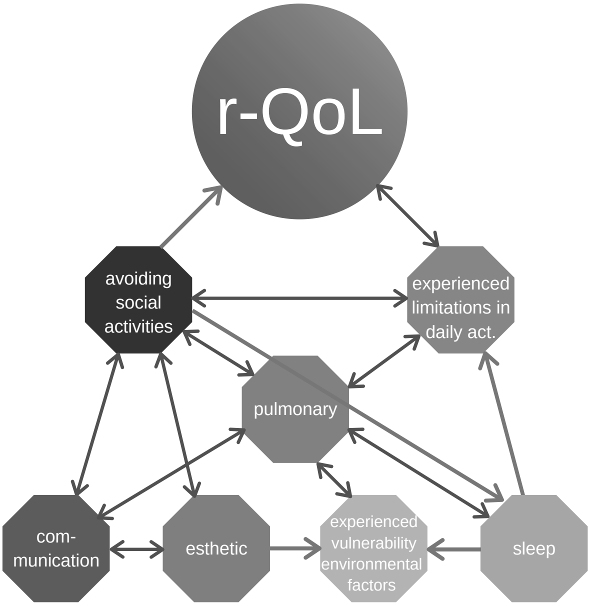

Figure 6.3: Correlations between themes and r-QoL. The single arrows indicate a one-way correlation (e.g. 'esthetics issues' only has an influence on the variance of the 'experienced vulnerability due to environmental factors', but not vice versa), the double arrows indicate an asymmetrical interdependent correlation. r-QoL= reported Quality of Life

Table 6.2: Correlations between themes and r-QoL

	Influence on variance [%]	Relative importance, Sum 100%							
Theme	% of theme variance (R^2) explained by other themes	r-QoL	Esthetic issues	Experienced limitations daily activities	Avoiding Social activities	Communica issues	Experienced vulnerability due to environ. factors	Pulmonary issues	Sleep issues
r-QoL	29	x		71	29				
Esthetic issues	4		X		40	60			
Experienced limitations in daily activities	39	41		X	20			25	14
Avoiding social activities	25		7	46	X	26		21	
Communication issues	20		10		37	X		53	
Experienced vulnerability due to environ. factors	18		6				X	60	34
Pulmonary issues	41			22	10	16	20	X	32
Sleep issues	25				15			85	X

6.4 Discussion

This study presents a large sample of total laryngectomized respondents from multiple countries, examining both a large range of participation and functional issues and r-QoL. The study data is unique and relevant, nevertheless the study lacks validated measures as well as respondents' information regarding, for example, marital status and type of voice rehabilitation. In order to reduce this limitation of the use on a non-validated questionnaire, the method of semantic clustering of questions was used, comparable to validated quality of life questionnaires (e.g. EORTC-C30) [22–24]. Clustering of both questions in r-QoL seems feasible, since it is known that independence is an attribute of the concept quality of life [23, 24].

The demographic characteristics of the respondents (the distribution of age, sex, educational level and time since laryngectomy) were comparable to other studies and correspond to the characterization of "the average total laryngectomized person" as a middle-aged man (around 65 years old, male-to-female ratio of $\sim$ 6:1) [13, 25–29].

Outcomes of the rating scales regarding overall health and independence both had an average rating of 7.2(on a scale 0-10). These two ratings are both notably high, since a TL procedure is usually associated with a lower QoL rating and depressive symptoms [13, 30]. However, retrospective study set-ups like this questionnaire tend to have a larger inclusion of nonproblematic patients than prospective studies, and thus, more positive results (e.g. a more positive r-QoL rating) [30]. However, the distribution of the independence rating is disconcerting: for example, 25% of respondents rate their independence a score of 5 or lower, which can be interpreted as being unable to participate in many daily activities.

Our study shows that respondents who were <2 years since TL, and respondents <60 years old in general report a lower r-QoL rating. For almost all the issue themes, sex, age and time since laryngectomy have a significant influence on (the number of) experienced participation and functional issues. Therefore, younger respondents (<60yrs), especially women, and those who have had the TL procedure less than two years ago, seem to be uniquely disadvantaged in terms of r-QoL ratings and reported participation and functional issues.

The influence of age, sex and time since TL procedure have also been underlined by other studies. Age as an influencing factor on r-QoL is supported by multiple studies reporting that indeed younger total laryngectomized patients experience a higher psychological distress, impacting their coping, since younger patients may have a better preoperative baseline functional status and activity level [4, 31, 32]. Time since laryngectomy as an influencing factor on reported issues has been supported by multiple papers: on average the reported problems decrease over time [5, 33]. The fact that after the first 2 years after the TL procedure the r-QoL improves, shows that rehabilitation and coping can be effective, but takes several years. Earlier studies have shown that gender differences are present in reported issues after total laryngectomy. Women are inclined to experience more post-operative complaints, and issues with social interaction due to stigmatization [4, 31, 34].

The 'pulmonary issues' seem to have a strong correlation to most other themes. Therefore these pulmonary issues might be partially responsible for other reported issues [5]. Although reported less frequently, the issues from the two themes 'avoiding social activities' and 'experienced limitations in daily activities' are the main influencers of the variance in r-QoL. Thus, the ability to fulfill meaningful activities seems to have a greater impact on r-QoL than the purely physical consequences of TL in general. The inclusion of the independence rating in the r-QoL rating could have introduced a bias in the correlation analysis. The concept quality of life includes development and improvement of life (adapting to changed health condition and finding new meaning), independence, achievement of goals and aspirations, and autonomy [22–24]. The study design has its limitations. The response rate was 21%; the questionnaire was sent via email and was shortly available online. Non-response bias might be present and can cause a bias in how well the data represents the actual total laryngectomized population. Additionally, the respondents of this questionnaire were all clients of Atos Medical AB. The selection bias concerns patients most likely using voice prostheses and/or HMEs, education level and internet use overall and across countries, financial status, and insurance or reimbursement systems between countries [3]. It is likely that Atos Medical clients with a higher age or lower educational level were less well reached with this online questionnaire in certain countries. This could explain the differences in education level between countries as well as the relatively high education level in this questionnaire.

6.5 Conclusion

Younger respondents (<60 years), especially women, and those who have had the TL procedure less than 2 years ago, seem to be uniquely disadvantaged in terms of r-QoL ratings and reported participation and functional issues. The experienced limitations in daily activities and avoiding social activities are associated with decrements in the respondents' r-QoL rating. The r-QoL rating is mainly influenced by the ability to do meaningful activities, and less by purely physical consequences of TL. Most issue themes are interdependently correlated. The theme "pulmonary issues" seems to have a strong correlation with most other themes and is key in most other reported issues. Therefore, pulmonary issues might be an underlying cause of many other issues, including experienced limitations in daily activities and avoiding social activities. To improve clinical practice, it is recommended to adequately prepare and monitor patients regarding their participation in social activities, meaningful activities, and pulmonary issues to enhance their QoL.

6.6 Acknowledgments

Atos Medical AB (Malmö, Sweden) and ReD Associates (Copenhagen, Denmark) are acknowledged for their work for making the database available to the Netherlands Cancer Institute—Antoni van Leeuwenhoek. The Netherlands Cancer Institute receives a research grant from Atos Medical AB (Malmö, Sweden), which contributes to the existing infrastructure for quality of life research of the Department of Head and Neck Oncology and Surgery.

6.7 Appendix A

Overview of the grouped questionnaire questions per theme. The percentages indicate the proportion of the total number of respondents (n=1705) who reported to have experienced the specific issue as stated in the question.

1. r-QoL rating
   (a) How would you rate your overall health from 0 to 10? 0 means worst imaginable health state, 10 means best imaginable health state.
   (b) On a scale from 0 to 10, how independent do you feel in completing the activities you want? 0 means the least imaginable independence in completing the activities you want, 10 means the most imaginable independence in completing the activities you want.
2. Esthetic issues
   (a) I have wished my HME looked more discreet: 54%
   (b) I have wished my adhesive, button or tube looked more discreet: 52%
   (c) Covering my stoma with an HME has made me feel more confident: 66%
   (d) Covering my stoma with a scarf or clothing has made me feel more confident: 38%
   (e) I would like to try the following product: an HME that is more discreet and that you wear for going out of the house: 70%
3. Experienced limitations in daily activities
   (a) I have been unable to do my favorite hobbies (e.g. gardening, woodwork, art, etc.): 27%
   (b) I have been unable to fulfill household or caring tasks anymore, such as cooking, taking care of children/grandchildren, or cleaning: 14%
   (c) I have avoided traveling or going on trips because of my stoma: 28%

(d) I have avoided physical exercise or sports (e.g. walking, taking the stairs, jogging, running, sports, dance or housework): 35%

(e) I have sat down or taken a break to catch my breath while wearing an HME: 48%

(f) I have taken out my HME to catch my breath: 51%

(g) I have become short of breath while walking: 42%

(h) I have become short of breath while running, exercising, playing sports, or dancing: 70%

(i) I have become short of breath while doing chores, housework, or manual labor: 56%

4. Avoiding social activities

(a) I have avoided everyday social situations, such as spending time with friends or going out for a meal, because I was worried about mucus or coughing: 35%

(b) I have avoided going to big social events (weddings, parties, etc.): 35%

5. Communication issues

(a) Others have had trouble understanding me when I talk: 63%

(b) I have wished my voice sounded more natural: 79%

(c) I have used methods other than my voice to communicate or get other people's attention (such as clapping, snapping, or gestures): 71%

(d) Other people have not understood that I need an extra second to start speaking: 58%

6. Experienced vulnerability due to environmental factors

(a) I have encountered crowded environments that may have germs (e.g. public transportation, restaurants, etc.): 80%

(b) I have encountered places with pollution (e.g. city streets): 77%

(c) I have encountered situations with dirt, dust, or irritating particles (e.g. factories, workshops, gardening, kitchens with lots of flour/spices): 71%

(d) I have felt irritation in my windpipe from dry air or air particles (sand, flour, dust, dirt, etc.): 45%

(e) I have felt irritation in my windpipe due to allergies: 17%

(f) I have had extra mucus in dry or windy air:45%

(g) I have had extra mucus in cold air: 53%

7. Pulmonary issues
   (a) I have experienced frequent coughing during the day: 47%
   (b) I have experienced increased coughing while being more active: 50%
   (c) My HME has popped out because of coughing: 45%
   (d) My adhesive, tube or button has come loose because of coughing: 51%
   (e) I have cleaned out mucus from the stoma or HME several times a day: 89%
   (f) I have had mucus flare-ups during mealtime or social occasions: 69%
   (g) I have used a nebulizer twice or more during the day: 20%
   (h) I have had to change HMEs during the day because of mucus: 71%
   (i) My adhesive has come loose because of mucus: 52%
   (j) My bed sheets have been soiled by mucus: 14%
   (k) My clothes have been soiled by mucus: 53%
   (l) I have had an infection in my windpipe: 15%
   (m) How often I experience coughing during the day: 47%
   (n) How often I have mucus flare-ups during the day: 40%
8. Sleep issues
   (a) I have woken up during the night from coughing: 48%
   (b) My coughing has woken up my partner or other family members during the night: 29%
   (c) I have struggled to fall asleep due to coughing: 24%

References

[1] C. Debry, A. Dupret–Bories, N. E. Vrana, P. Hemar, P. Lavalle, and P. Schultz, "Laryngeal replacement with an artificial larynx after total laryngectomy: The possibility of restoring larynx functionality in the future," *Head & Neck*, vol. 36, no. 11, pp. 1669–1673, 2014.

[2] P. Boscolo–Rizzo, F. Maronato, C. Marchiori, A. Gava, and M. C. Da Mosto, "Long-term quality of life after total laryngectomy and postoperative radiotherapy versus concurrent chemoradiotherapy for laryngeal preservation," *The Laryngoscope*, vol. 118, no. 2, pp. 300–306, 2008.

[3] E. B. van der Houwen, T. A. van Kalkeren, W. J. Post, F. J. M. Hilgers, B. F. A. M. van der Laan, and G. J. Verkerke, "Does the patch fit the stoma? a study on peristoma geometry and patch use in laryngectomised patients," *Clinical Otolaryngology*, vol. 36, no. 3, pp. 235–241, 2011.

[4] S. R. Cox, J. A. Theurer, S. J. Spaulding, and P. C. Doyle, "The multidimensional impact of total laryngectomy on women," *Journal of Communication Disorders*, vol. 56, pp. 59–75, 2015.

[5] F. J. M. Hilgers, A. H. Ackerstaff, N. K. Aaronson, P. F. Schouwenburg, and N. van Zandwijk, "Physical and psychosocial consequences of total laryngectomy," *Clinical Otolaryngology & Applied Sciences*, vol. 15, no. 5, pp. 421–425, 1990.

[6] A. H. Ackerstaff, F. J. M. Hilgers, N. K. Aaronson, and A. J. M. Balm, "Communication, functional disorders and lifestyle changes after total laryngectomy," *Clinical Otolaryngology & Applied Sciences*, vol. 19, no. 4, pp. 295–300, 1994.

[7] A. Öztürk and M. Mollaoğlu, "Determination of problems in patients with post-laryngectomy," *Scandinavian Journal of Psychology*, vol. 54, no. 2, pp. 107–111, 2013.

[8] F. Shiraz, E. Rahtz, K. Bhui, I. Hutchison, and A. Korszun, "Quality of life, psychological wellbeing and treatment needs of trauma and head and neck cancer patients," *British Journal of Oral and Maxillofacial Surgery*, vol. 52, no. 6, pp. 513–517, 2014.

[9] M. Wells, M. Cunningham, H. Lang, S. Swartzman, J. Philp, L. Taylor, and J. Thomson, "Distress, concerns and unmet needs in survivors of head and neck cancer: a cross-sectional survey," *European Journal of Cancer Care*, vol. 24, no. 5, pp. 748–760, 2015.

[10] B. M. R. Op de Coul, A. H. Ackerstaff, C. J. Van As, F. J. A. Van Den Hoogen, C. A. Meeuwis, J. J. Manni, and F. J. M. Hilgers, "Quality of life assessment in laryngectomized individuals: do we need additions to standard questionnaires in specific clinical research projects?," *Clinical Otolaryngology*, vol. 30, no. 2, pp. 169–175, 2005.

[11] R. Müller, J. Paneff, V. Köllner, and R. Koch, "Quality of life of patients with laryngeal carcinoma: a post-treatment study," *European Archives of Oto-Rhino-Laryngology*, vol. 258, no. 6, pp. 276–280, 2001.

[12] S. Singer, H. Danker, A. Dietz, U. Kienast, F. Pabst, E. F. Meister, J. Oeken, A. Thiele, and R. Schwarz, "Sexual problems after total or partial laryngectomy," *The Laryngoscope*, vol. 118, no. 12, pp. 2218–2224, 2008.

[13] H. Danker, D. Wollbrück, S. Singer, M. Fuchs, E. Brähler, and A. Meyer, "Social withdrawal after laryngectomy," *European archives of Oto-Rhino-Laryngology*, vol. 267, no. 4, pp. 593–600, 2010.

[14] F. Hilgers, A. H. Ackerstaff, C. van As, A. Balm, M. van den Brekel, and I. Tan, “Development and clinical assessment of a heat and moisture exchanger with a multi-magnet automatic tracheostoma valve (Provox Free-Hands HME) for vocal and pulmonary rehabilitation after total laryngectomy,” *Acta Oto-Laryngologica*, vol. 123, pp. 91–99, 2003.

[15] F. J. M. Hilgers, N. K. Aaronson, A. H. Ackerstaff, P. F. Schouwenburg, and N. Van Zandwijk, “The influence of a heat and moisture exchanger (HME) on the respiratory symptoms after total laryngectomy,” *Clinical Otolaryngology & Applied Sciences*, vol. 16, no. 2, pp. 152–156, 1991.

[16] S. Bień, S. Okła, C. J. van As-Brooks, and A. H. Ackerstaff, “The effect of a heat and moisture exchanger (Provox HME) on pulmonary protection after total laryngectomy: a randomized controlled study,” *European Archives of Oto-Rhino-Laryngology*, vol. 267, no. 3, pp. 429–435, 2010.

[17] A. H. Ackerstaff, F. J. M. Hilgers, N. K. Aaronson, M. F. de Boer, C. A. Meeuwis, P. P. M. Knegt, H. A. A. Spoelstra, N. van Zandwijk, and A. J. M. Balm, “Heat and moisture exchangers as a treatment option in the post-operative rehabilitation of laryngectomized patients,” *Clinical Otolaryngology & Applied Sciences*, vol. 20, pp. 504–509, 1996.

[18] J.-C. Mérol, A. Charpiot, T. Langagne, P. Hémar, A. H. Ackerstaff, and F. J. M. Hilgers, “Randomized controlled trial on postoperative pulmonary humidification after total laryngectomy: External humidifier versus heat and moisture exchanger,” *The Laryngoscope*, vol. 122, no. 2, pp. 275–281, 2012.

[19] K. E. van Sluis, L. van der Molen, R. J. J. H. van Son, F. J. M. Hilgers, P. A. Bhairosing, and M. W. M. van den Brekel, “Objective and subjective voice outcomes after total laryngectomy: a systematic review,” *European Archives of Oto-Rhino-Laryngology*, vol. 275, no. 1, pp. 11–26, 2018.

[20] R Core Team, “R: A language and environment for statistical computing. R Foundation for Statistical Computing,” 2018.

[21] U. Grömping *et al.*, “Relative importance for linear regression in R: the package relaimpo,” *Journal of Statistical Software*, vol. 17, no. 1, pp. 1–27, 2006.

[22] A. C. Åberg, B. Sidenvall, M. Hepworth, K. O’Reilly, and H. Lithell, “On loss of activity and independence, adaptation improves life satisfaction in old age–a qualitative study of patients’ perceptions,” *Quality of Life Research*, vol. 14, no. 4, pp. 1111–1125, 2005.

[23] S. Pinto, L. Fumincelli, A. Mazzo, S. Caldeira, and J. C. Martins, “Comfort, well-being and quality of life: Discussion of the differences and similarities among the concepts,” *Porto Biomedical Journal*, vol. 2, no. 1, pp. 6–12, 2017.

[24] M. Puts, N. Shekary, G. Widdershoven, J. Heldens, P. Lips, and D. Deeg, "What does quality of life mean to older frail and non-frail community-dwelling adults in The Netherlands?," *Quality of Life Research*, vol. 16, no. 2, pp. 263–277, 2007.

[25] S. T. Massa, N. Osazuwa-Peters, E. Adjei Boakye, R. J. Walker, and G. M. Ward, "Comparison of the financial burden of survivors of head and neck cancer with other cancer survivors," *Otolaryngology - Head and Neck Surgery*, vol. 145, no. 3, pp. 239–249, 2019.

[26] F. G. R. Souza, I. C. Santos, A. S. de Freitas, L. C. S. Thuler, A. Bergmann, E. Q. Freitas, and F. L. Dias, "Comparative analysis of quality of life in advanced laryngeal and oral cancer undergoing extensive surgeries," *Archives of Head and Neck Surgery*, vol. 47, no. 3, p. e029, 2018.

[27] S. H. Choi, J. E. Terrell, K. E. Fowler, S. A. McLean, T. Ghanem, G. T. Wolf, C. R. Bradford, J. Taylor, and S. A. Duffy, "Socioeconomic and other demographic disparities predicting survival among head and neck cancer patients," *PLOS ONE*, vol. 11, no. 3, p. e0149886, 2016.

[28] N. B. Oozeer, S. Owen, B. Z. Perez, G. Jones, A. R. Welch, and V. Paleri, "Functional status after total laryngectomy: cross-sectional survey of 79 laryngectomees using the performance status scale for head and neck cancer," *The Journal of Laryngology & Otology*, vol. 124, no. 4, pp. 412–416, 2010.

[29] R. Kazi, J. Cordova, J. Kanagalingam, R. Venkitaraman, C. M. Nutting, P. Clarke, P. Rhys Evans, and K. Harrington, "Quality of life following total laryngectomy: Assessment using the UW-QOL scale," *ORL*, vol. 69, no. 2, pp. 100–106, 2007.

[30] E. Armstrong, K. Isman, P. Dooley, D. Brine, N. Riley, R. Dentice, S. King, and F. Khanbhai, "An investigation into the quality of life of individuals after laryngectomy," *Head & Neck*, vol. 23, no. 1, pp. 16–24, 2001.

[31] M. S. Graham and A. Palmer, "Gender difference considerations for individuals with laryngectomies," *Contemporary Issues in Communication Science and Disorders*, vol. 29, pp. 59–67, 2002.

[32] T. D. Woodard, A. Oplatek, and G. Petruzzelli, "Life after total laryngectomy: A measure of long-term survival, function, and quality of life," *Archives of Otolaryngology - Head and Neck Surgery*, vol. 133, pp. 526–532, 2007.

[33] S. Singer, D. Wollbrück, A. Dietz, J. Schock, F. Pabst, H. Vogel, J. Oeken, A. Sandner, S. Koscielny, and K. Hormes, "Speech rehabilitation during the first year after total laryngectomy," *Head & Neck*, vol. 35, no. 11, pp. 1583–1590, 2013.

[34] K. E. van Sluis, A. F. Kornman, L. van der Molen, M. W. M. van den Brekel, and G. Yaron, "Women's perspective on life after total laryngectomy: a qualitative study," *International Journal of Language & Communication Disorders*, vol. 55, no. 2, pp. 188–199, 2020.

CHAPTER 7

Expiratory muscle strength training in patients after total laryngectomy; a feasibility pilot study

Abstract

Objectives: Expiratory muscle strength training (EMST) is a threshold based device-driven treatment for improving expiratory pressure. EMST proved to be effective in different patient groups to improve cough function. To date, EMST has not been tested in the total laryngectomy population (TL).

Methods: This prospective, randomized case-series study examined feasibility, safety, and compliance of EMST in a group of TL participants and its effects on pulmonary function, physical exertion, fatigue, and vocal functioning. Ten TL participants were included in the study to perform a four till eight weeks of EMST. Objective and subjective outcome measures included manometry, spirometry, cardio pulmonary exercise testing (CPET), voice recordings, and patient reported outcome measures. Group means were reported and estimates of the effect are shown with a 95% confidence interval, using single sample t-tests.

Results: Nine participants completed the full study protocol. Compliance to the training program was high. All were able to perform the training, although

[0]Appeared as: van Sluis, K.E., Kornman, A.F., Groen, W.G., van den Brekel, M.W.M., van der Molen, L., Hoffman-Ruddy, B., Stuiver, M.M. "Expiratory muscle strength training in patients after total laryngectomy; a feasibility pilot study." *Annals of Otology, Rhinology & Laryngology*, epub ahead of print, 2020.

it requires adjustments of the device and skills of the participants. Maximum expiratory pressure (MEP) and vocal functioning in loudness improved over time. After EMST no changes were seen in other objective and subjective outcomes.

Conclusions: EMST appears to be feasible and safe after total laryngectomy. MEP improved over time but no improvement in the clinically relevant outcome measures were seen in this sample of relatively fit participants. Further investigation of the training in a larger group of participants who report specifically pulmonary complaints is recommended to investigate if the increase in MEP results in clinical benefits.

7.1 Introduction

Total laryngectomy (TL), which involves surgical removal of the larynx, leads to lifelong changes in voice, swallowing and airway. Pulmonary driven speech can be re-established with insertion of a voice prosthesis, also called tracheoesophageal speech[1]. Compared to healthy individuals the voices of tracheoesophageal speakers have a rough voice quality, reduced loudness, and limited range [2]. After TL pulmonary condition is affected as air entering the lungs via the stoma is not warmed and humidified by the nose and upper respiratory tract, which leads to impaired mucociliary clearance. Impaired mucociliary function is the main cause of pulmonary complaints such as increased mucus production and forced mucus expectoration after TL [3–5]. Warming and humidifying the inhaled air is a key factor in improving pulmonary condition. Over the years several options were developed: external humidifiers, stoma cloths (e.g. bibs) and heat-moisture-exchange devices (HME-devices) [5–8]. Both stoma cloths and HME-devices are proved to be effective in everyday life and can reduce pulmonary problems in this patient group [9–11]. Nevertheless, pulmonary complaints including increased mucus production and forced mucus expectoration are still present after TL [12, 13]. The altered anatomy results in an physiological altered cough function after TL. Removal of the larynx eliminates the ability to generate subglottic pressure prior to cough onset. Cough requires high expiratory airflows to aerosolize and remove material that is unable to be removed by mucociliary action. Cough "strength" is determined by the ability to generate high expiratory pressures immediately prior to cough onset and by the volume of air that is expelled from the airways [14]. With the absence of the ability to build up subglottic pressure, it is assumed that after a TL increased cough strength can only be achieved by increasing the volume and speed of cough expiratory airflow. To date, cough function after TL and the potential benefit of rehabilitation techniques have not been studied. A well-known rehabilitation technique to improve cough function is expiratory muscle strength training (EMST). Recently, the use of EMST has been evaluated in several studies [15], including different patient groups such as Parkinson's disease [14, 16, 17], multiple sclerosis [18], sleep apnea [19], head and neck cancer patients with dysphagia [20], stroke [21, 22], amyotrophic lateral sclerosis [23], supracricoid partial laryngectomy [24], and healthy participants [25–27]. In several populations, to date, EMST has been shown to increase expiratory pressure generation by 30 to 150%, with an average increase of approximately 50% in a 4-week period of time [14, 16, 18, 20, 22, 23, 27]. Most of the clinical studies cited above aimed at improving cough strength to improve clearing of aspiration of oral intake or saliva. [16–18, 20, 22, 24] EMST has not been tested in the TL population. The main objective of the present study is to evaluate feasibility, safety, and compliance of EMST in individuals who have undergone TL. Subsequently, objective and subjective outcomes of the effects of EMST on pulmonary function, physical exertion, fatigue, and vocal functioning are assessed.

7.2 Materials and Methods

This prospective, randomized case-series study examined feasibility of EMST and its effects on pulmonary function and voice, in a group of male TL participants, who were at least a half year post-surgery and, if applicable, post-operative (chemo) radiation. Participants were screened and recruited from the head and neck department of the Netherlands Cancer Institute, Amsterdam, The Netherlands. Exclusion criteria were: recurrence of head and neck cancer, a history of lung cancer, severe asthma, tuberculosis, uncontrolled or untreated hypertension, a heart attack in the last year, or abdominal hernia. The study was approved by the medical ethical review committee of the Netherlands Cancer Institute (registration nr. NL60167.031.16.).

7.2.1 Expiratory muscle strength training and adjustment to use after total laryngectomy

EMST150 (Aspire Products) is a threshold based device-driven treatment for improving expiratory pressure generating capacity. EMST employs a handheld training device consisting of a plexiglass tube. Inside the device is a variable tension spring controlling a valve that is calibrated in pressure, adjustable from 30 till 150cmH20. When enough pressure is developed, the valve opens, allowing air to flow through the trainer. The patient is instructed to exhale forcefully through the device. In this way, EMST targets the rectus abdominis and internal intercostal muscles through a program of progressive overloading which is an exercise stimulus specific for forceful expiration.

To adjust the use of the EMST150 device for TL participants, we developed and produced an adaptor to connect the device at the adhesive baseplate in front of the tracheostoma. This adaptor contains an opening on top to allow inhalation which can be occluded manually, a small lumen for use of the voice

Figure 7.1: EMST150 connected to a baseplate with the adapter and voice prosthesis plug present.

Figure 7.2: EMST150 during practice, connected to the baseplate attached to the tracheostoma. The index finger is used to occlude the opening on top during forced exhalation.

prosthesis plug, and a tube which enables connection with a manometer (Figure 7.1 and 2).

7.2.2 Emst study training protocol

All participants performed four weeks of EMST (period A) which consisted of five sets of five repetitions across five days per week. Strength measurements and subsequent adjustment of the pressure of the EMST device were performed at baseline and after each training week. Participants were instructed to perform their training sessions at home and log every training and details in a diary to check compliance. During the weekly visits to the hospital the researchers monitored the participants' compliance by discussing their training results. The pressure of the device was set at approximately 80% of the participant's mean Maximum Expiratory Pressure (MEP). In the following four weeks (period B), participants were allocated to different protocols depending on the group they were randomized to. Participants in group 1 discontinued the training. Participants in group 2 continued EMST for another four weeks, but with a lower frequency of two days per week. Block-randomization in blocks of four with stratification for age was performed with help of Alea software.

Adjustments in training procedures with the EMST150 device were needed. Before each training session, the participants' voice prosthesis had to be blocked with a plug to avoid air escaping through the voice prosthesis into the esophagus. Participants were instructed to connect the EMST device at the adhesive in front of the tracheostoma, take a deep breath, close the opening on top of the adaptor with a finger, and exhale forcefully into the EMST device until enough pressure was built up to open the valve inside.

7.2.3 Feasibility, safety, and compliance

The main objective of the study was to evaluate feasibility, safety, and compliance with the EMST program. Any difficulties regarding participants' performance of the training and assessment procedures were documented. Participants were instructed to log every training session and reflect on their experiences in a diary. At the end of the training program participants filled in a short questionnaire on whether they found the training feasible and whether they could stay motivated during the period of training.

7.2.4 Objective and subjective outcome measures

The effects of EMST on pulmonary function, physical exertion, fatigue, and vocal functioning are assessed with manometry, spirometry, cardio pulmonary exercise testing (CPET), voice recordings, and questionnaires. Time points of the assessments are shown in Table 7.1. MEP in cmH_2O was obtained with a calibrated digital manometer (Druck DPI 705) connected to the adapter whilst the EMST device was adjusted to the maximum pressure of 150 cmH_2O and connected to the tracheostoma. Participants were instructed to sit, take a deep breath, occlude the adapter and exhale as forcefully as possible. Peak Expiratory Flow (PEF) in L/min was obtained with a Micro l spirometer combined with a Microgard ll filter (PT Medical) which was placed directly on the baseplate of the stoma. Participants were instructed to inhale calm but deep and then exhale as forcefully and fast as they could. A series of three forced expirations was used to obtain MEP and PEF, the mean of the three trials was used for the analysis. With a calibrated ergospirometry system (Jaeger Masterscreen CPX, Houten, The Netherlands), connected to the stoma, vital capacity (VC), and forced expiratory volume in the first second (FEV1) were measured. Participants performed a Cardio Pulmonary Exercise Test (CPET) on an electronically braked cycle ergometer (Lode Corival, ProCare, Groningen, The Netherlands). An adapter was made to fit the flow turbine directly to the baseplate in front of the tracheostoma and a headband for fixation was used to support the adapter and flow turbine. The ventilatory efficiency was defined by minute volume of expired air relative to volume of CO_2 produced (VE/VCO_2). Patients cycled till they had reached a respiratory exchange ratio of 1.0 thus precluding them from a maximal exertion. Directly after finishing the test, participants were asked to rate their perceived level of exertion and dyspnea on a Borg scale [28, 29]. Voice recordings were made with a head mounted microphone and recorded with Audacity software[30]. Participants were instructed to perform a sustained vowel /a/ as long as possible to measure maximum phonation time (MPT). Vocal range in Herz (Hz) and dynamic range in decibel (dB) were measured as an outcome of the difference between the lowest and highest and softest and loudest /a/ produced, respectively. The best of three attempts was used for each value. All voice recordings were acoustically analyzed with PRAAT software [31]. Self-reported vocal functioning is assessed with the Voice Handicap Index-10 (VHI-10) [32], self-reported fatigue with the Short Fatigue Questionnaire (SFQ) [33], and self-reported pulmonary functioning with the clinical COPD questionnaire (CCQ)[34].

Table 7.1: Overview of the Assessed Outcome Measures per Time Point, Visualization of Period A and B Over Time.

	Wk0	**Wk1**	**Wk2**	**Wk3**	**Wk4**	**Wk8**
	Period A					Period B
Manometry - MEP	X	X	X	X	X	X
Spirometry - PEF	X	X	X	X	X	X
Spirometry - VC, FEV1	X				X	
CPET—VE/VCO_2, Borg exertion-, Borg dyspnea scale	X				X	
Voice assessment—Vocal range (Hz and dB), MPT	X				X	X
Short fatigue questionnaire [33]	X				X	X
Clinical COPD Questionnaire [34]	X				X	X
Voice Handicap Index-10 [32]	X				X	X

Abbreviations: CPET, Cardio Pulmonary Exercise Testing; FEV1, Forced Expiratory Volume in the first second; MEP, Maximum Expiratory Pressure; MPT, Maximum Phonation Time; PEF, Peak Expiratory Flow; VC, Vital Capacity.

7.2.5 Statistical analysis

Due to lack of preliminary data with regard to pulmonary exercise programs in TL patients, there was no meaningful way to perform sample size calculations related to the quantitative outcome. All data was analyzed using SPSS software [35]. Accordingly, no significance tests were performed. Group means were estimated with a 95% confidence interval, using single sample t-tests.

7.3 Results

Ten participants were included and signed informed consent. Participants age ranged from 50-73 years, and all were tracheo-esophageal speakers (characteristics are presented in 7.2).

Table 7.2: Participant Characteristics.

Participant	Age in Years	Time since TL in Years	Timing Radiotherapy	Flap Reconstruction	Neck Dissection	Tracheo Esophageal Speech: Manually/Hands Free Device	Randomization Group 1 = Stop Training 2 = Continue Training
1	65	13	Post-surgery	No	Both sides	Freehands	2
2	67	18	Pre-surgery	Unknown†	Unknown†	Manually	1
3	50	3	Pre-surgery	No	Both sides	Both, alternating	1
4	68	1	Pre-surgery	PM-flap	No	Manually	2
5	57	15	Pre-surgery	No	Both sides	Freehands	1
6	63	7	Pre-surgery	No	No	Manually	2
7	68	3	Pre-surgery	No	Both sides	Both, alternating	1
8	57	10	Post-surgery	No	Both sides	Freehands	2
9	78	3	Post-surgery	No	Left side	Manually	1
10	73	20	Pre-surgery	Unknown†	Unknown†	Manually	2

Abbreviations: PM-flap, Pectoralis Major flap; TL, total laryngectomy.
†No surgical information was available.

7.3.1 Feasibility, safety, and compliance

Using the EMST device with help of an adapter on the tracheostoma of the TL participants appeared to be generally feasible. All participants could perform the training. One participant withdrew from the study after one week due to unrelated medical reasons. For the remaining nine participants, compliance to the allocated training program was 95.5%. According to the final short questionnaire, participants did not experience problems to stay motivated during the training weeks. Minor problems included leakage of air underneath the adhesive, which was reported by four of the participants.Two participants occasionally experienced dizziness during and shortly after the training. Three participants were not able to plug the voice prosthesis prior to the training, and continued training without plug. From the six participants who did plug their voice prosthesis before training, three mentioned that plugging was a hassle.

Regarding to safety, one adverse event occurred with one of the participants in his fourth EMST training week. After a training session with the EMST, the voice prosthesis was not in situ anymore. A new voice prosthesis was placed and an X-thorax was made which showed that the voice prosthesis was not in the lungs. All participants and the medical ethical review committee were informed.

Table 7.3: Mean, Standard Deviation, and Confidence Interval for the Outcome Measures of the Total Group of Participants at Baseline, Week 4 and Week 8.

	Baseline		Wk 4		Wk 8	
		95% CI		95% CI		95% CI
	Mean (SD)	Lower–Upper	Mean (SD)	Lower–Upper	Mean (SD)	Lower–Upper
MEP—cmH_2O	125.5 (19.6)	110.4-140.6	174.8 (28.6)	152.8-196.8	164.9 (30.0)	141.8-189.0
PEF—L/min	455.9 (116.4)	374.8-537.0	445.5 (116.432)	356.0-535.0	405.3 (112.7)	318.8-492.0
CPET—VE/VCO_2	29.5 (4.7)	25.9-33.1	28.4 (2.8)	26.3-30.57	-	-
Borg exertion scale	13.0 (1.7)	11.7-14.3	12.4 (1.3)	11.4-13.5	-	-
Borg dyspnea scale	2.7 (2.2)	1.0-4.4	2.2 (1.9)	0.8-3.7	-	-
FEV1—L	3.2 (0.7)	2.6-3.7	3.0 (0.6)	2.5-3.5	-	-
VC—L	4.3 (1.0)	3.5-5.1	4.1 (0.8)	3.5-4.8	-	-
Vocal range—Hz	144.6 (98.9)	62.0-227.3	147.0 (63.0)	94.4-199.6	149.0 (101.4)	64.3-233.7
Vocal range—dB	26.4 (5.6)	21.7-31.0	31.9 (3.7)	28.8-35.0	27.6 (5.4)	23.1-32.2
MPT—sec	12.5 (7.9)	5.9-19.0	12.2 (8.5)	5.1-19.3	12.0 (6.8)	6.2-17.7
VHI-10[32]	15.6 (10.0)	7.3-24.0	16.0 (9.7)	8.0-24.2	16.5 (8.8)	9.1-23.9
SFQ[33]	8.1 (5.1)	4.2-12.0	8.44 (4.2)	5.2-11.7	9.1 (4.7)	5.5-12.7
CCQ total score[34]	1.2 (0.8)	0.6-1.8	1.3 (0.7)	0.7-1.9	1.2 (0.6)	0.8-1.7

Abbreviations: CCQ, Clinical COPD Questionnaire; CPET, Cardio Pulmonary Exercise Testing; FEV1, forced expiratory volume in the first second; MEP, Maximum Expiratory Pressure; MPT, Maximum Phonation Time; PEF, Peak Expiratory Flow; SFQ, Short Fatigue Questionnaire; VC, Vital capacity; VHI-10, Voice Handicap Index 10 item version.

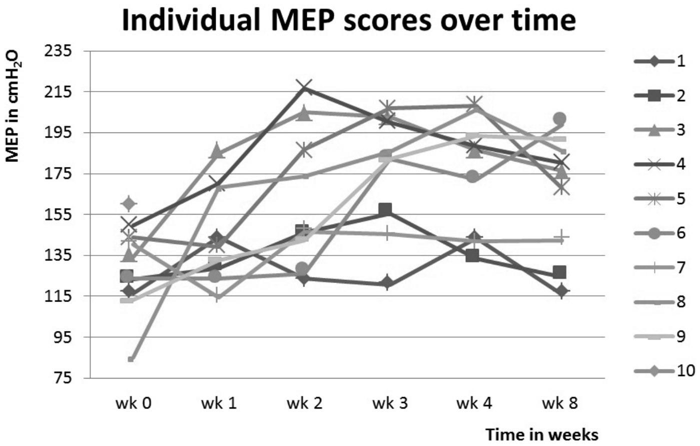

Figure 7.3: Individual Maximum Expiratory Pressure (MEP) scores over time in cmH_2O.

7.3.2 Objective and subjective outcome measures

Results on pulmonary function, physical exertion, fatigue, and vocal functioning are presented in Table 7.3. We observed a non-linear increase in the manometry outcome MEP. During the first four training weeks mean MEP increased from 125.5 cmH_2O to 174.8 cmH_2O during the first 4 weeks (95% CI baseline 110.4–140.6, 4wk 152.8–196.8). After four weeks, MEP stabilized, with no evidence for differences in detraining between group 1 and group 2 (Figure 7.3). Overall, MEP decreased slightly to a mean of 164.9 cmH_2O (95% CI 141.8–189.0) at the end of the follow-up. Three of the nine participants achieved a MEP-score above 187.5 cmH_2O during the first four training weeks. For those participants, training at 80% of the mean MEP was not possible from that moment onward, since the maximum setting of the EMST is 150 cmH_2O. These participants continued their training on the maximum setting of the EMST150-device. The three participants who used no plug during training showed increase of the MEP as well.

We observed no effect on spirometry outcome PEF over time (Figure 7.4). Mean PEF values were 455.9L/min at baseline and 445.5L/min after four weeks (95%-CI:baseline 374.8–537.0, 4wk 356.0–535.0). In this sample, no effect of detraining was seen in PEF after reducing or stopping the training with a mean PEF value of 405.3 at wk 8 (95%-CI:318.8–492.0).

No changes over time were found for self-reported pulmonary problems (CCQ), vocal functioning (VHI-10), fatigue (SFQ). No differences were found for outcomes in physical exertion (CPET, Borg scales). Objective assessment of vocal functioning showed no differences in MPT and vocal range in Hz. Exception was dynamic range in dB, which increased from 26.4dB (95% CI 21.7 –

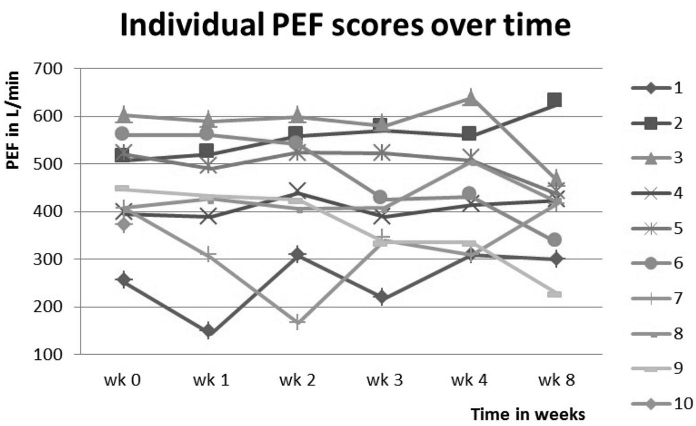

Figure 7.4: Individual Peak Expiratory Flow (PEF) scores over time in liter per minute.

31.0) to 31.9dB (95% CI 28.8 – 35.0) after four weeks of training. After period B vocal range in dB returned back to baseline values (mean 27.6dB 95% CI 23.1 to 32.2).

7.4 Discussion

To date, this is the first study to investigate feasibility, safety, and compliance of EMST in a group of TL participants. EMST appears to be somewhat challenging, but feasible. MEP improved over time but this did not seem clinically relevant for this group of relatively fit participants.

Challenges encountered the need for an adapter, skills of the participants with plugging the voice prosthesis, placing the EMST150 with an adaptor on the tracheostoma, and creating an airtight seal. One safety issue occurred with a participant presenting without a voice prosthesis after a training session. Nevertheless, compliance to the training program was high (>95%).

The results show a clear increase in MEP over time in contrast to PEF values in which no change was seen. The increase of MEP as a result of EMST is consistent with both biological rationale and former findings [14, 16, 18, 20, 22, 23, 25]. Baseline MEP scores were higher than predicted for eight out of ten participants when compared with reference values for healthy adults [36]. No normative MEP values for TL patients are found. The higher than expected MEP scores might be the effect of frequent coughing and forced expectoration which is present after TL. The high MEP values are in contrast with the study of Hutcheson et al. [20], in which head and neck cancer patients suffering from chronic aspiration were included. Their group showed reduced average MEPs at baseline [20]. Palmer et al. [24] reported a good MEP at baseline in their group after partial laryngectomy and reasoned their candidates must have good pulmonary support to tolerate some amount of aspiration during recovery.

We observed no changes over time for PEF outcomes. Contrary to the MEP scores, baseline scores were lower than predicted for seven participants compared with reference values [37]. No normative values for PEF in TL patients are found. In the group of partial laryngectomy patients performing EMST a significant increase in peak cough flow (L/min) was found, from below normal prior to the intervention to within normal range after training.24 Peak expiratory flow rate also increased significantly with training in a group of elderly.26 It is worth questioning why PEF values did not improve in the studied group of TL participants. A possible reason is impaired cough technique. As there is no glottic closure, the higher built up pressure does not lead to increased flow. Besides, no special attention was paid at the exhaling technique (i.e. generating the force primarily using the rectus abdominis muscle, and keeping the thorax maximally expanded during the first part of the forced expiration) during the training.

The increase in MEP and vocal range in dB did not lead to an improvement in the clinical relevant outcome measures regarding pulmonary function, physical exertion, fatigue, and vocal functioning. It was disappointing to see that no clinically relevant benefits were found in the self-reported outcome measures. This might be the result of a sample of relatively fit TL participants which showed high baseline values in MEP and who mentioned no specific pulmonary complaints. Despite an improvement in loudness, related to the dynamic range in dB after four weeks of training, no changes in self-reported vocal functioning (VHI-10) were seen.

Although this pilot study offered useful insights, there are some limitations which should be mentioned related to the included group, use of EMST device and performance of training and measurements. Because of the small number of participants the results of this first EMST study in a TL group must be interpreted with caution. The outcomes cannot be generalized to the entire TL population. In particular, because the participants in our sample were all male, all fluent tracheoesophageal speakers, and relatively fit from the start, they may have progressed less and perceived less benefit, compared to what might be expected from TL patients who are less fit and report explicit coughing problems. Participants who achieved a MEP-score above 187.5 cmH_2O continued the training with the device set at the maximum work load of 150 cmH_2O. It remains unclear if the increase in MEP would have been even larger in case of training with an EMST device with a wider range. To measure the effect of EMST with manometry, spirometry and CPET, adjustments were needed for use on the tracheostoma. Problems creating an airtight seal and plugging the voice prosthesis and the resulting air leakage could have influenced the results of the training and the measured outcomes negatively. VHI-10, SFQ, CCQ and Borg scales were best available questionnaires, although not specifically validated for TL population, and should therefore be interpreted with caution.

For future research, it should be considered to test EMST in a group of TL participants without plugging the voice prosthesis prior to the training. If this also results in improvements, users can be spared the hassle of plugging. Next to this, special attention should be given towards exhaling technique used during the training. It is recommended to assess effectiveness of the training in a large group of participants which includes patients who explicitly report pulmonary problems, less fit elderly participants, and women. Finally, there are still unanswered questions about the association between patients characteristics (e.g. time since TL, flap reconstruction, neck dissection etc.) and effectiveness of EMST.

7.5 Conclusion

This pilot feasibility study indicates that an EMST program is generally safe and feasible in individuals following TL, although it requires adjustments of the device and skills of the participants to perform the training. Compliance to the training program was high. The EMST leads to an increase MEP, no evident changes in PEF outcomes were found. An increase in dynamic range in dB was seen, but this did not result in less reported voice handicap measured with VHI-10. No effects were found in the voice parameters MPT and vocal range in Hz, and ventilatory efficiency during exercise. For this relatively fit group of TL participants, clinical relevant benefits measured with self-reported clinical outcomes could not be determined. It is recommended to assess EMST in a less fit TL population who specifically report pulmonary problems to further investigate potential clinical benefit.

7.6 Acknowledgements

The authors acknowledge Atos medical (Malmö, Sweden) for their research grant, which contributes to the existing infrastructure for quality of life research of the Department of Head and Neck Oncology and Surgery. The authors acknowledge Aspire Products for providing EMST training devices free of charge to study participants. FJM Hilgers emeritus professor and emeritus head and neck surgeon at the Netherlands Cancer Institute and University of Amsterdam is greatly acknowledged for his innovative idea, enthusiasm, and support to test EMST in total laryngectomy patients.

References

[1] K. E. van Sluis, L. van der Molen, R. J. van Son, F. J. Hilgers, P. A. Bhairosing, and M. W. van den Brekel, "Objective and subjective voice outcomes after total laryngectomy: a systematic review," *European Archives of Oto-Rhino-Laryngology*, vol. 275, no. 1, pp. 11–26, 2018.

[2] T. Drugman, M. Rijckaert, C. Janssens, and M. Remacle, "Tracheoesophageal speech: A dedicated objective acoustic assessment," *Computer Speech & Language*, vol. 30, no. 1, pp. 16–31, 2015.

[3] M. Maurizi, G. Paludetti, G. Almadori, F. Ottaviani, and T. Todisco, "Mucociliary clearance and mucosal surface characteristics before and after total laryngectomy," *Acta Oto-Laryngologica*, vol. 102, no. 1-2, pp. 136–145, 1986.

[4] F. J. Hilgers and A. H. Ackerstaff, "Comprehensive rehabilitation after total laryngectomy is more than voice alone," *Folia Phoniatrica et Logopaedica*, vol. 52, no. 1-3, pp. 65–73, 1999.

[5] C. van den Boer, S. H. Muller, V. van der Noort, R. V. Olmos, A. Minni, C. Parrilla, F. J. Hilgers, M. W. van den Brekel, and S. van der Baan, "Effects of heat and moisture exchangers on tracheal mucociliary clearance in laryngectomized patients: a multi-center case–control study," *European Archives of Oto-Rhino-Laryngology*, vol. 272, no. 11, pp. 3439–3450, 2015.

[6] S. Bień, S. Okła, C. J. van As-Brooks, and A. H. Ackerstaff, "The effect of a heat and moisture exchanger (provox® hme) on pulmonary protection after total laryngectomy: a randomized controlled study," *European Archives of Oto-Rhino-Laryngology*, vol. 267, no. 3, pp. 429–435, 2010.

[7] C. Parrilla, A. Minni, H. Bogaardt, G. F. Macri, M. Battista, R. Roukos, M. Pandolfini, G. Ruoppolo, G. Paludetti, and L. D'Alatri, "Pulmonary rehabilitation after total laryngectomy a multicenter time-series clinical trial evaluating the provox xtrahme in hme-naïve patients," *Annals of Otology, Rhinology & Laryngology*, pp. 706–713, 2015.

[8] F. Hilgers, N. Aaronson, P. Schouwenburg, and N. Zandwijk, "The influence of a heat and moisture exchanger (hme) on the respiratory symptoms after total laryngectomy," *Clinical Otolaryngology & Allied Sciences*, vol. 16, no. 2, pp. 152–156, 1991.

[9] G. Quail, J. J. Fagan, O. Raynham, H. Krynauw, L. R. John, and H. Carrara, "Effect of cloth stoma covers on tracheal climate of laryngectomy patients," *Head & Neck*, vol. 38, no. Suppl 1, pp. E480–E487, 2016.

[10] J. C. Mérol, A. Charpiot, T. Langagne, P. Hémar, A. H. Ackerstaff, and F. J. Hilgers, "Randomized controlled trial on postoperative pulmonary humidification after total laryngectomy: external humidifier versus heat and moisture exchanger," *The Laryngoscope*, vol. 122, no. 2, pp. 275–281, 2012.

[11] L. Lansaat, C. van den Boer, S. H. Muller, V. van der Noort, M. W. van den Brekel, and F. J. Hilgers, "Ex vivo humidifying capacity and patient acceptability of stoma cloths in laryngectomized individuals," *Head & Neck*, vol. 39, no. 5, pp. 921–931, 2017.

[12] J. Bickford, J. Coveney, J. Baker, and D. Hersh, "Validating the changes to self-identity after total laryngectomy," *Cancer Nursing*, vol. 42, no. 4, pp. 314–322, 2019.

[13] B. Op de Coul, A. Ackerstaff, C. Van As, F. Van den Hoogen, C. Meeuwis, J. Manni, and F. Hilgers, "Quality of life assessment in laryngectomized individuals: do we need additions to standard questionnaires in specific clinical research projects?," *Clinical Otolaryngology*, vol. 30, no. 2, pp. 169–175, 2005.

[14] T. Pitts, D. Bolser, J. Rosenbek, M. Troche, M. S. Okun, and C. Sapienza, "Impact of expiratory muscle strength training on voluntary cough and swallow function in parkinson disease," *CHEST Journal*, vol. 135, no. 5, pp. 1301–1308, 2009.

[15] H. Laciuga, J. C. Rosenbek, P. W. Davenport, and C. M. Sapienza, "Functional outcomes associated with expiratory muscle strength training: narrative review," *Journal of Rehabilitation Research & Development*, vol. 51, no. 4, pp. 535–546, 2014.

[16] Y.-C. Kuo, J. Chan, Y.-P. Wu, J. R. Bernard, and Y.-H. Liao, "Effect of expiratory muscle strength training intervention on the maximum expiratory pressure and quality of life of patients with parkinson disease," *NeuroRehabilitation*, vol. Preprint, pp. 1–8, 2017.

[17] M. S. Troche, M. S. Okun, J. C. Rosenbek, N. Musson, H. H. Fernandez, R. Rodriguez, J. Romrell, T. Pitts, K. M. Wheeler-Hegland, and C. M. Sapienza, "Aspiration and swallowing in parkinson disease and rehabilitation with EMST: a randomized trial," *Neurology*, vol. 75, no. 21, pp. 1912–1919, 2010.

[18] E. P. Silverman, S. Miller, Y. Zhang, B. Hoffman-Ruddy, J. Yeager, and J. J. Daly, "Effects of expiratory muscle strength training on maximal respiratory pressure and swallow-related quality of life in individuals with multiple sclerosis," *Multiple Sclerosis Journal – Experimental, Translational and Clinical*, vol. 3, no. 2, p. 2055217317710829, 2017.

[19] Y.-C. Kuo, T.-T. Song, J. R. Bernard, and P. Y.-H. Liaoo, "Short-term expiratory muscle strength training attenuates sleep apnea and improves sleep quality in patients with obstructive sleep apnea," *Respiratory Physiology & Neurobiology*, no. 243, pp. 86–91, 2017.

[20] K. A. Hutcheson, M. P. Barrow, E. K. Plowman, S. Y. Lai, C. D. Fuller, D. A. Barringer, G. Eapen, Y. Wang, R. Hubbard, S. K. Jimenez, L. G. Little, and J. S. Lewin, "Expiratory muscle strength training for radiation-associated aspiration after head and neck cancer: A case series," *The Laryngoscope*, vol. 128, no. 5, pp. 1044–1051, 2018.

[21] J. S. Park, D. H. Oh, M. Y. Chang, and K. M. Kim, "Effects of expiratory muscle strength training on oropharyngeal dysphagia in subacute stroke patients: a randomised controlled trial," *Journal of Oral Rehabilitation*, vol. 43, no. 5, pp. 364–372, 2016.

[22] K. W. Hegland, P. W. Davenport, A. E. Brandimore, F. F. Singletary, and M. S. Troche, "Rehabilitation of swallowing and cough functions following stroke: An expiratory muscle strength training trial," *Archives of Physical Medicine and Rehabilitation*, vol. 97, no. 8, pp. 1345–1351, 2016.

[23] E. K. Plowman, L. Tabor-Gray, K. M. Rosado, T. Vasilopoulos, R. Robison, J. L. Chapin, J. Gaziano, T. Vu, and C. Gooch, "Impact of expiratory strength training in amyotrophic lateral sclerosis: Results of a randomized, sham-controlled trial," *Muscle Nerve*, vol. 59, no. 1, pp. 40–46, 2019.

[24] A. D. Palmer, R. K. Bolognone, S. Thomsen, D. Britton, J. Schindler, and D. J. Graville, "The safety and efficacy of expiratory muscle strength training for rehabilitation after supracricoid partial laryngectomy: A pilot investigation," *Annals of Otology, Rhinology & Laryngology*, p. 3489418812901, 2018.

[25] S. Baker, P. Davenport, and C. Sapienza, "Examination of strength training and detraining effects in expiratory muscles," *Journal of Speech, Language, and Hearing Research*, vol. 48, no. 6, pp. 1325–1333, 2005.

[26] J. S. Park, D. H. Oh, and M. Y. Chang, "Effect of expiratory muscle strength training on swallowing-related muscle strength in community-dwelling elderly individuals: a randomized controlled trial," *Gerodontology*, vol. 34, no. 1, pp. 121–128, 2017.

[27] J. Kim, P. Davenport, and C. Sapienza, "Effect of expiratory muscle strength training on elderly cough function," *Archives of Gerontology and Geriatrics*, vol. 48, no. 3, pp. 361–366, 2009.

[28] G. Borg, *Borg's perceived exertion and pain scales*. Champaign, IL: Human Kinetics, 1998.

[29] G. A. Borg, "Psychophysical bases of perceived exertion," *Medicine & Science in Sports & Exercise*, vol. 14, no. 5, pp. 377–381, 1982.

[30] D. Mazzoni and R. Dannenberg, "Audacity [software]," *The Audacity Team, Pittsburg, PA, USA*, 2000.

[31] P. Boersma, "Praat, a system for doing phonetics by computer," *Glot International*, vol. 5:9/10, pp. 341–345, 2001.

[32] C. A. Rosen, A. S. Lee, J. Osborne, T. Zullo, and T. Murry, "Development and validation of the voice handicap index-10," *The Laryngoscope*, vol. 114, no. 9, pp. 1549–1556, 2004.

[33] M. Alberts, E. Smets, J. Vercoulen, B. Garssen, and G. Bleijenberg, "'Verkorte vermoeidheidsvragenlijst': een praktisch hulpmiddel bij het scoren van vermoeidheid," *Nederlands Tijdschrift voor Geneeskunde*, vol. 141, no. 31, pp. 1526–1530, 1997.

[34] T. Van der Molen, B. W. Willemse, S. Schokker, N. H. Ten Hacken, D. S. Postma, and E. F. Juniper, "Development, validity and responsiveness of the clinical COPD Questionnaire," *Health and Quality of Life Outcomes*, vol. 1, p. 13, 2003.

[35] "IBM SPSS Statistics for Windows." [computer program] Version 22.0. Armonk, NY: IBM Corp; 2013.

[36] J. A. Evans and W. A. Whitelaw, "The assessment of maximal respiratory mouth pressures in adults," *Respiratory Care*, vol. 54, no. 10, pp. 1348–1359, 2009.

[37] "ERS - Normal Values. 2011-2019." https://vitalograph.co.uk/resources/ers-normal-values. Accessed: July 26, 2019.

CHAPTER 8

Women's perspective on life after total laryngectomy: a qualitative study

Abstract

Background: Physical and psychosocial challenges are common after total laryngectomy. The surgery leads to lifelong changes in communication, airway, swallowing and appearance. As we move towards health models driven by patient-centred care, understanding the differential impacts of surgical procedures on subgroups of patients can help improve our care models, patient education and support systems. In this article, we discuss the experiences of women following total laryngectomy.

Aims: This study aims to gain insight into the impact of total laryngectomy on women's daily life while identifying their specific rehabilitation needs.

Methods & Procedures: This article is based on in-depth, semi-structured interviews with eight women who had undergone total laryngectomy. These interviews were conducted with women at least one year after they had undergone total laryngectomy, and the participants did not have recurrent disease. Using an interview guide, participants were encouraged to discuss their everyday experiences, while also focusing on issues typical to women. The transcribed

[0]Appeared as: van, Sluis, K.E., Kornman, A.F., van der Molen, L., van den Brekel, M.W.M., and Yaron, G. "Women's perspective on life after total laryngectomy: a qualitative study." *International Journal of Language & Communication Disorders*. 55,2: 188-199, 2020.

interview data were analysed by the method of Thematic Analysis, taking interpretative phenomenological analysis as a lead.

Outcomes & Results: The interviews revealed three main themes: disease and treatment as a turning point, re-establishing meaningful everyday activities, and persistent vulnerability. Participants reported experiencing challenges in their rehabilitation process due to physical disabilities, dependency on others and experienced stigma. Women-specific challenges arose in dealing with the altered appearance and voice, performing care activities, and the spousal relationship (including intimacy).

Conclusions & Implications: Women who undergo total laryngectomy are likely to experience issues in returning to work, the performance of informal care work, the spousal relationship, intimacy, and social interaction due to stigmatization. Medical pre-treatment counselling and multidisciplinary rehabilitation programs should help patients form realistic expectations and prepare them for the changes they will face. A gender- and age-matched laryngectomized patient visitor can contribute to this process. Rehabilitation programs should incorporate the partner and offer psychosocial support for women following total laryngectomy to return to their former roles in family life, social life and work-related activities.

8.1 What this paper adds

8.1.1 What is already known on the subject?

Total laryngectomy is a major surgery, which has a significant impact on affected individuals' quality of life. Patients are confronted with long term consequences including communication problems, altered breathing, swallowing issues, changed appearance, and psychosocial issues. The current, limited body of qualitative literature on the impact of total laryngectomy predominantly features the perspective of men. As a result, little is known about experiences specific to women.

8.1.2 What this paper adds to existing knowledge

Using a qualitative approach, this paper shows that women after total laryngectomy experience problems in returning to work, the performance of informal care work, the spousal relationship, and spousal intimacy. Women dislike their new voice and experience stigma due to their unusual voice and appearance. This study demonstrates the value of peer support in pre-treatment counselling and rehabilitation, especially with gender-matched and age-matched individuals.

8.1.3 What are potential or actual clinical impications of this work?

To adequately prepare female patients undergoing total laryngectomy, they should receive counselling on the impact of the procedure on their everyday life. A gender- and age-matched laryngectomized patient visitor can contribute to this process. The provided counselling and support should involve the patient's partner so that the impact of the surgery on the relationship and family-roles can be discussed. This will allow these females to form realistic expectations, prepare them for the changes they will face, and help them re-integrate into their former roles in family life, social life, and working towards work-related activities.

8.2 Introduction

Total laryngectomy has major consequences for essential physical functions including airway, swallowing, and speech. With the removal of the larynx, the natural voice is lost and patients have to rehabilitate speech, often with help of a voice prosthesis [1]. In addition, patients also have a changed appearance due to the creation of a permanent tracheostoma. The physical alterations lead to ongoing functional issues including reduced physical fitness, problems with food intake, frequent coughing, and communication problems. As a result, patients may be confronted with psychosocial issues [2, 3]. Since survival and complication rates improved over the last decades, scholarly attention has shifted towards providing patients with optimal supportive care [1, 4, 5]. As we move in the direction of health models driven by patient-centred care, understanding the differential impacts of surgical procedures on subgroups of our patients is important to help improve our care models, patient education and support systems. Since the majority of patients undergoing this surgery is male, studies into total laryngectomy are typically male-dominated. However, the proportion of female laryngectomees is rising; therefore, research should explore possible woman-specific needs and issues in laryngectomees.

Quantitative studies investigating gender-specific issues after total laryngectomy show that women might be at risk for a lower global health status, experience a greater impact on their relationship, and have more stigma-related problems [6–9]. Graham and Palmer (2002) show that their studied group of female laryngectomees were younger at time of surgery, were more likely to be working, more often had no partner, had a lower income, had more postoperative physical complaints, and relied more often on support from family and friends instead of support groups compared to their male counterparts [10]. Women run a greater risk of a negative impact of total laryngectomy on their relationship than men after total laryngectomy [8]. Sexual problems and problems with intimacy after total laryngectomy are present in both male and female laryngectomees [8, 11]. The study of Offerman et al. (2015) demonstrates that women experience more sexual problems than their male counterparts do following total laryngectomy. Furthermore, these female laryngectomees report that there is less openness in discussing the consequences of their condition in the family [8]. It is shown that women have a lower global health status and quality of life after total laryngectomy [7]. The procedure might impact women more in social functioning than men because of the lower pitched voice and altered appearance [6].

Qualitative research can contribute in gaining insight in women-specific needs, since it highlights issues and needs from the patient's perspective. However, the number of qualitative studies into the daily impact of total laryngectomy is still limited [12–15]. Existing studies focus on daily functional and psychological difficulties of patients, demonstrating that they struggle with the cancer diagnosis, psychological issues (e.g. anxiety, depression), and adjusting to life after total laryngectomy. Because this body of literature predominantly features the perspective of men, little is known about experiences specific to women. Gardner (1966) first describes the impact of total laryngectomy in affected women. Their focus on proper wifely duties and attitudes is fairly outdated, although it does confirm the presence of gender-specific issues [16]. Dooks et al. (2012) present a single female participant in their qualitative study as a contrast case, since she was the most tearful and depressed [13]. Bickford et al. (2013) suggest that women after total laryngectomy have specific needs, since one of the female participants in their qualitative study indicates that she would appreciate having contact with other young female laryngectomees [12]. Several authors state that future research should address the specific issues and needs of women [6, 12, 13].

Addressing this gap in the literature, the present article aims to provide insight in the everyday impact of total laryngectomy on women, while examining their specific rehabilitation needs. This study uses a qualitative approach to investigate women's experiences following total laryngectomy, with a special focus on the period surrounding the procedure and long-term functional and psychosocial outcomes.

8.3 Methods

This article is based on a qualitative descriptive study focusing on women's experiences following total laryngectomy. Interpretative phenomenological analysis as a qualitative approach is used in order to explore and understand the experience of a particular phenomenon, in this case experiences of women living with total laryngectomy. Braun et al.'s method for thematic analysis was used as a guideline to structure the analysis process [17].

After obtaining ethical approval, possible participants were selected from The Netherlands Cancer institute Antoni van Leeuwenhoek. Inclusion criteria were: female; the total laryngectomy was at least one year ago; and they did not have recurrent disease. There were 32 possible candidates. A maximum sampling strategy was used to include a diverse group of women in terms of, for instance, relationship status, number of years since total laryngectomy, ethnicity. Another consideration was to include participants who had acted as a patient visitor. Eight candidates were approached for the study. They were provided with the participant information form, and contacted one week later to determine intent to participate. All agreed to participate. The number of participants was determined by saturation of the data. Participants were between 60 and 77 years old and were operated 1 to 31 years ago. Two participants had undergone total laryngectomy as a primary treatment, three a salvage procedure, and three due to a dysfunctional larynx after previous treatment. All used a voice prosthesis to communicate. Seven participants were able to speak in fluent sentences, one participant was limited in her verbal communication and had a poor intelligibility. One participant needed nutritional support, and one was limited in her verbal communication and also partly tube feeding dependent. At time of the interview, six respondents had a partner, one was single, and one was a widow (see Table 8.1 for participants' characteristics).

The interviews took place between December 2017 and March 2018, were conducted at respondents' private homes and lasted around 90 minutes. Before the start of the interview participants signed informed consent forms. In four cases the partner was also present during the interview.

The interviews were mostly conducted by two interviewers ((alternating KS, AK and GY). KS, MSc, works as a speech therapist and PhD-student and has a background in health sciences. AK, MSc, works as a speech therapist and junior researcher and has a background in health sciences. KS and AK were already familiar with some of the participants due to their clinical work as speech therapists. MB, MD, PhD, is a head and neck surgeon. LM, PhD, works as a postdoctoral fellow and speech therapist. GY, PhD, works as a postdoctoral fellow and has a background in medical humanities and qualitative methods in healthcare research.

The interviews were conducted using an interview guide, a so-called semi-structured approach (see appendix 1). The guide was developed by deriving topics from the existing qualitative literature [12, 13, 15] and reviewed by an expert panel of healthcare professionals. A Roland Edirol digital recorder and Logitech HD Webcam C510 was used to obtain audio and video recordings of the interviews. Video recordings were used to support intelligibility in case of poor voice outcomes. All interviews were transcribed verbatim. After each interview, participants were asked whether they would be willing to answer follow-up questions via email. All eight participants agreed, three were subsequently approached. Two of them gave a written response.

Table 8.1: Participants included in the study

Pt	**Indication total laryngectomy (TL)**	**Age at TL**	**Current age**	**???**	**Type of speech**	**Intelligibility**	**Diet**	**Tube feeding**	**Highest education**	**Working Before TL / at time of interview**
1	Salvage	67	68	Y	TES	Good	Oral intake, avoid specific food, nutritional support	N	Secondary education	No/No
2	Dysfunctional larynx	71	74	Y	TES	Poor	Oral intake, soft diet	Y	University	No/No
3	Salvage	54	67	Y	TES	Good	Oral intake, normal diet	N	Higher vocational education	Yes/No
4	Primary	47	65	N	TES	Good	Oral intake, avoid specific food	N	Higher vocational education	Yes/Yes
5	Dysfunctional larynx	69	74	Y	TES	Good	Oral intake, avoid specific food	N	Vocational education	No/No
6	Dysfunctional larynx	52	76	Y	TES	Good	Oral intake, avoid specific food	N	Lower education	Yes/No
7	Salvage	29	60	Y	TES	Good	Oral intake, normal diet	N	Secondary education	Yes/Yes
8	Primary	47	62	N*	TES	Good	Oral intake, normal diet	N	Vocational education	Yes/Yes

Abbreviatons: TL = total laryngectomy, TES = tracheo-esophageal speech

*Participant has been married during and after the period of TL, but is widow since one year

The interview data were analysed following the six phases described in the method of thematic analysis [17], while also taking interpretative phenomenological analysis as a lead [18]. This method for the analysis of interviews emphasizes how respondents make sense of their own subjective experiences, in particular those following a life-changing event. In accordance with both methodological approaches, the first step includes familiarization with the data by repeatedly reading the transcriptions. In addition, the interviewers all wrote short reflection reports following each interview (phase 1, [17]). Coding and analysis were initiated by highlighting potential codes together. Weekly meetings were held to discuss project's progress, initial coding, and the analysis (phase 2, [17]). Next, members of the project team each coded a share of the interviews, while also regularly discussing the coding process together. After going through the first two interviews by hand, NVivo software (v11) was used to structure the coding process, creating a long list of codes along the data set. Inter-coder reliability was monitored by keeping a master list of all codes, tracking all changes, and regularly discussing these. Next, in an iterative process encompassing several joint sessions, codes were sorted into potential themes (phase 3, [17]). To visualize relationships between codes and themes, a 'mind map' was created (phase 4, [17]). This process resulted in a set of three fully worked-out themes, each encompassing three sub-themes (see Figure 8.1) (phase 5, [17]). Themes cover recurring issues, as well as issues that respondents indicated were meaningful to them. The themes and categories were checked against the entire data set by AK. The report of the analysis served as a basis for the present article (phase 6, [17]). After finishing the first draft of the report, a two-page summary was created to perform participant checking. The participants received this summary by mail and were invited to respond via a response form or telephone. No comments were provided by the participants.

8.4 Results

Three main themes were identified: "Disease and treatment as a turning point", "Re-establishing meaningful everyday activities", and "Persistent vulnerability" (visualized in Figure 8.1). This section discusses each theme. The first theme is presented more briefly due to overlap with existing qualitative literature on the experiences of individuals undergoing total laryngectomy [12, 13, 15].

8.4.1 Disease and treatment as a turning point

This first theme highlights the fact that respondents experienced their disease and its treatment as a turning point in their lives. This shift already started before the procedure, when participants were first confronted with symptoms and functional issues, received their diagnosis, and were offered a treatment plan. When presented with the treatment plan, some respondents initially felt an aversion to total laryngectomy:

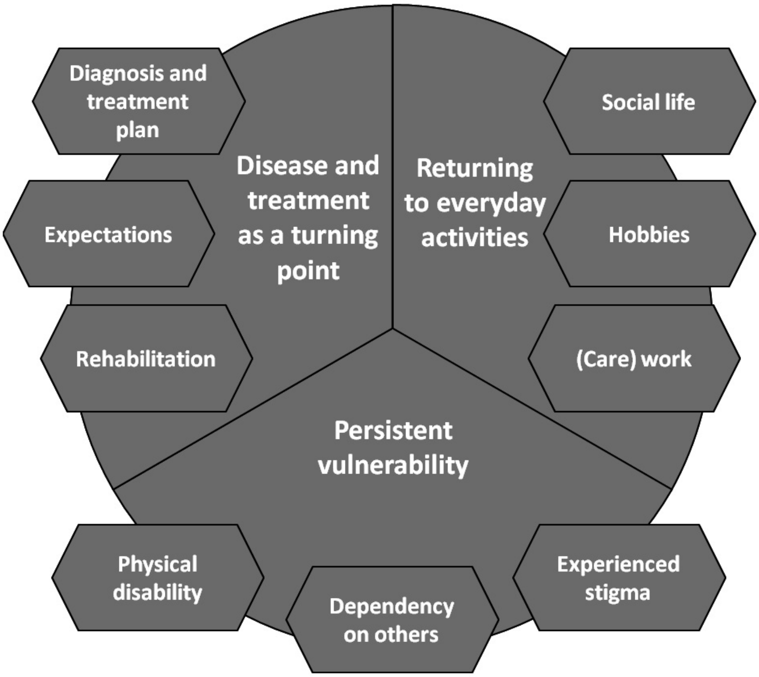

Figure 8.1: Three main themes and sub-themes that resulted from the analysis of the interviews.

> I used to say: 'I'll have everything [medically necessary] done but not that.' Because I had seen people [with total laryngectomy] walk around in the hospital, and then I said: 'That's not for me.' I'm not going to walk around with a hole in my throat. (Participant 5)

Some needed to go into surgery quickly, which was distressing and required last-minute practical arrangements. One participant spoke about the impact the procedure had on her family life; her adolescent children and partner had to make a lot of effort to combine normal family life, school, and work with visiting the hospital and caring for their mother/wife.

Participants' expectations regarding the surgery and its outcome were shaped by the preoperative counselling they received in the hospital, information provided by patient visitors (a laryngectomized individual), and information they had sought themselves. Two respondents said the counselling they received from the hospital was too technical; they missed receiving information about the daily implications of having total laryngectomy. All participants remembered their pre-operative meeting with the laryngectomized patient visitor vividly.

Many were relieved to meet someone who had undergone the surgery, and could still speak, eat and function well. This provided a positive, hopeful image of life after total laryngectomy. Three respondents stated that this meeting convinced them to undergo the procedure. However, some respondents indicated that, in hindsight, the laryngectomized patient visitor had provided an overly positive image of the surgery's outcome and that they neglected to discuss ongoing difficulties. Most emphasized the importance of meeting an individual who had already undergone the surgery to obtain an understanding of the procedure's outcomes (e.g. voice, functioning).

During their stay in the hospital and the first period at home, respondents experienced challenges in adjusting to their altered body. After the surgery, when recovering on the hospital's ward, they were not able to speak. To compensate, respondents used mouthing, gestures, and writing to communicate. They indicated that being unable to speak can be severely distressing because it interferes with the capacity to communicate one's concerns clearly and in real time. Before and during standard medical procedures (e.g. suctioning mucus out of a tracheostoma), participants experienced little opportunity to ask questions, and their voiceless expressions of discomfort were brushed aside. Learning to speak again was often a physical and emotional challenge. All respondents disliked the sound of their new voice; they described their voices as “unfeminine”, “unpleasant,” “gruff”, “scratchy” or “grumpy”. Functional issues respondents mention included reduced physical fitness, the need for use of medical aids, swallowing difficulties, and managing mucus discharge from the tracheostoma.

8.4.2 Returning to everyday activities

The second theme captures respondents' experiences as they tried to return to everyday activities after recovering sufficiently. Their re-integration into everyday life took place in three main domains: social life, hobbies, and (care) work.

Social life

Respondents described the impact of total laryngectomy on the relationship with their partner. Married respondents mostly experienced their husband as helpful and understanding, although they described becoming more dependent on their spouse, as the total laryngectomy reduced their ability to fulfil care activities such as cooking and looking after (grand)children. This led to shifts in the relationship's give-and-take balance, alterations in familiar role patterns, and changes in expectations between laryngectomee and partner:

> I came home quite disabled, (. . .) I couldn't speak well [and experienced] much coughing. [So, my husband] really thought: ‘my God, I‘m stuck with an old woman’. (. . .) At one point I thought: I am

> going to leave him. (...) I refused to be a wife who is just being tolerated. (...) Like some kind of glorified housemaid, you know. No way. Take it or leave it. (Participant 7)

Three respondents discussed their experiences regarding sexual intercourse. One respondent started having intercourse again after bringing this subject up with her husband; she stated that she needed intimacy. Two respondents no longer had sexual intercourse after total laryngectomy. One indicated that she did not feel comfortable to have sex because of her changed voice and the risk of coughing. The single participant had her total laryngectomy at the age of 45 and did not find a partner after surgery. She started to date after her rehabilitation period. Before the procedure, she experienced being quite popular with the opposite sex. After being rejected repeatedly because of her voice after the surgery, she stopped dating. One respondent gave birth to two children after her surgery. She discussed her wish to conceive with her specialist, since she worried about her physical abilities (e.g. push during labour). In hindsight, she indicates that she did not experience any problems from her condition during this period, except for the voice prosthesis which was dislocated due to hormonal changes.

Participants experienced changes in social contact with unfamiliar people. Participants frequently experienced negative reactions or assumptions regarding their voice, in direct interactions as well as over the telephone. Many respondents discussed being self-conscious about their appearance and how they present themselves to others, by covering their stoma, for example. Some mentioned listeners associate the sound of their altered voice with being grumpy or interpret it as masculine. As one participant stated:

> the post office about a delivery, [they'd say]: 'Sir, stop complaining.' And then when I would come pick it up, they would say: 'Oh...'. They always think I am a man [before seeing me]. (...) People [also] treat me gruffly. Then, when they see me in real life, they suddenly become very friendly. (Participant 4) (When I would telephone)

Among family and friends, participants often experienced understanding and support. Nevertheless, these intimates also needed to get used to the alterations following total laryngectomy. In particular, respondents' new method of speaking, which is slower, less clear, and less loud, affected their ability to interact with others. All respondents experienced limitations in social situations that require shouting, whispering, singing, reading out loud, providing quick retorts, or expressing emotions vocally (e.g. laughing, crying). Consequently, some participants eventually withdrew from social situations that involved groups:

> is different when everyone is over here [at my house] for my birthday. You barely hear me talking then. I'll just be sitting there, watching everything. It is harder for me to be intelligible [in a group]. [So] I just don't [speak]. (Participant 1) (Communicating)

The total laryngectomy also led to new social contacts, especially with peers (individuals who had undergone the same surgery). The respondent who had her surgery at age of twenty-nine realized that she could not relate to the experience of elderly laryngectomized individuals when she visited a patient meeting:

> I went [to a patient meeting] twice. First of all [. . .], they were all much older than me. Discussions were like: "My grandchildren will not hear my own voice". And I thought: "Well, come on, my children have never heard my voice. You know what, [this setup] didn't fit my experiences." (Participant 7)

Hobbies

To take up their former hobbies, participants either made practical adjustments so they could still perform these after total laryngectomy or took up new leisure activities. One respondent modified her boat to decrease the risk of falling into the water. Many participants developed new interests and hobbies that would be easy to engage in despite their condition (e.g. sewing clothes, making necklaces to cover the stoma, gardening). Two participants joined a special choir for head and neck cancer patients. This enabled them to enjoy singing and performing and to have fun with their voice again.

(Care) work

Four respondents had a job before the total laryngectomy, but none of them was able to return to it following their surgery since these positions required much speaking (e.g. teacher, secretary, saleswoman). Two participants actively sought a job after their total laryngectomy and got rejected repeatedly during the application process.

> The first years after my laryngectomy I still applied [for vacancies] (. . .). The first years, I thought [my condition] wasn't too bad. I didn't realize how severe it was. How many disabilities I had (. . .). After five or six years I stopped [applying], thinking: 'This isn't going to happen'." (Participant. 4)

All respondents of working age were declared unfit to work. Nevertheless, five respondents started work-related activities. Some started volunteering in their community or within the Dutch organization for head and neck cancer patients (e.g. patient-visitor, providing lessons). They valued their ability to help others by sharing their own experiences. Others set up their own business in order to remain financially secure; one provided painting lessons to small groups, and one started her own bed and breakfast.

Most participants were still able to perform (some) informal care-work, including household activities and taking care of others. Nevertheless, those who required assistance from their husband, relatives or a household help reported they had difficulties transitioning from being the care provider to being its recipient. They found it hard to delegate or outsource tasks such as cooking or cleaning:

> Taking care of the household and not wanting to delegate [tasks] (...) is really very difficult for me. (...) I now need a whole week for things that previously took me one day to accomplish. (Participant 5)

Respondents who regularly took care of children, especially (grand)mothers, mentioned that it was difficult for them to warn a child in a dangerous situation, and experienced challenges while singing or reading books. Just calling out to capture a child's attention in public also proved potentially embarrassing—and therefore more challenging:

> And with my kids, (...) at the shops, I would clap my hands and then they'd know: 'Oh, mama is calling'. Because I didn't want to [shout] then, in such a shopping mall. (...) You really don't want to stand out [with your unusual voice], then. (Participant 7)

8.4.3 Persistent vulnerability

The third theme identified in the interviews was persistent vulnerability. Although participants found various ways to return to the daily activities, they experienced vulnerability, including physical disabilities, dependence on others for support, and stigma.

Physical disabilities

Participants were confronted with a number of ongoing physical changes in breathing, swallowing, speaking and appearance. The physical limitations of the tracheo-esophageal voice made respondents apprehensive about situations where they were dependent on their verbal abilities. One participant noted that she feared being unable to call for help during an emergency due to living alone:

> It is a lot harder when you're single; (...) being unable to talk is very frightening (...) when you're home alone and thinking that you're choking or something (Participant 4)

All respondents reported consciousness of dealing with water in daily settings, such as being on a boat, cycling near open water or taking a shower.

Dependency on others

Due to the changes after total laryngectomy described above, some participants reported a continuous dependency on their social network and on healthcare. In some situations, such as using the telephone, participants needed family members or partners to speak for them. One participant spoke about the support she required in caring for her young children:

> When we would go swimming with [my] kids, I always took a friend with me, because well, I had to be able to send someone into the water [if necessary] since I can't do it myself. (Participant 7)

Some participants received financial support from relatives. Others remained dependent on family members and friends for transport to the hospital or other forms of medical care. Respondents reported a dependence on healthcare as well, since they kept requiring medical support, e.g. managing problems with the stoma and the voice prosthesis. When travelling, participants preferred being treated in a specialized centre, since healthcare professionals are not always familiar with total laryngectomy. For a number of respondents, this all meant they were hesitant to leave town or go abroad.

Experienced stigma

Participants reported experiencing stigma regularly, especially in the form of unwanted attention: other people staring at their stoma, making (offensive) comments or asking (intrusive) questions. One participant shared her experiences of being bullied by neighbourhood children:

> (...) Children ringing the doorbell, [shouting]: 'Witch, witch, you cannot talk!' (. . .) 'Say something, say something!' (...) To me that is awful. (Participant 4)

The respondents assumed that this unwanted attention resulted from other people's unfamiliarity with total laryngectomy. Managing unwanted attention was bothersome or difficult for some participants.

When asked directly, most participants indicated that a total laryngectomy impacts men and women in similar ways due to the physical issues involved. Nevertheless, some respondents proposed that the challenges affected women face in daily life are bigger, especially in social situations. They presumed women speak more and are judged on their appearance and voice more than men. Participants gave examples of women-specific stigma they had experienced, including being taken for a man over the telephone and receiving comments about how they look or about their low-pitched voice.

8.5 Discussion

As the stories of women with total laryngectomy reveal, both their disease and its treatment formed a turning point in their lives. Respondents made efforts to re-integrate into the social communities they belonged to before their surgery and return to former everyday activities. All found various ways to participate in daily life again. Nevertheless, participants remain persistently vulnerable after total laryngectomy, because of the effort required by foreseeing and managing functional, social, and health-related issues. As a result, participants are required to continuously manage their disabilities as well as other people's responses to them. The disease and its treatment mean that both body and self are radically and irrevocably altered.

This study explored the perceptions of females on their life after total laryngectomy and shows the presence of women-specific issues. Interestingly, when asked directly, most participants did not think there were substantial differences between the experiences of men and women after total laryngectomy. Similarly, Graham and Palmer (2002) found that responses of men and women after total laryngectomy were more similar than dissimilar [10]. The study of Lee et al. (2010) showed that females following total laryngectomy had significantly lower global health status than males and lower levels of physical, emotional, cognitive and social functioning [7]. Our findings illustrate that some of the challenges participants experience are due to their inability to adhere with societal expectations regarding feminine roles and activities (e.g. taking care of others and the home, looking attractive, having a high pitched voice). Many respondents discussed being self-conscious about their appearance and how they present themselves to others. In addition, they all disliked the sound of their voice, which several described as unfeminine. Our findings herein therefore corroborate Cox et al.'s (2015) hypothesis that changes in voice, sound, and speech quality after total laryngectomy lead to a loss of femininity [6].

Our study also demonstrates that the relationship of women with their partner is influenced by the total laryngectomy. Cancer treatment in general impacts the spousal relationship [19, 20]. Offerman et al. (2015), showed that female laryngectomees run a greater risk of a negative impact of total laryngectomy on their relationship than their male counterparts [8]. Specifically, our results show that the surgery and its aftermath leads to a noticeable shift in give-and-take balance between partners. This shift is presumably associated with the fact that women in a heterosexual relationship still mainly perform the majority of household chores [21, 22].

In line with this point, this study also revealed women-specific issues in participants' performance of informal care-work. Some respondents found it hard to depend on others who cared for them or delegate household chores after the surgery. Problems with informal care-work were also a recurring topic in interviews for participants who were (grand)mothers. Earlier studies on the impact of total laryngectomy on affected individuals did not investigate how care-work activities take shape in daily life after the procedure. Again, since the majority of child-care is still performed by women [23], (informal) care-work is a domain in which gender differences are still very much relevant. To further investigate the impact of total laryngectomy on women therefore required addressing both shifts in the spousal relationship and care-work.

Issues pertaining to femininity also arose in the areas of intimacy and sexuality. Research shows that libido and sexual problems are common after treatment for laryngeal cancer in both male and female patients [11]. Singer et al. (2008) showed that having reduced physical strength ("not enough stamina") rather than the changes in appearance and/or the sputum are considered as most problematic for having sexual intercourse [11]. By contrast, our study suggests that the issues women experience herein involve both their altered appearance and functioning, i.e. physical disabilities. Thus, our respondents worried that the noises their stoma produces and their coughing fits might be unattractive to their partner or interfere with the intercourse. Similarly, a Offerman et al.'s (2015) study on the impact of total laryngectomy on the spousal relationship showed that female laryngectomees experience more deterioration in their sexual relationship and more sexual problems when compared to their male counterparts [8]. Further research should explore the role of appearance, functionality and physical strength in reduced sexuality in male as well as female patients.

The women in our study who were of working age could not return to their former job after surgery. After a period of recuperation and adaption, most wanted to re-integrate into work activities nevertheless. In the end, two respondents started a business, and three turned to volunteer work (one combined the two). In a study conducted by Graham and Palmer (2002), the female laryngectomy participants included were younger at time of surgery [10]. As a consequence, more females compared to males were working or on disability leave after surgery [10]. This highlights the importance of returning to work-related activities for women following total laryngectomy. Well's et al.'s (2013) systematic review of qualitative studies exploring return to work after cancer demonstrates that successfully returning to work depends on shifts and adjustments in each aspect of what is already a complex set of factors at the individual, organizational and societal level [24]. Successful return to work after total laryngectomy might be influenced by former work activities, functional outcomes and adjustment, and age at total laryngectomy. As our results suggest, it also depends on gender. Our respondents—and women generally—are still predominantly employed in areas such as teaching, administration, or care, which involve communication and service-orientation.

All eight respondents in this study have (had) contact with peers. Peer support is highly satisfying for people with cancer [25, 26]. Peer support provided by a laryngectomized patient visitor can play an important role in both information giving and counselling before surgery [1, 27, 28]. Indeed, all eight participants of this study remembered the pre-operative meeting with a laryngectomized patient visitor vividly. Although the female participants of our study did not explicitly indicate that they would have liked to meet a female patient, we suggest such a match might aid prospective women laryngectomees form realistic expectations concerning voice quality, appearance, and other outcomes that may affect women differently than men. Similarly, one respondent who received her surgery at age 29 indicated she could not relate to elderly peers. Earlier research, too, shows that younger female laryngectomees prefer contact with younger peers [12].

8.5.1 Recommendations for clinical practice

The insights provided by this study may help improve clinical practice by enabling a more patient-centred approach in pre-treatment counselling and rehabilitation. As we move towards health models driven by patient-centred care, understanding the differential impacts of surgical procedures on subgroups of our patients, in this case women, is critically important to help improve our care models, patient education and support systems. We support the conclusions of earlier research recommending personalized rehabilitation provided by a multidisciplinary team and recognize the presence of unmet supportive care needs in women undergoing total laryngectomy [2, 3, 7, 9, 12, 29–31].

First, we recommend that standard pre-treatment counselling procedures be re-evaluated. Counselling involves exploring patients' priorities, values and concerns about the treatment and the outcome, as well as possible benefits, risks and implications of the treatment options [27, 28, 32]. However, we recommend that the information given also includes expected everyday issues after the procedure, specifically as experienced from a patients' point of view. This implies providing information on gender-specific issues. It therefore seems prudent to match patients and patient visitors based on both gender and age to facilitate realistic expectations and enhance peer support possibilities.

Secondly, we recommend tailored rehabilitation programs following total laryngectomy, with attention to gender-specific issues. Our study suggests that women face specific problems while reintegrating on an individual as well as on a societal level after total laryngectomy. Lee et al. (2010) propose that women may particularly benefit from rehabilitation programs which aim to improve emotional and social functioning post-laryngectomy, since they found this to be significantly impacted in female laryngectomees. Our study shows the persistent vulnerability on the long term including the dependence on others, stigma-related problems, and problems with returning to the former job. Thus, rehabilitation programs should include psychosocial and practical support to pursue participation in social activities and work-related activities.

Thirdly, it is important to involve the partner in the treatment and rehabilitation process. Women following total laryngectomy experience a greater impact of the procedure on their relationship than men [8]. Our study specifies this with the reported shift in give-and-take balance of the spousal relationship, challenges in performing care work, and problems with intimacy. We therefore highlight the importance of involving partners in the process of rehabilitation and recommend that clinicians create openness in discussing the impact of the surgery on the relationship and family roles. This will allow females following to re-integrate into their former roles in as well family, social life as in work.

This study has some limitations. Qualitative research in general is never representative for the entire population, because it draws on the experiences of a relatively small sample. Also, since only women were recruited, experiences specific to men were not addressed in this study. However, the main themes "Disease and treatment as a turning point", "Returning to everyday activities" and ”Persistent vulnerability” are likely to occur with men as well. Furthermore, the study necessarily included participants who had a relatively high level of speaking proficiency with their voice prosthesis and were therefore able to express themselves verbally. Despite this bias, our findings illustrate a wide range of issues beyond communication. This paper therefore provides an impression of an extensive range of possible issues women may be confronted with after total laryngectomy.

8.6 Conclusions

The accounts provided by the women interviewed in this study show that undergoing total laryngectomy has a major impact on their everyday life. The disease and its treatment form a turning point in their lives, after which they experience challenges in while returning to everyday activities and persistent vulnerability. Women-specific issues are present in returning to work, the performance of informal care work, the spousal relationship, intimacy, and social interaction (mostly stigmatisation due to voice and appearance). The aim of this study was to reveal women-specific issues, but we nevertheless assume that some of the reported issues are present for men as well. Based on the findings of this study it is recommended to include information on changes in daily functioning in the pre-treatment counselling procedures, and the contribution of an aged-matched and gender-matched patient visitor would be valuable. Tailored rehabilitation for women following total laryngectomy is recommended to enhance participation in social roles and work-related activities. Within this process of treatment and rehabilitation, it is important to involve the partner in order to discuss the impact of the surgery on the relationship and family roles. To fully understand the gender-specific aspects of total laryngectomy, future research should focus on the particular ways in which the procedure affects both women and men in specific areas of daily life (e.g. returning to work, care activities, and intimate relationships).

8.7 Acknowledgements

The Netherlands Cancer Institute received a research grant from Atos Medical (Malmö, Sweden), which contributes to the existing infrastructure for quality of life research of the Department of Head and Neck Oncology and Surgery. The authors have no other funding, financial relationships or conflicts of interest to disclose.

References

[1] L. van der Molen, A. F. Kornman, M. N. Latenstein, M. W. van den Brekel, and F. J. Hilgers, “Practice of laryngectomy rehabilitation inter-

ventions: a perspective from Europe/the Netherlands," *Current Opinion in Otolaryngology & Head and Neck Surgery*, vol. 21, no. 3, pp. 230–238, 2013.

[2] J. M. Bickford, J. Coveney, J. Baker, and D. Hersh, "Support following total laryngectomy: Exploring the concept from different perspectives," *European Journal of Cancer Care*, vol. 27, no. 3, p. e12848, 2018.

[3] A. Perry, E. Casey, and S. Cotton, "Quality of life after total laryngectomy: functioning, psychological well-being and self-efficacy," *International Journal of Language & Communication Disorders*, vol. 50, no. 4, pp. 467–475, 2015.

[4] V. M. Rosa, J. M. L. Fores, E. P. F. da Silva, E. O. Guterres, A. Marcelino, P. C. Nogueira, W. R. M. Baia, and M. A. V. Kulcsar, "Interdisciplinary interventions in the perioperative rehabilitation of total laryngectomy: an integrative review," *Clinics*, vol. 73(Suppl. 1), p. e484s, 2018. Epub September 06, 2018.

[5] J. Zenga, T. Goldsmith, G. Bunting, and D. G. Deschler, "State of the art: Rehabilitation of speech and swallowing after total laryngectomy," *Oral Oncology*, vol. 86, pp. 38–47, 2018.

[6] S. R. Cox, J. A. Theurer, S. J. Spaulding, and P. C. Doyle, "The multidimensional impact of total laryngectomy on women," *Journal of Communication Disorders*, vol. 56, pp. 59–75, 2015.

[7] M. T. Lee, S. Gibson, and K. Hilari, "Gender differences in health-related quality of life following total laryngectomy," *International Journal of Language & Communication Disorders*, vol. 45, no. 3, pp. 287–294, 2010.

[8] M. P. Offerman, J. F. Pruyn, M. F. de Boer, J. J. Busschbach, and R. J. Baatenburg de Jong, "Psychosocial consequences for partners of patients after total laryngectomy and for the relationship between patients and partners," *Oral Oncology*, vol. 51, no. 4, pp. 389–398, 2015.

[9] F. Jansen, S. E. J. Eerenstein, B. I. Lissenberg-Witte, C. F. v. Uden-Kraan, C. R. Leemans, and I. M. V. Leeuw, "Unmet supportive care needs in patients treated with total laryngectomy and its associated factors," *Head & Neck*, vol. 40.12, pp. 2633–2641, 2018.

[10] M. S. Graham and A. D. Palmer, "Gender difference considerations for individuals with laryngectomies," *Contemporary Issues in Communication Science and Disorders*, vol. 29, pp. 59–67, 2002.

[11] S. Singer, H. Danker, A. Dietz, U. Kienast, F. Pabst, E. F. Meister, J. Oeken, A. Thiele, and R. Schwarz, "Sexual problems after total or partial laryngectomy," *The Laryngoscope*, vol. 118, no. 12, pp. 2218–2224, 2008.

[12] J. Bickford, J. Coveney, J. Baker, and D. Hersh, "Living with the altered self: a qualitative study of life after total laryngectomy," *International Journal of Speech-Language Pathology*, vol. 15, no. 3, pp. 324–333, 2013.

[13] P. Dooks, M. McQuestion, D. Goldstein, and A. Molassiotis, "Experiences of patients with laryngectomies as they reintegrate into their community," *Supportive Care in Cancer*, vol. 20, no. 3, pp. 489–498, 2012.

[14] J. Mertl, E. Zackova, and B. Repova, "Quality of life of patients after total laryngectomy: the struggle against stigmatization and social exclusion using speech synthesis," *Disability and Rehabilitation: Assistive Technology*, vol. 13, no. 4, pp. 342–352, 2018.

[15] B. J. Noonan and J. Hegarty, "The impact of total laryngectomy: the patient's perspective," in *Oncology Nursing Forum*, vol. 37, pp. 293–301, 2010.

[16] W. H. Gardner, "Adjustment problems of laryngectomized women," *Archives of Otolaryngology*, vol. 83, no. 1, pp. 31–42, 1966.

[17] V. Braun, V. Clarke, and G. Terry, "Thematic analysis," *Qualitative Research in Clinical Health Psychology*, vol. 24, pp. 95–114, 2014.

[18] J. A. Smith, P. Flowers, and M. Larkin, *Interpretative Phenomenological Analysis: Theory, Method and Research.* London: Sage, 2009.

[19] Q. Li and A. Y. Loke, "A literature review on the mutual impact of the spousal caregiver–cancer patients dyads:'communication','reciprocal influence', and 'caregiver–patient congruence'," *European Journal of Oncology Nursing*, vol. 18, no. 1, pp. 58–65, 2014.

[20] C. Pitceathly and P. Maguire, "The psychological impact of cancer on patients' partners and other key relatives: a review," *European Journal of cancer*, vol. 39, no. 11, pp. 1517–1524, 2003.

[21] G. N. Powell and J. H. Greenhaus, "Sex, gender, and decisions at the family → work interface," *Journal of Management*, vol. 36, no. 4, pp. 1011–1039, 2010.

[22] B. A. Shelton and D. John, "The division of household labor," *Annual review of sociology*, vol. 22, no. 1, pp. 299–322, 1996.

[23] N. Folbre, "Measuring care: Gender, empowerment, and the care economy," *Journal of Human Development*, vol. 7, no. 2, pp. 183–199, 2006.

[24] M. Wells, B. Williams, D. Firnigl, H. Lang, J. Coyle, T. Kroll, and S. MacGillivray, "Supporting 'work-related goals' rather than 'return to work' after cancer? a systematic review and meta-synthesis of 25 qualitative studies," *Psycho-Oncology*, vol. 22, no. 6, pp. 1208–1219, 2013.

[25] L. M. Hoey, S. C. Ieropoli, V. M. White, and M. Jefford, "Systematic review of peer-support programs for people with cancer," *Patient Education and Counseling*, vol. 70, no. 3, pp. 315–337, 2008.

[26] E. Watson, "The mechanisms underpinning peer support: a literature review," *Journal of Mental Health*, pp. 1–12, 2017.

[27] E. Fitzgerald and A. Perry, "Pre-operative counselling for laryngectomy patients: a systematic review," *Journal of Laryngology and Otology*, vol. 130, no. 1, pp. 15–20, 2016.

[28] N. Raol, E. Lilley, Z. Cooper, J. Dowdall, and M. A. Morris, "Preoperative counseling in salvage total laryngectomy: Content analysis of electronic medical records," *Otolaryngology–Head and Neck Surgery*, vol. 157, no. 4, pp. 641–647, 2017.

[29] T. Eadie, M. Kapsner-Smith, S. Bolt, C. Sauder, K. Yorkston, and C. Baylor, "Relationship between perceived social support and patient-reported communication outcomes across communication disorders: a systematic review," *International Journal of Language & Communication Disorders*, vol. 53.6, pp. 1059–1077, 2018.

[30] E. Passchier, M. M. Stuiver, L. van der Molen, S. I. Kerkhof, M. W. van den Brekel, and F. J. Hilgers, "Feasibility and impact of a dedicated multidisciplinary rehabilitation program on health-related quality of life in advanced head and neck cancer patients," *European Archives of Oto-Rhino-Laryngology*, vol. 273, no. 6, pp. 1577–1587, 2016.

[31] G. Sharpe, V. C. Costa, W. Doubé, J. Sita, C. McCarthy, and P. Carding, "Communication changes with laryngectomy and impact on quality of life: a review," *Quality of Life Research*, vol. 28, no. 4, pp. 863–877, 2019.

[32] O. Laccourreye, D. Malinvaud, F. C. Holsinger, S. Consoli, M. Ménard, and P. Bonfils, "Trade-off between survival and laryngeal preservation in advanced laryngeal cancer: the otorhinolaryngology patient's perspective," *Annals of Otology, Rhinology & Laryngology*, vol. 121, no. 9, pp. 570–575, 2012.

CHAPTER 9

General discussion

The typical clinical trajectory for total laryngectomy patients is oriented towards curative treatment, restoring communication, and reducing functional issues patients have to cope with. In this thesis we investigated the multidimensional impact of this surgery by exploring voice and speech outcomes (section 1) and functional issues and well-being (section 2) following total laryngectomy. Within this discussion, the implications of the studies are discussed, and recommendations for further research and clinical practice are made.

9.1 Voice and speech outcomes

9.1.1 Summary of the findings

Speech rehabilitation methods

With the systematic literature review in **Chapter 2**, we investigated the different types of speech rehabilitation methods after total laryngectomy and report objective and subjective outcome measures. For several decades detailed studies on voice outcomes have been published, investigating acoustic and perceptual outcomes. Most studies include small groups of participants and there is a difference in offered speech rehabilitation methods. The earlier studies primarily show mostly outcomes on esophageal speech and electrolarynx speech, while more recent studies do evaluate tracheoesophageal speech.

The studies included in this systematic literature review show that tracheoesophageal speech offers the most natural speech result after total laryngectomy. Nevertheless, values for acoustic, perceptual, and speech related quality of life (QoL) are abnormal compared to healthy laryngeal speakers. Tracheoesophageal speech has favourable outcomes on acoustic features as fundamental frequency, loudness and maximum phonation time. For both tracheoesophageal speech and esophageal speech the sound source is the PE-segment. In tracheoesophageal speech, the mucosal vibrations of the neoglottis are driven by the lungs. It is likely that this pulmonary air stream leads to more fluency, longer phonation time and louder speech, resulting in a better acoustic and perceptual outcome. For electrolarynx speech it is not relevant to perform acoustic analysis, since fundamental frequency and loudness are mechanically produced.

Results in speech-related QoL were variable between the different speech rehabilitation methods that are offered. There are only few studies which compare different groups of speakers and within these studies different instruments are used. Several of them have only included excellent speakers, which leads to an inclusion bias. As a consequence, we had to conclude that in all three groups of speakers abnormality is observed in their reported speech-related QoL. Training period and satisfaction of the reached result might be of influence in the reported handicap. It is the case that, esophageal speakers have to train for a long period, most patients need six months to accomplish intelligible and fluent speech. Therefore, when they are evaluated after this training period they are proud and satisfied with their accomplishment. In contrast, electrolarynx speech is mostly possible in one session, although intelligibility issues occur. The training period for tracheoesophageal speech differs from a few sessions up to several months of training with a speech language pathologist (SLP). The amount of needed training and level of succes might have an influence on reported speech- related QoL.

Prospective evaluation of voice outcomes

In **Chapter 3** we prospectively evaluated voice and speech outcomes from before up to one-year post-laryngectomy. This study has been conducted over two countries and five hospitals; all patients who were admitted to total laryngectomy in the study time frame have been approached for inclusion. Aiming to overcome an inclusion bias by only including excellent speakers, we chose to prospectively assess patients over time. The study shows the variation between speakers. Voice outcomes decrease from pre- to post-surgery, which is seen in the AVQI scores and perceptual outcomes. Voice rehabilitation is insufficient in most participants who decease within the first year following surgery. This means that they often have no verbal communication method in the palliative stage of their life. Nevertheless, the study participants who were included in the study at one year post-surgery do reach acceptable tracheoesophageal speech, this was the case in 20 of the 22 included participants at the 12 months post-surgery assessment acquired tracheoesophageal speech. This is also reflected in the speech-related QoL outcomes, and overall reported QoL, which improves in the long term.

Intelligibility issues in tracheoesophageal speech

In **Chapter 4** we investigated the intelligibility issues in tracheoesophageal speech by exploring acoustic features. This dedicated acoustic study consists of a within-subjects study design in which participants are studied pre- and post-surgery. The study shows that the acoustic contrast is reduced after surgery, with the features of the word-initial, voiced, /d/, becoming more like, unvoiced, /t/. The presence of prevoicing decreases and the burst-duration of this initial consonant increases. Thus, the difference between word-initial /d/ and /t/ becomes smaller after treatment, which explains a part of the intelligibility issues which occur in this patient population. With the dataset created with our prospective study, valuable speech data including pre- and post-laryngectomy recordings are obtained. This can help future work to further investigate intelligibility issues, as well as develop clinical applications and interventions to enhance understand ability of tracheoesophageal speech.

Long-term stability of tracheoesophageal speech

In **Chapter 5** we assessed long-term speech outcomes for tracheoesophageal speech. With the ongoing improvements in long-term oncological treatment, survival rates increase. Therefore, patients who have undergone total laryngectomy have to speak a substantial number of years with tracheoesophageal speech. The study on long-term voice outcomes of tracheoesophageal speakers in chapter 5 shows the long-term stability of this type of speech. On average, a slight decrease in voice quality and intelligibility is seen after a follow -up 7 to 8 years after total laryngectomy, however some participants showed better voice quality and intelligibility outcomes on the long-term. The general slight decrease of voice quality and intelligibility might be the effect of aging factors. A lot remains to be learned about the impact of aging on voice quality and the (details about) physiological changes of the PE-segment for total laryngectomy patients.

With the general increased age expectancy, there can also be co-morbidities in total laryngectomy patients, which affect speaking as well. All morbidities which affect general fitness, pulmonary functioning, or posture are likely to influence the capability of the laryngectomized person to use tracheoesophageal speech.

9.1.2 Measuring voice and speech outcomes after total laryngectomy

To adequately measure and monitor voice and speech outcomes following total laryngectomy, assessment tools are needed. We will discuss the instruments used in the various studies which are included in this thesis, their usability, reliability and validity in evaluating speech after total laryngectomy.

Automated acoustic outcome measures

The studies included in the systematic literature review presented in Chapter 2 use a great number of acoustic outcome measures. Therefore, in this thesis we chose to evaluate a range of traditional acoustic measures including fundamental frequency (F0), jitter, shimmer, intensity, and spectral tilt, as well as signal-to-noise ratio computations including Harmonics to Noise Ratio (HNR), and percentage of voicedness (%voiced) and maximum phonation time (MPT) [1, 2]. These outcome variables are frequently used in the literature, although the validity and clinical utility of many of these acoustic measures has been strongly debated [3–5]. Some studies discussed in Chapter 2only included excellent speakers and excluded audio samples of tracheoesophageal speakers which were too aperiodic to perform acoustic analysis [6–10]. With the aperiodic characteristics of tracheoesophageal speech, the use of these standard acoustic outcomes is debatable [1, 2, 11, 12]. It is troublesome to compare groups of speakers (over time, between studies etc.) with this high number of acoustic parameters. Each study applies different assessment protocols, recording equipment, and software for analysis which reduces the comparability of outcomes between studies.

In chapter 3 and chapter 5, we therefore decided to use the Acoustic Voice Quality Index. To date, AVQI has not often been utilised for evaluating voice quality after total laryngectomy, however it provides several benefits in structured voice evaluation and therefore is a promising tool. Since it produces one outcome score rather than many individual scores, it has a better usability and comparability within the subject as well as in-between subjects' study designs. Chapter 3 shows the need for separate validation of this scale for use in tracheoesophageal speech. The AVQI scale has originally been developed to assess dysphonia in laryngeal speakers [5]. In our prospective study we show that patients admitted to total laryngectomy already have abnormal AVQI values before surgery, which worsen after surgery when evaluating the tracheoesophageal speech samples. The AVQI scores do correlate well with the perceptual rated voice quality and intelligibility scores. Therefore, AVQI provides consistent information on both perceived voice quality and intelligibility.

Patient-reported outcome measures

In this thesis chapters 2 and 3 include patient-reported outcome measures (PROMs). With the help of these PROM questionnaires the patients' functioning has been assessed. This is an essential part in evaluating communicative functioning, including voice problems and intelligibility issues [13]. The studies in the systematic review show the use of different questionnaires, including VR-QOL, VHI-30, and VHI-10. None of these questionnaires have been specifically validated for total laryngectomy patients, but since these questionnaires are the best available instruments and widely used in dysphonia research, they are also used often in this patient group. However, there is a need for a questionnaire which does evaluate the specific problems patients after total laryngectomy face in intelligibility, voice quality and daily communication.

With the shift towards patient-centred care, it is essential to incorporate the patients' view on the communicative outcome. Using PROMs has advantages; it is not time-consuming and can be used multiple times over time to detect changes during the rehabilitation process. It has to be noted that in our prospective study, we seemed to detect a response shift. A response shift refers to the shift in internal standards, values, and meaning of QoL. After surgery, some participants express they are happy to be alive and satisfied with their communicative abilities. This results in normal PROM scores, whilst they reported abnormal scores before surgery and the objective measures show a worsening of voice quality. The response shift can be explained by the ability of human beings to adapt to life events. In future research investigating this response shift, specifically for the head and neck cancer group, is valuable to define what speech-related QoL means for individuals before and after treatment. Further developments in generating condition-specific questionnaires and computer-adaptive testing can improve targeted monitoring of QoL outcomes.

Perceptual instruments

In the studies in this thesis we focus on a general score for perceptually rated voice quality and intelligibility. These parameters are derived from the IINFVo scale, where impression, intelligibility, noise, fluency and voicing are evaluated [14]. The IINFVo scale has been developed to evaluate tracheoesophageal speech, which requires different aspects of evaluation. The scale includes specific aspects which are interesting for tracheoesophageal speech, for example the unintended additive noise during speaking. For usability reasons, we chose to simplify the perceptual outcome to overall impression of voice quality and intelligibility, in both chapter 3 and 5. We preferred the advantage of creating one outcome score for voice quality and one outcome score for intelligibility. Within these studies there were groups of up to ten SLP's who rated the perceptual outcomes.

Perceptual outcomes are subjective by nature, and outcomes are influenced by listeners' understanding of the scales, and listeners' expectations, and tiredness in a long listening experiment. Chapter 5 shows that reported perceptual results are scattered between the expert raters (Figure 5.1). Perceptual rating are still considered as the gold standard by many researchers, but with the promising automatic acoustic evaluation tools as well as the upcoming role of PROMs this might be reconsidered for evaluating speech after total laryngectomy.

Recommendations for multi-dimensional voice and speech analysis

Ideally, outcomes in studies (prospectively) evaluating total laryngectomy speech have to be combined with meta-analyses. Chapter 2 showed that the included number of total laryngectomy patients remains small and heterogeneity in offered speech rehabilitation is present. Prospectivally studying this group proved to be complex because of the change in outcomes over time, drop-out of participants, and the need for educated staff to continuously perform the research assessments. In order to generate consistency in evaluating speech outcomes within studies and over different study groups, it is helpful to create consensus in the multidimensional assessment tools that are used. Having a combined final score for perceptual outcome, (automated) acoustic scores and PROMs improves quality and comparability of the studies. The body of the studies incorporated in this thesis is not strong enough to provide definite recommendations towards the instruments which should be used, but with help of a dedicated expert panel consensus on multi-dimensional analysis of tracheoesophageal speech can be reached.

9.2 Functional issues and well-being after total laryngectomy - Section 2

The chapters incorporated in the second part of this thesis focus on functional issues and psychosocial functioning after total laryngectomy. With the removal of the larynx, speaking, breathing, swallowing and smell function is altered. The functional issues resulting from this are diverse and continually impact daily functioning. In the studies incorporated in the second section we use a variety of research methods, including a questionnaire, functional training, and qualitative research, to investigate the problems that total laryngectomized individuals face.

9.2.1 Summary of the findings

The influence of functional issues on the reported quality of life

In **Chapter 6** we explored functional and participation issues and their interaction with reported QoL after total laryngectomy. With a worldwide questionnaire we assessed interactions between reported issues. Although the response rate was 21%, the total of 1705 respondents is uniquely high in this specific patient group. Respondents were situated in different countries, over several continents. As no validated questionnaires were used, we chose to restructure the questions in issue themes. An advanced analysis was performed to correlate the issue themes to each other and the reported QoL. Interestingly, the issues from the two themes “avoiding social activities” and “experienced limitations in daily activities” are the main influencers of the variance in reported QoL. The ability to fulfil meaningful activities seems to have a greater impact on reported QoL than the purely physical consequences of total laryngectomy in general. Another interesting finding was the strong correlation which was found between “pulmonary issues” and almost all other themes. Pulmonary issues might be partially responsible for other reported issues.

Offering expiratory muscle strength training

In further work on these pulmonary issues, a specific training is provided in the study presented in **Chapter 7**. In this chapter we have investigated the feasibility and safety of expiratory muscle strength training in total laryngectomy patients. It was very promising to see that after the nescessary adaptions, this training device could be used with patients who had undergone total laryngectomy. This study was conducted as a pilot study in testing the feasibility, safety, and compliance of this training program. As the expiratory Muscle Strength Training device was developed for use by mouth, we had to make adaptions to it to make it suitable for use on a tracheostoma.

The outcomes show high compliance to the training protocol. Performing EMST is feasible, although there are some challenging issues and one safety issue occurred. Encountered challenges included the need for an adapter, skills of the participants with plugging the voice prosthesis, placing the EMST150 with an adaptor on the tracheostoma, and creating an airtight seal. One safety issue occurred when a participant appeared without a voice prosthesis after a training session. Nevertheless, compliance to the training program was high (>95%). The results show a clear increase in Maximum Expiratory Pressure (MEP) over time in contrast to Peak Expiratory Flow (PEF) values in which no change was seen. The high baseline MEP values for the total laryngectomized participants were interesting when compared with reference values for healthy adults. The higher than expected MEP scores might be the effect of frequent coughing and forced expectoration which is present after total laryngectomy. None of the secondary outcome measures reached statistical significance, which is probably influenced by the low number of participants and the fact that we chose to include a sample of relatively fit participants.

It would be valuable to continue research regarding pulmonary issues and investigate possibilities to improve the pulmonary condition. As seen in chapter 6, pulmonary issues correlate strongly with other issues such as participating in social activities. This points towards a correlation between the presence of pulmonary issues and problems with performing meaningful daily activities and social activities. It is important to detect patients who suffer from pulmonary issues and provide optimal medical aids and interventions to enhance their pulmonary functioning.

Exploring rehabilitation needs

In **Chapter 8** we explored the impact of a total laryngectomy on women and investigated specific rehabilitation needs. This study further builds on previous qualitative work, which was predominantly carried out with male participants. Our study shows that the disease and its treatment formed a turning point in the lives of the respondents. After the surgery, they made efforts to reintegrate into the social communities they belonged to before their surgery and return to former everyday activities. All respondents found various ways to participate in daily life again. Nevertheless, participants remain persistently vulnerable after total laryngectomy, because of the effort required by taking care and managing functional, social and health-related issues.

The interviews showed that women-specific issues are present in returning to work, the performance of informal care-work, the spousal relationship, intimacy, and social interaction. In social interaction the presented issues were mostly stigmatization due to voice and appearance. For some respondents it was very helpful to engage in peer-support activities. Peer-support can play an important role in counselling and rehabilitation since it offers an opportunity to relate to the experiences of others and help form realistic expectations regarding outcomes.

9.3 Future research perspectives

The prospective study still leaves unanswered questions regarding the prediction of individual speech outcomes. Despite all the advances in speech rehabilitation, we are still not able to predict individual voice outcomes. Previous research has shown that individual factors, such as social status and coping strategy, in addition to treatment factors, such as type of surgery and post-operative radiotherapy, are related to vocal functioning. However, in the prospective study (chapter 3) we did not find an effect of these oncological treatment variables, which might be influenced by sample size and variability of our cohort. If, however, correlations between certain factors and voice outcomes exist, those factors might be used to develop a predictive model. With individualized and patient-centered medicine in mind, there is a demand for methods that can present probable speech outcomes to a patient during the planning phase of cancer treatment. This would allow appropriate predictions of voice and speech quality as well as an audio rendering of the probable speech sound after total laryngectomy. Subsequently, insight in factors which improve voice outcomes would contribute to clinicians' decision making. Future work can further build on this systematic evaluation of voice outcomes, to create a larger dataset and work towards personalized predictions.

There is a need for validated instruments to evaluate functional issues including voice outcomes and participation issues after total laryngectomy. Validation of the AVQI for tracheoesophageal speech specifically contributes to objective outcomes with acoustic analysis. Further development of perceptual scales to assess voice quality and intelligibility are needed, as well as the development of patient-reported outcome measures and/or validating existing instruments which assess functional and participation issues following total laryngectomy.

Pulmonary rehabilitation after total laryngectomy remains an important concern. Although efforts have been made to reduce pulmonary complaints over the past decades, they are still present. The survey study (chapter 6) showed the high correlation of pulmonary issues with other experienced problems. It seems that presence of pulmonary issues influences participation in (social) activities. Providing EMST, proved to be safe and feasible in total laryngectomy patients. It is recommended to further investigate the effectiveness of this training program in a large group of participants who explicitly report pulmonary problems.

9.4 Clinical recommendations

9.4.1 Pre-treatment counselling and tailored rehabilitation

We recommend to provide tailored rehabilitation programs for this group of patients from the initial preoperative period to the postoperative period and through the rehabilitative period. Pre-operative counselling is needed to inform patients who are admitted to total laryngectomy about the procedure and the consequences of the procedure. The process by which the patient is admitted and is informed by healthcare professionals is leading towards more shared decision-making, which refers to the process in which patients and healthcare professionals make health-related choices together. Therefore, it is important to explore patients' priorities, values and concerns about the treatment and the outcome, as well as possible benefits, risks and implications of the treatment options. The SLP should offer pre-operative counselling on the available speech rehabilitation methods and (communicative) functioning after total laryngectomy. Creating realistic expectations about the rehabilitation trajectory and functional outcomes is important to help patients cope with their condition after surgery.

Within this rehabilitation program the focus should be on patient-centered care, in which the provided support is adapted to the needs of the total laryngectomy patient and their loved ones [15–17]. It is important to involve the partner of the laryngectomee, since a shift in give-and-take balance of the spousal relationship can be expected. Clinicians should create openness in discussing the impact of the surgery on the relationship. Ultimately, a high level of patient activation is stimulated. Patient activation is defined as patients' knowledge, confidence, and ability to manage the disease and manage their life again [15, 18, 19].

The SLP should be involved in the rehabilitation trajectory from preoperative counselling up to long-term functioning, since most patients need ongoing management of their communication issues. During the data collection for the prospective study, the SLP's stated that the structured study assessments provided them an opportunity to re-evaluate the rehabilitation process. During this consultation they preferred to discuss overall well-being as well as functional problems in communication, swallowing and smell function. When patients did report problems in these domains the SLP would schedule follow-up appointments. Therefore, structured follow-up assessments during the first year post-laryngectomy are recommended to facilitate these evaluations.

9.4.2 Offering speech rehabilitation

This thesis shows that tracheoesophageal speech offers the most natural speech rehabilitation method after total laryngectomy and, when successful, provides a stable long-term verbal communication method. When possible, patients deserve tracheoesophageal speech rehabilitation in the early stage post-surgery to allow them to regain verbal speech as to enhance QoL as quickly as possible.

Worldwide there are still differences in offered speech rehabilitation methods. In some countries caretakers are reluctant to offer tracheoesophageal speech, since they associate this with a higher complication rate. Not all healthcare systems cover the costs of tracheoesophageal speech, leading to reimbursement issues. This is problematic in a patient group in which individuals are often from a lower socio-economic background. In some cases, patients prefer a speech rehabilitation method which makes them less dependent on the hospital. In tracheoesophageal speech, regular hospital visits are needed for voice prosthesis changes.

With the advances in oncological treatment, these patients deserve the best available rehabilitation options for speech and general functioning to optimize their daily life after surgery. Healthcare professionals should offer guidance and interventions to help total laryngectomized individuals cope with their condition. To assure that the best rehabilitation is offered, clinicians and researchers have to make ongoing efforts to map treatment outcomes, work out reimbursement issues, and implement new developments in the clinical trajectory.

References

[1] P. H. Dejonckere, M. B. J. Moerman, J. P. Martens, J. Schoentgen, and C. Manfredi, "Voicing quantification is more relevant than period perturbation in substitution voices: an advanced acoustical study," *European Archives of Oto-Rhino-Laryngology*, vol. 269, pp. 1205–1212, Apr. 2012.

[2] C. J. van As-Brooks, F. J. Koopmans-van Beinum, L. C. Pols, and F. J. Hilgers, "Acoustic Signal Typing for Evaluation of Voice Quality in Tracheoesophageal Speech," *Journal of Voice*, vol. 20, pp. 355–368, Sept. 2006.

[3] M. S. De Bodt, F. L. Wuyts, P. H. Van de Heyning, and C. Croux, "Test-retest study of the grbas scale: influence of experience and professional background on perceptual rating of voice quality," *Journal of Voice*, vol. 11, no. 1, pp. 74–80, 1997.

[4] I. R. Titze, *Vocal fold physiology: Frontiers in basic science*, vol. 7. Singular Publishing Group, 1993.

[5] Y. Maryn, P. Corthals, P. Van Cauwenberge, N. Roy, and M. De Bodt, "Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels," *Journal of Voice*, vol. 24, no. 5, pp. 540–555, 2010.

[6] M. R. Arias, J. L. Ramón, M. Campos, and J. J. Cervantes, "Acoustic analysis of the voice in phonatory fistuloplasty after total laryngectomy," *Otolaryngology—Head and Neck Surgery*, vol. 122, no. 5, pp. 743–747, 2000.

[7] M. H. Bellandese, J. W. Lerman, and H. R. Gilbert, "An Acoustic Analysis of Excellent Female Esophageal, Tracheoesophageal, and Laryngeal Speakers," *Journal of Speech, Language, and Hearing Research*, vol. 44, pp. 1315–1320, Dec. 2001.

[8] G. W. Blood, "Fundamental frequency and intensity measurements in laryngeal and alaryngeal speakers," *Journal of Communication Disorders*, vol. 17, pp. 319–324, Oct. 1984.

[9] M. L. Ng, C.-L. I. Kwok, and S.-F. W. Chow, "Speech performance of adult cantonese-speaking laryngectomees using different types of alaryngeal phonation," *Journal of Voice*, vol. 11, pp. 338–344, Sept. 1997.

[10] L. Širić, D. Šoš, M. Rosso, and S. Stevanović, "Objective assessment of tracheoesophageal and esophageal speech using acoustic analysis of voice," *Collegium antropologicum*, vol. 36, no. 2, pp. 111–114, 2012.

[11] R. J. Baken and R. F. Orlikoff, *Clinical measurement of speech and voice*. San Diego: Singular Thomson Learning, 2nd ed., 2000.

[12] M. Moerman, J.-P. Martens, and P. Dejonckere, "Multidimensional assessment of strongly irregular voices such as in substitution voicing and spasmodic dysphonia: A compilation of own research," *Logopedics Phoniatrics Vocology*, vol. 40, pp. 24–29, Jan. 2015.

[13] P. H. Dejonckere, P. Bradley, P. Clemente, G. Cornut, L. Crevier-Buchman, G. Friedrich, P. Van De Heyning, M. Remacle, and V. Woisard, "A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques," *European Archives of Oto-rhino-laryngology*, vol. 258, no. 2, pp. 77–82, 2001.

[14] M. Moerman, J.-P. Martens, L. Crevier-Buchman, E. Haan, S. Grand, C. Tessier, V. Woisard, and P. Dejonckere, "The INFVo perceptual rating scale for substitution voicing: development and reliability," *European Archives of Oto-Rhino-Laryngology*, vol. 263, pp. 435–439, May 2006.

[15] J. Bickford, J. Coveney, J. Baker, and D. Hersh, “Living with the altered self: a qualitative study of life after total laryngectomy,” *International Journal of Speech-Language Pathology*, vol. 15, no. 3, pp. 324–333, 2013.

[16] J. M. Bickford, J. Coveney, J. Baker, and D. Hersh, “Support following total laryngectomy: Exploring the concept from different perspectives,” *European Journal in Cancer Care (Engl)*, vol. 27, no. 3, p. e12848, 2018.

[17] M. P. Offerman, J. F. Pruyn, M. F. de Boer, J. J. Busschbach, and R. J. Baatenburg de Jong, “Psychosocial consequences for partners of patients after total laryngectomy and for the relationship between patients and partners,” *Oral Oncology*, vol. 51, no. 4, pp. 389–98, 2015.

[18] F. Jansen, S. E. J. Eerenstein, B. I. Lissenberg-Witte, C. F. v. Uden-Kraan, C. R. Leemans, and I. M. V. Leeuw, “Unmet supportive care needs in patients treated with total laryngectomy and its associated factors,” *Head & Neck*, vol. 40.12, pp. 2633–2641, 2018.

[19] B. J. Noonan and J. Hegarty, “The impact of total laryngectomy: the patient’s perspective,” in *Oncology nursing forum*, vol. 37, pp. 293–301, 2010.

Summary

9.1 English summary

Total laryngectomy Exploring voice outcomes and functional issues

In **Chapter** 1 a general introduction to head and neck cancer is provided. Total laryngectomy as a treatment modality is introduced, as well as the consequences of this procedure regarding voice, speech and quality of life. A total laryngectomy leads to physiological changes and the functional problems. To rehabilitate speech, three main substitute voice methods are available: tracheoesophageal speech, esophageal speech, and electrolarynx speech. These substitute voice methods are described and the functioning of these methods is illustrated. To evaluate voice and speech, a multi-dimensional evaluation is recommended; which entails acoustic evaluation, perceptual evaluation, and patient-reported evaluation. The need for dedicated studies to investigate voice outcomes and functional issues after total laryngectomy is stated.

In **Chapter 2** a systematic literature review evaluating comparative acoustic, perceptual, and patient-reported outcomes for esophageal speech, tracheoesophageal speech, and electrolarynx speech in comparison with healthy speakers is presented. We have performed a systematic search in several databases, followed by a selection process, twenty-six articles are included. In most studies, methodological quality was low. It is likely that in many studies an inclusion bias is present, studies report including only exceptional speakers. When comparing the different groups of speakers, statistically significant better outcomes are reported for tracheoesophageal speakers compared to esophageal speakers for the acoustic parameters, fundamental frequency, maximum phonation time, and intensity. In studies presenting auditory-perceptual outcomes, tracheoesophageal speech is rated with a better voice quality and intelligibility compared to esophageal speech and electrolarynx speech. For self-reported voice outcome, none of the three substitute voice groups reported evidently better outcomes; all speaker groups report a degree of voice handicap. We conclude that tracheoesophageal speech is the favorable speech rehabilitation method according to acoustic and perceptual outcomes. The study emphasizes the need for standardized measurement tools for systematic evaluations of substitute voice speakers.

In **Chapter 3** entails a study in which we have prospectively assessed the course of voice outcomes in total laryngectomy patients from before surgery up to one year post-surgery. Patients who were admitted to total laryngectomy were approached to participate. Data is collected over five hospitals situated in The Netherlands and Australia. The study contained four time-points for each participant: prior to total laryngectomy, three months, six months, and twelve months post-surgery. Our study assessments included perceptual evaluation, voice recordings, and self-reported outcomes. The Voice Handicap Index 10 item version (VHI-10) is used for self-reported outcomes. The Acoustic Voice Quality Index (AVQI) is used for acoustic evaluation, AVQI score combines several acoustic measures in one score. Forty-three participants were included in the study, of which 93% received a voice prosthesis. Nevertheless, satisfactory voice rehabilitation with tracheoesophageal speech was not accomplished in all cases, mostly as a result of medical complications. Thirty percent of the participants deceased within the first year after surgery. For the total group, after surgery, a worsening of all voice related values is seen, which gradually improves over time up to one year post surgery. When outcomes for the multidimensional assessment methods assessed over time are pooled, strong correlations are found between the dimensions of voice related outcomes. Worst quality of life score was reported before surgery, after surgery this score gradually improved towards the reference value of healthy people in the same age group. In self-reported voice handicap and quality of life, a response shift was seen, indicating a change in internal standards, values, and meaning of quality of life.

We present a detailed acoustic study investigating the changes of the word initial consonants /t/ and /d/ in Dutch speaking individuals before and after total laryngectomy in **Chapter 4**. Speech recordings of seventeen participants before and after total laryngectomy are used. Participants did read aloud a Dutch text of neutral content, *Tachtig dappere fietsers* ('eighty brave cyclists'). This text contained eighteen tokens with /t/ or /d/ in initial position. For each token prevoicing, burst duration and duration of the vowel following the word initial consonant was analyzed. The results show that the acoustic contrast of the /t/ and /d/ in word initial position is reduced after total laryngectomy. In tracheoesophageal speech, the presence of prevoicing of /d/ decreases, and burst duration increases. In the post-operative situation, therefore, /d/ becomes more similar to /t/ acoustically. For clinical practice, special attention towards intelligibility issues is recommended to enhance understandability of tracheaesophageal speech.

Chapter 5 presents the outcomes of longitudinal collected voice recordings. We assessed differences in voice quality and intelligibility of thirtheen tracheoesophageal speakers using speech samples that were recorded with at least seven years in between. Intelligibility and voice quality are perceptually evaluated by ten experienced speech and language pathologists. In addition, automatic speech evaluations are performed with tools from Ghent University as well as AVQI. No statistical significant group effect was found for changes in voice quality and intelligibility. The results show a wide inter-speaker variability. Perceptual ratings show a high variation for the perceptual ratings by the speech language pathologists. Tools for automatic evaluation of voice quality and intelligibility correlated strongly, these tools are promising for analyzing trends within individual speakers over time. We can conclude that intelligibility and voice quality of tracheoesophageal speech is mostly stable over a period of seven to 18 years.

In **Chapter 6** we present the results of a worldwide survey assessing functional and participation issues after total laryngectomy. A questionnaire was sent out to 8119 recipients of whom 1705 (21%) responded. The questionnaire included 26 questions, including demographic information and regarding medical product use of the respondents, their experienced overall health and independence, and functional and participation issues. For the analysis, respondents are grouped based on sex, age, time since total laryngectomy, educational level, and country of residence. Questions are grouped in a measure of reported-Quality of Life (r-QoL) and seven issue themes (esthetic issues, experienced limitations in daily activities, avoiding social activities, communication issues, experienced vulnerability due to environmental factors, pulmonary issues and sleep issues) which allowed us to assess the underlying relations. The outcomes of the study show that a higher amount of functional and participation issues and a lower r-QoL was reported in the group of younger respondents (<60yrs), women, and respondents who have had the total laryngectomy procedure less than two years ago. The ability to participate in meaningful and social activities was a major factor determining r-QoL. A strong correlation was found between pulmonary issues and most other themes. Pulmonary issues might be an underlying cause of many other issues and therefore negatively influence r-QoL.

We have investigated the effect of an Expiratory Muscle Strength Training (EMST) program in **Chapter 7**. EMST is a threshold based device-driven treatment for improving expiratory pressure and cough function in order to diminish pulmonary complaints. Ten male participants who had undergone total laryngectomy were included. For this pilot study we aimed to examine feasibility, safety, and compliance of the EMST training program. Objective and subjective outcome measures included manometry, spirometry, cardio pulmonary exercise testing, voice recordings, and patient reported outcome measures. Nine participants completed the full study protocol. Compliance to the training program was high. All were able to perform the training, although it requires adjustments of the device and skills of the participants. The results show that maximum expiratory pressure (MEP) and vocal functioning in loudness improve over time. After EMST, no changes were seen in other objective and subjective outcomes. EMST appeared to be feasible and safe in use after total laryngectomy. MEP improved over time during the training period but no improvement in the clinically relevant outcome measures are seen in our sample of relatively fit participants. We recommend further investigation of the training in a larger group of (less fit) participants who report specifically pulmonary complaints to investigate if this increase in MEP can result in a clinical benefits such as less frequent coughing.

In **Chapter 8** we present the results of a qualitative study in which we interviewed female patients on their perspective of life after total laryngectomy. The study aimed to gain insight into the impact of total laryngectomy on women's daily life while identifying their specific rehabilitation needs. In-depth, semi-structured interviews with eight women who had undergone total laryngectomy were conducted. Using an interview guide, participants were encouraged to discuss their everyday experiences, while also focusing on issues typical to women. The transcribed interview data is analysed by the method of Thematic Analysis, taking interpretative phenomenological analysis as a lead. The interviews reveal three main themes: disease and treatment as a turning point, re-establishing meaningful everyday activities, and persistent vulnerability. Study participants report that they experience challenges in their rehabilitation process due to physical disabilities, dependency on others, and experienced stigma. Women who undergo total laryngectomy are likely to experience issues in returning to work, the performance of informal care work, the spousal relationship, intimacy, and social interaction due to stigmatization. Medical pre-treatment counselling and multidisciplinary rehabilitation programs should help patients form realistic expectations and prepare them for the changes they will face. A gender- and age-matched total laryngectomized patient visitor can contribute to this process. Rehabilitation programs should incorporate the partner and offer psycho-social support for women following total laryngectomy to return to their former roles in family life, social life and work-related activities.

In **Chapter 9** the studies which are incorporated in this thesis are discussed and recommendations for clinical practice and future research are provided. To work towards uniformity in evaluating speech outcomes after total laryngectomy, consensus is needed regarding the assessment tools used. Having one final score for each multidimensional outcome including perceptual outcome, (automated) acoustic scores, and patient-reported outcomes helps to investigate change over time within patients and improves comparability between studies. Our studies show that automated assessment are promising. The AVQI is a potential instrument to classify voice outcomes and also correlates highly with perceptual voice quality and intelligibility outcomes. In future research, a specific cut-off value for AVQI and VHI-10 outcomes in total laryngectomy patients can be determined with help of a validation study. Future work can build further on this systematic evaluation of voice outcomes, to create a larger data-set and work towards personalized predictions. For clinical practice we recommend to provide tailored rehabilitation programs. Pre-treatment counselling should incorporate information on expected outcomes to facilitate shared decision-making. Patients need to be informed about the different substitute speech methods. We have shown that tracheoesophageal speech offers the most natural speech rehabilitation method; when possible patients should be offered tracheoesophageal speech as a verbal rehabilitation method. Throughout the full trajectory of counselling and rehabilitation the impact of the surgery on daily life has to be discussed by clinicians. This facilitates the possibility to determine personalized rehabilitation goals on functional and participation level.

9.2 Dutch summary

Totale laryngectomie Stemuitkomsten en functionele problemen

Hoofdstuk 1 van dit proefschrift introduceert de onderwerpen van dit proefschrift. We geven een inleiding over hoofhalskanker en de behandeling hiervan, met in het specifiek de behandeloptie totale laryngectomie. Een totale laryngectomie is een operatie waarbij het strottenhoofd wordt verwijderd. Hiermee verliest de patiënt zijn natuurlijke stem en verandert de ademweg met het stoma dat wordt aangelegd in de hals. We leggen uit wat voor effecten de operatie heeft op de anatomie en fysiologie en uitkomsten in stem, spraak en kwaliteit van leven. Om het spreken te revalideren zijn de drie meest gebruikte methodes: tracheo-oesofageale spraak, slokdarmspraak en electrolarynxspraak. We introduceren deze 3 substituut- spraakmethoden met illustraties van de werking. Voor het evalueren van deze stemmen wordt gebruik gemaakt van akoestische analyse, perceptuele analyse en patiënt-gerapporteerde uitkomsten. In meerdere hoofdstukken wordt gebruik gemaakt van deze manieren van evaluatie. Dit proefschrift bestaat uit twee delen. De hoofdstukken in het eerste deel van dit proefschrift geven inzicht in uitkomsten in stemkwaliteit en verstaanbaarheid na een totale laryngectomie. Het tweede deel bevat hoofdstukken over algeheel functioneren, specifieke revalidatiebehoeften van vrouwelijke gelaryngectomeerden en naar de behandeling van hoestklachten na een totale laryngectomie.

In **hoofdstuk 2** presenteren we een systematisch opgezet literatuuronderzoek waarin we de verschillende stemrevalidatiemethoden met elkaar vergelijken. We vergelijken de drie substituut-spraakmethoden met spraak van gezonde sprekers, waarbij akoestische, perceptuele en patiëntgerapporteerde uitkomsten meegenomen worden. Voor deze literatuurstudie zijn verschillende databases doorzocht en is een systematisch opgezet selectieproces doorlopen. Er konden 26 artikelen worden geïncludeerd die voldeden aan de selectiecriteria. De meeste van deze artikelen hadden een lage methodologische kwaliteit waardoor het aannemelijk is dat in veel studies een zogenaamde bias aanwezig is. Een aantal beschrijft bijvoorbeeld dat alleen hele goede sprekers gevraagd zijn om deel te nemen aan hun studie wat zorgt voor een vertekening van de resultaten. Echter, rekening houdend met deze bias, laten de resultaten van deze vergelijkende literatuurstudie zien dat tracheo-oesofageale sprekers een betere uitkomst laten zien ten opzichte van slokdarmsprekers op perceptuele en akoestische uitkomsten. Dit is het geval voor de akoestische parameters fundamentele frequentie, maximale fonatieduur en intensiteit. Voor de patiëntgerapporteerde uitkomsten zien we dat geen van de revalidatiemethoden een duidelijk betere uitkomst laat zien. De literatuurstudie laat zien dat tracheo-oesofageale spraak de voorkeur heeft als revalidatiemethode voor akoestische en perceptuele uitkomsten maar dat voor patiëntgerapporteerde uitkomsten geen duidelijke voorkeursmethode naar voren komt.

In **hoofdstuk 3** hebben we een prospectieve studie opgezet met gestandaardiseerde meetmethoden voor akoestische, perceptuele en patiëntgerapporteerde uitkomsten. Het doel van deze studie was om het verloop van de spraakrevalidatie na een totale laryngectomie van voor de operatie tot een jaar na de operatie vast te leggen. In vijf ziekenhuizen in Nederland en Australië werden patiënten die een totale laryngectomie als behandeloptie voorgesteld kregen gevraagd om mee te doen aan deze studie. Voorafgaand aan de laryngectomie en drie, zes en twaalf maanden na de operatie heeft de logopedist een assessment uitgevoerd. Ieder assessment bestond uit perceptuele evaluatie van de stem, stemopnames en vragenlijsten. De Voice Handicap Index 10 (VHI-10) is gebruikt om de patiënt-gerapporteerde stemhandicap vast te leggen. De stemopname is gebruikt om de Acoustic Voice Quality Index (AVQI) score te berekenen. De uiteindelijke analyse bevat 43 deelnemers. Bij 40 van de deelnemende patiënten (93%) is een stemprothese geplaatst om te revalideren met tracheo-oesofageale spraak. Over de gehele groep zien we drie maanden na de operatie een verslechtering van alle stemuitkomsten. Hierna wordt een geleidelijke verbetering gezien. Na één jaar zijn er nog 23 deelnemers in de studie, waarvan 20 succesvol kunnen spreken met tracheo-oesofageale spraak. Degenen waarbij de spraak niet succesvol is gerevalideerd hadden medische complicaties. Van de groep deelnemers overleed 30% in het eerste jaar na de operatie. Algehele kwaliteit van leven wordt voor de groep deelnemers het slechtst beoordeeld voorafgaand aan de operatie. Over tijd is er een verbetering zichtbaar. Binnen de zelf-gerapporteerde uitkomsten zien we een response shift optreden. Response shift verwijst naar verandering in de betekenis van kwaliteit van leven en communicatief functioneren, veroorzaakt door een verandering van de interne standaarden. Het gevolg is een discrepantie in de patiëntgerapporteerde stemhandicap en de akoestisch gemeten stemkwaliteit. Sprekers waarderen de tracheo-oesofageale spraak ondanks de afwijkende klank. Dit is van belang om mee te nemen in de voorlichting. Wat verder uit de resultaten bleek is, indien er sprake is van een slechte oncologische prognose dit leidt tot een grotere kans om niet te kunnen spreken in de laatste levensfase. Ook dit is relevant om in de voorlichting mee te nemen. Onze studie laat zien dat het meetinstrument, de AVQI, een potentieel sterk instrument is om een objectieve classificatie van de stemuitkomst te geven. Voor de toekomst is het zinvol om een validatiestudie op te zetten voor zowel de AVQI als de VHI-10, met als doel een afkapwaarde bepalen voor stemprothesespraak. Hiermee kunnen vergelijkingen over tijd en tussen groepen patiënten gemaakt worden om de uitkomsten beter in kaart te brengen.

In **hoofdstuk 4** onderzoeken we in detail de akoestische verschillen in de woord initiële medeklinkers /t/ en /d/ in Nederlandse sprekers voorafgaand en na een totale laryngectomie. We gebruiken hiervoor opgenomen spraakopnames van zeventien studiedeelnemers die voor en na totale laryngectomie de tekst *Tachtig dappere fietsers* hebben voorgelezen. Deze tekst bevat achttien tokens (itemwoorden) waarbij de /t/ en /d/ in initiële positie van het woord staan. Voor ieder token hebben we de akoestische metingen uitgevoerd op prevoicing, burst duration en duur van de klinker volgend op deze initiële consonant. De uitkomsten van deze analyses laten zien dat het akoestisch contrast tussen de woord initiële consonanten /t/ en /d/ minder wordt na totale laryngectomie. In tracheo-oesofageale spraak is er een vermindering van de aanwezigheid van prevoicing en een langere burst duration voor de /d/. Hierdoor worden de akoestische eigenschappen van de /d/ meer vergelijkbaar met een /t/. Dit leidt tot meer misverstanden in spraakverstaan en hiermee een slechtere verstaanbaarheid. Voor de klinische dagelijkse praktijk is het van belang om aandacht te besteden aan deze en breder bestaande verstaanbaarheidsproblemen.

In **hoofdstuk 5** presenteren we lange termijn uitkomsten van tracheo-oesofageale sprekers. Voor dit onderzoek hebben we spraakopnames van dertien sprekers gebruikt waarbij minimaal zeven jaar tussen de eerste en tweede opname zat. Het verschil in stemkwaliteit en verstaanbaarheid over tijd is gemeten met akoestische en perceptuele analyse. Tien logopedisten hebben door middel van een luisterexperiment de spraakopnames perceptueel beoordeeld. Daarnaast zijn de samples met tools voor automatische akoestische spraakanalyse, ELIS, ELISALF en de AVQI, geanalyseerd. Voor de gehele groep, wordt over tijd geen statistisch significant verschil gevonden voor de perceptueel en automatisch beoordeelde stemkwaliteit en verstaanbaarheid. Wel bleek dat er een grote variabiliteit zat tussen de sprekers. De perceptuele beoordelingen uitgevoerd door de logopedisten lieten ook een grote variatie zien, desondanks was er een sterke correlatie tussen de perceptueel beoordeelde stemkwaliteit en perceptueel beoordeelde verstaanbaarheid. Daarnaast bleken de AVQI uitkomsten zijn sterk gecorreleerd met de perceptuele beoordelingen. Dit maakt dat automatische tools, zoals de AVQI, veelbelovend zijn. Met deze tools kunnen veranderingen in stem en spraak van individuele sprekers over tijd geanalyseerd worden. We concluderen dat de stemkwaliteit en verstaanbaarheid voor tracheo-oesofageale sprekers stabiel genoemd kan worden over een periode van zeven tot achttien jaar na de operatie.

Het tweede deel van dit proefschrift bevat hoofdstukken die gericht zijn op functionele uitkomsten, longproblemen en vrouwspecifieke problemen die aanwezig kunnen zijn na totale laryngectomie.

In **hoofdstuk 6** laten we de resultaten zien van een wereldwijd uitgevoerde studie naar functionele en participatieproblemen na een totale laryngectomie. Hiervoor werd een vragenlijst verstuurd naar 8119 gelaryngectomeerden, van wie er 1705 (21%) reageerden. De vragenlijst bevat 26 vragen; er is demografische informatie uitgevraagd, en gevraagd naar het gebruik van medische hulpmiddelen het ervaren van de algehele gezondheid en vragen ten aanzien van functionele en participatiekwesties. Voor de analyse zijn de respondenten gegroepeerd op basis van geslacht, leeftijd, tijd sinds totale laryngectomie, opleidingsniveau en het land waar men woont. De uitkomstmaat 'mate van gerapporteerde kwaliteit van leven' (r-QoL) en zeven geselecteerde probleemthema's zijn gebruikt voor de analyse. De probleemthema's zijn: esthetische problemen, ervaren beperkingen in dagelijkse activiteiten, vermijden van sociale activiteiten, communicatieproblemen, ervaren kwetsbaarheid door omgevingsfactoren, longproblemen en slaapproblemen. Met de analyse konden we onderliggende relaties tussen de r-QoL en de probleemthema's onderzoeken. Uit dit onderzoek blijkt dat meer functionele en participatieproblemen als ook een lagere r-QoL worden gerapporteerd in de groep jongere respondenten (<60 jaar), vrouwen en respondenten die minder dan twee jaar geleden de totale laryngectomie hebben ondergaan. Het vermogen om deel te kunnen nemen aan zinvolle en sociale activiteiten is een belangrijke factor in de uitkomst van r-QoL. De meeste deelnemers rapporteren longproblemen. Deze pulmonale klachten hebben een sterke correlatie met de meeste andere thema's. Vanwege deze correlatie lijkt het aannemelijk dat pulmonale klachten een onderliggende oorzaak kunnen zijn van veel andere problemen en daarmee de gerapporteerde kwaliteit van leven negatief beïnvloeden.

In **hoofdstuk 7** onderzoeken we het effect van een hoestkrachttraining, de Expiratory Muscle Strength Training (EMST). Dit trainingsprogramma is uitgevoerd door een groep patiënten die een totale laryngectomie hebben ondergaan. De EMST-trainer is een klein apparaat met een spanningsveer waarin de patiënt via de mond moet uitblazen tegen weerstand. De training heeft als doel het verbeteren van de uitademingsdruk en hiermee de hoestfunctie, wat uiteindelijk kan leiden tot minder pulmonale klachten. In onze studie hebben we tien deelnemers met laryngectomie geïncludeerd. Er is een koppelstuk gemaakt om de EMST te gebruiken op het stoma in de hals. Voor deze pilotstudie wilden we de haalbaarheid, veiligheid en therapietrouw van het EMST-trainingsprogramma onderzoeken. Om de uitkomst van de training te meten hebben we een combinatie van objectieve en subjectieve uitkomstmaten gebruikt: manometrie, spirometrie, een cardiopulmonale inspanningstest, spraakopnames en patiëntgerapporteerde uitkomstmaten. Negen deelnemers voltooiden het volledige studieprotocol. Het trainingsprogramma bleek een hoge therapietrouw te hebben. Ondanks dat het handigheid van de deelnemers vereiste om de EMST op het stoma te plaatsen, waren de deelnemers in staat de training uit te voeren. De resultaten laten zien dat maximale druk tijdens expiratie (MEP) en luidheid van de stem na het voltooien van het trainingsprogramma duidelijk verbeterden. Echter, er werden na het trainingsprogramma geen significante veranderingen gezien in andere objectieve en subjectieve uitkomsten. Concluderend kunnen we zeggen dat de EMST haalbaar en veilig is om te gebruiken na een totale laryngectomie. MEP uitkomsten verbeterden gedurende de trainingsperiode, maar er werd geen verbetering gezien in de klinisch relevante uitkomstmaten, zoals minder vaak hoesten, in onze groep van relatief fitte deelnemers. Het wordt aanbevolen deze training verder te onderzoeken en uit te voeren bij een grotere groep (minder fitte) deelnemers die specifiek pulmonale klachten aangeven. Hiermee kan onderzocht worden of ook hier een toename in MEP wordt gezien en of dit kan leiden tot klinische voordelen zoals minder vaak hoesten.

In **hoofdstuk 8** presenteren we een onderzoek waarin we vrouwelijke patiënten interviewden over hun perspectief op leven na een totale laryngectomie. Het doel van dit kwalitatief onderzoek was om inzicht te krijgen in de impact van totale laryngectomie op het dagelijkse leven van vrouwen en om specifieke revalidatiebehoeften te identificeren. We hebben semi-gestructureerde diepte interviews gehouden met acht vrouwen die een totale laryngectomie hadden ondergaan. Met behulp van een interviewlijst werd de deelnemers gevraagd naar hun ervaringen in de aanloop naar de operatie toe, de revalidatie-periode en het dagelijks leven nu. Daarnaast hebben we gevraagd naar vrouw-specifieke ervaringen. De getranscribeerde interviewgegevens hebben we geanalyseerd met behulp van thematische analyse, waarbij interpretatieve fenomenologische analyse als leidraad is genomen. Er kwamen drie hoofdthema's naar voren vanuit de interviews: ziekte en behandeling als keerpunt, terugkeer naar zinvolle dagelijkse activiteiten en blijvende kwetsbaarheid. Studiedeelnemers vertelden dat ze het moeilijk hadden tijdens hun revalidatieproces door fysieke problemen, de afhankelijkheid van anderen en ervaren stigma. Het is waarschijnlijk dat vrouwen die een totale laryngectomie ondergaan door stigmatisering problemen ondervinden bij bijvoorbeeld terugkeer naar werk. Daarnaast gaven de vrouwen aan moeilijkheden te ervaren in het uitvoeren van zorgtaken, de relatie met hun partner met o.a. intimiteit en sociale interactie. Deze informatie meenemen in de voorlichting voorafgaand aan de behandeling en het aanbieden van multidisciplinaire revalidatie kan patiënten helpen realistische verwachtingen te vormen en hen beter voor te bereiden op de veranderingen die ze zullen gaan doormaken. Men kan hierbij denken aan het inzetten en scholen van patiëntvoorlichters, die een belangrijke bijdrage aan dit voorlichtingsproces leveren. Het is ook van belang de eventuele partner of naasten te betrekken in het revalidatieproces. Tot slot kan in dit proces specifieke aandacht gegeven worden aan het aanbieden van psychosociale ondersteuning aan vrouwen na een totale laryngectomie en hun partners. Dit kan mogelijk de terugkeer naar hun vroegere rol in het gezinsleven, het sociale leven en werk gerelateerde activiteiten bevorderen.

In **hoofdstuk 9** volgt een discussie van de onderzoeken uit dit proefschrift, formuleren we aanbevelingen voor de klinische praktijk en geven suggesties voor vervolgonderzoek. Om meer uniformiteit te bereiken in het evalueren van stem- en spraakuitkomsten na een totale laryngectomie is consensus nodig over de te gebruiken meetmethoden. Het lijkt waardevol om meetmethoden te gebruiken die één score geven voor perceptuele- en akoestische uitkomst en zelf-gerapporteerde uitkomst. Dit zorgt ervoor dat het makkelijker is om veranderingen over tijd te onderzoeken en studies met elkaar te vergelijken. Automatische spraakevaluatietools zoals de AVQI zijn veelbelovend. De AVQI lijkt een waardevol instrument om stemuitkomsten na totale laryngectomie te classificeren aangezien de uitkomsten sterk correleren met akoestische en perceptuele uitkomsten. In vervolgonderzoek zou het nuttig zijn om gespecificeerde afkapwaarden te berekenen voor totale laryngectomie patiënten voor de uitkomsten op AVQI en VHI-10. In toekomstige studies is het ook van belang om het systematisch vastleggen en evalueren van stemuitkomsten over tijd voort te zetten. Hiermee kan een grotere dataset opgesteld worden en kan een voorspelmodel ontwikkeld worden om de stemuitkomsten te voorspellen voorafgaand aan de laryngectomie. Uiteindelijk kan dit bijdragen aan eerlijke en realistische voorlichting. In de voorlichting voorafgaand aan de operatie is het de aanbeveling om in te gaan op de verwachte uitkomst zodat de patiënt een weloverwogen keuze kan maken voor de behandeling.

Concluderend kunnen wij stellen dat tracheo-oesofageale stemgeving de voorkeursmethode is voor spraakrevalidatie. Vanuit dit onderzoek raden wij aan dat, indien mogelijk, deze methode aangeboden wordt aan patiënten die een totale laryngectomie ondergaan. Gedurende het hele traject van voorlichting en revalidatie is het van belang om de impact van de operatie op het dagelijks leven te bespreken. Hiermee kunnen patiënt-specifieke en interdisciplinaire revalidatiedoelen geformuleerd worden om bij te dragen aan een zo goed mogelijke uitkomst op het gebied van dagelijks functioneren en participatie.

About the author

Klaske van Sluis was born July 30th, 1988 in Leeuwarden (Friesland, The Netherlands). After finishing her pre-vocational secondary education, she did a secondary vocational education degree at the ROC Friese Poort in Leeuwarden, where she trained to be a teaching assistant in primary education. During the several internships for this degree she taught children in groups, set up learning goals and provided individual teaching. She developed an interest in providing one-to-one guidance to children as well as an interest in language development. Consequently, she decided to study speech language pathology. At the Hanze University of Applied Sciences Groningen she did the bachelor's degree in speech pathology. During the second and third year of this degree she took courses in linguistics at The University of Groningen. After finishing her degree, she started as a speech language pathologist in a nursery home. Simultaneously, she did a master's degree at the University of Utrecht. For her master's thesis, Klaske studied voice quality in early stage laryngeal cancer patients, comparing data before and after cancer treatment. By the time she had finished her master's degree, she had developed an interest in speech language pathology for patients who suffer from head and neck cancer. Therefore, immediately after graduation, she started working as a speech language pathologist at The Netherlands Cancer Institute - Antoni van Leeuwenhoek. In the next year, she started a PhD program at the University of Amsterdam, Amsterdam Center for Language and Communication and The Netherlands Cancer Institute. During her PhD, Klaske was supervised by Prof. dr. Paul Boersma, Prof. dr. Michiel van den Brekel, Dr. Rob van Son, and Dr. Lisette van der Molen. As part of her PhD studies, Klaske set-up a multicenter study in The Netherlands and Australia, for which she visited hospitals in Sydney area. She completed courses in research skills, provided guest lectures at the university, and supervised students. She published and presented the results of her PhD studies in international scientific journals and at conferences. Furthermore, she is an active member of an association for Dutch linguistics, het Werkverband Amsterdamse Psycholinguisten, for which she writes in the news bulletin.

Klaske currently works as a speech language pathologist and as a clinical project manager at The Netherlands Cancer Institute - Antoni van Leeuwenhoek.

Dankwoord

Promoveren doe je niet alleen. Ik al helemaal niet. De afgelopen jaren is mijn drang tot samenwerken alleen maar sterker geworden. Ik hou van samen sparren, leren van andere mensen, nieuwe werkvelden leren kennen en samen het eindresultaat vieren. Zowel bij het Antoni van Leeuwenhoek als bij het Amsterdam Center for Language and Communication van de UvA heb ik een heel plezierige tijd gehad. In dit dankwoord wil ik een heel aantal mensen bedanken voor de samenwerking en hun steun tijdens dit traject.

Allereerst wil ik mijn team van promotoren en co-promotoren bedanken. Michiel, dank voor alle vrijheid die je me gaf. In je manier van leidinggeven staat vertrouwen voorop. Je steunt me altijd, bij nieuwe ideeën, bij levensveranderingen en je stelt me met trots voor aan andere onderzoekers in het werkveld. Paul, de eerste twee jaren van mijn promotietraject heb ik weinig gebruik gemaakt van je expertise. Met een werkplek in het ziekenhuis en maar een half been in het ACLC zagen we elkaar weinig. Ik was vooral heel trots dat jij als mijn promotor volle zalen trok op het PEVOC congres. Op de terugweg in de Thalys en met een maaltijd van McDonalds leerden we elkaar beter kennen. Hierna werkten we samen aan de ICPhS paper. Bedankt voor je kritische blik en de complimenten over het eindresultaat. Rob, zonder jou was dit proefschrift er zeker niet geweest. Wat heb ik veel van je geleerd de afgelopen jaren. Statistiek, programmeren, akoestische analyse, plaatjes maken, je bent van alle markten thuis. Jouw gedrevenheid voor het onderzoek heeft me heel erg geholpen. Samen studenten begeleiden is een goede combi, want we vullen elkaars eigenschappen goed aan. Lisette, als gepromoveerd logopedist was je meteen een voorbeeld voor me. Je bent een powervrouw die laat zien dat een carrière in de wetenschap te combineren is met het gezinsleven. Met je warme persoonlijkheid en pragmatische blik heb je me altijd enthousiast weten te houden.

Geachte leden van de promotiecommissie: Ellen Gerrits, Gwen van Nuffelen, Julia van Weert, Judith Rispens en Fons Balm, hartelijk dank dat jullie in de commissie zitting willen nemen. Ik kijk er naar uit om met jullie de uitkomsten van mijn proefschrift te bediscussiëren tijdens de promotieplechtigheid.

Heel veel dank gaat uit naar de patiënten die met mijn onderzoeken hebben meegedaan. Het krijgen van de diagnose kanker is vreselijk. Toch heb ik gezien wat een aanpassingsvermogen en levenslust er in mensen naar boven kan komen. Dat jullie als patiënten het onderzoek naar deze ziekte verder willen helpen en lotgenoten steunen vind ik fantastisch. Ik ben trots op alle mannen die hard hebben getraind bij de hoestkrachttraining en geraakt door de vrouwen uit de interviewstudie.

Ik wil alle co-auteurs bedanken voor hun bijdragen aan het tot stand komen van dit proefschrift. I would also like to thank the Australian co-workers for their contribution and the warm welcoming dinner in Sydney. Richard, geweldig dat je het initiatief nam om een multi-center onderzoek op te zetten in Australië. Mijn ervaringen daar waren uniek, erg leuk om jou en je gezin in die setting op te kunnen zoeken. Frans Hilgers, dank voor de bijdrage aan het systematische literatuuronderzoek. Door jouw jarenlange gedrevenheid om de stemprothesespraak van laryngectomiepatiënten vooruit te helpen kunnen we deze mooie onderzoekslijn voortzetten. Maartje, samen schreven we een artikel over veel en complexe data. Door onze wekelijkse meetings kregen we een resultaat om trots op te zijn, dank voor alle life-hacks. Gili, samen met jou kan ik uren op een stationnetje van tevredenheid staan. Jij hebt met je chaos, feminisme en filosofische blik een frisse wind laten waaien op onze afdeling. Martijn en Wim, dank voor jullie bijdrage aan de EMST-studie. Martijn, je liet ons het hoofd vaak duizelen met je wervelwind aan ideeën en analyses over hoestkracht maar het eindresultaat mag er zijn. Wim, we hebben je uit je comfort zone gehaald met het begeleiden van de fietstest bij de gelaryngectomeerden, toch hebben we iets unieks neergezet met deze studie.

Ton Vlasveld en Saar Muller, dank voor jullie bijdrage aan de hoestkrachtstudie.

Alle collega's van het AvL, bedankt voor de prettige samenwerking. Toen ik in 2014 het AVL binnenliep werd ik meteen opgenomen in het hechte logopedieteam met Anne, Merel en Lisette. Merel, je bent altijd betrokken en gedreven voor de logopedie, veel dank voor je hulp bij het verzamelen van de spraakopnames. Door de jaren heen zijn Brigitte, Els, Nadya en Malou onderdeel geweest van dit dreamteam, ook jullie heel erg bedankt voor de leuke logopediedagen die het onderzoekswerk afwisselend maakten. Baris, Bing, Frans, Fons, Pim, Lot, Luc, Michiel, Martin, Richard, Lilly-Ann, Peter, oftwel alle (oud) hoofd-halschirurgen van de afdeling wil ik bedanken. Verder wil ik Ellen, Hannah, Marene, Renske, Robert, Maarten, de collega's van polikliniek 1D en de collega's van 5C heel erg bedanken voor de prettige samenwerking. Marion, dank voor alle vergaderruimtes die je voor me geregeld hebt en de talloze overige acties zoals het checken van de agenda van Michiel. Collega-onderzoekers uit het AVL, gezelligheid in het U, in het O, ergens op een congres of op een reisje. Ann-Jean, Anne Heirman, Alejandra, Danique, Hanneke, Jos, Joris, Maartje, Martijn, Marije, Matthijs, Pieter, Rebecca, Tessa, Thomas en Zeno, veel dank voor de leuke tijd. Ik heb genoten van de ongedwongen gezelligheid. Coralie, Dominique en Marise heel veel succes met jullie PhD en leuk dat ik betrokken mag zijn bij jullie projecten.

Tijdens mijn promotietraject heb ik met veel plezier studenten begeleid. Met hun onderzoekswerk hebben ze mijn projecten verder geholpen. Heel veel dank daarvoor aan Thalita, Eric, Marijn, Merel, Marit en Keely. Eric, thank you for your extensive work, the conversations, and your interest in the Dutch culture. I really enjoyed our reunion in Melbourne.

Ik wil alle collega's van het ACLC bedanken. Met veel plezier kijk ik terug op het contact met jullie, de LOT-winterschool en vooral het jaarlijkse ACLC-uitje. Folkert Kuiken, dank voor de jaarlijkse begeleidingsgesprekken. Je liet ons ieder jaar reflecteren op de voortgang en de begeleiding, zonder oordelen maar met een positieve doch kritische inslag, erg prettig. Dirk-Jan Vet, veel dank voor je hulp bij het vinden van de juiste microfoon en het maken van de koppelstukjes voor de dataverzameling. Het is altijd leuk om bij je binnen te wandelen in je rariteitenkabinet aan technische apparaten. Mijn mede PhD studenten, Imme, Sybren, Darlene, Iris, Ulrika, Sune, Nathalia, Thom, Klaas en Marieke, dank voor de warme ontvangst in het PC Hoofthuis. Manon, mijn AVL-ACLC PhD-maatje, ondanks dat we in een hele andere fase van ons leven een PhD-traject doen zitten we in hetzelfde schuitje, dit schept een band. David, Mirjam, Camilla, Silke en Paul dank voor de onverwachte avonden tafelen met goede gesprekken tijdens de reis naar Melbourne voor het ICPhS congres.

Mijn WAP-redactiegenootjes, Imme, Darlene, Saskia, Linda, Jael en Tessa. Ieder kwartaal maken we weer de vertaalslag van de wetenschap naar de praktijk door het nieuwsbulletin. Dank voor de gezellige redactievergaderingen.

Naast alle collega's wil ik ook graag vrienden bedanken die doordat zij in mijn leven zijn me hebben gebracht tot waar ik nu ben. Esther en Amanda jullie hebben me al van jongs af aan de waarde van vriendschap laten inzien. Mijn studiegenootjes vanuit de logopedie uit Nederland en Duitsland met Catharina, Hannah, Laura en Margreet, dank voor alle liefste meisjes momenten die we hebben. Mijn jaarclub, Anouk, Jeske, Jessie, Laura, Sanne, Liesbeth, Floor, Mienke en Helen. We hebben onze clubnaam omgedoopt naar 'jc pimp je cv', we zijn een groep gedreven meiden en zijn er voor elkaar. Liesbeth, sinds het begin van onze studententijd nemen we samen stappen in het leven en zo moet het zijn, voor altijd. Mieke, mijn logopediewetenschap vriendin, met jou en Johan hebben we prachtige vakanties gehad, ik ben dankbaar voor onze vriendschap. Femke, van huisgenoot tot goede vriendin en nu ontwerper van de cover van dit boek. Door jou heeft iedereen een mooie aanwinst in de boekenkast. Meiden van vrouwenwielrennen Amsterdam, dank voor de fietsende ontspanning.

Dank aan mijn paranimfen. Anne, je bent mijn collega logopedist, co-auteur op de twee beste artikelen ooit en mijn maatje. We voelen elkaar goed aan, leven met elkaar mee en hebben humor om te lachen. Waar ik met grote stappen aan de slag wil gaan, neem jij de tijd om op een rijtje te zetten wat er moet gebeuren. In de studie naar de ervaringen van vrouwen hebben we onze compassie voor deze patiëntengroep kunnen omzetten naar wetenschappelijke output. Marije, op dezelfde datum begonnen we aan dit avontuur. En naast het onderzoekswerk zijn we heel wat andere avonturen aangegaan, van wekelijkse sportlesjes tot een reis door Argentinië. Je brengt altijd een verfrissende gezelligheid met je mee en bent een echte levensgenieter. Ik wil jullie allebei heel erg bedanken voor de steun tijdens mijn PhD traject en ben trots dat jullie naast me willen staan tijdens de verdediging.

Lieve schoonfamilie, Engelien, Jos, Linda, Bennie en de kinderen. Bedankt dat jullie me in de familie hebben opgenomen. Ik ben graag bij jullie. Niet iedereen heeft zo'n sportieve schoonpa om mee op de racefiets te toeren. Het is leuk dat we extra raakvlakken hebben door ons werk in de zorg.

Leave heit en mim, van kinds af aan hebben jullie me mijn eigen studiekeuzes laten maken en deze onvoorwaardelijk gesteund. Het doorlopen van ongeveer alle onderwijsvormen heeft me gevormd tot wie ik ben. Ik ben jullie dankbaar voor de liefde en zorg waarmee wij zijn grootgebracht. Heit, ik waardeer de oprechte interesse en warmte waarmee jij contact maakt met de mensen om je heen. Mim, je gedrevenheid om iedereen met elkaar verbonden te houden en activiteiten te initiëren maakt dat we een fijne band hebben als gezin. Froukje, als ik een zus zou mogen kiezen, koos ik jou. Ik ben blij met je en bewonder je veerkracht. Dat jij en Ids nu een kindje verwachten is iets waar ik erg naar uit kijk. Paulus, mijn kleine broertje, je bent me vooruitgesneld in het behalen van je PhD, fantastisch. Jouw liefde voor taal deel je met je liefde Femke en zij heeft mij geweldig geholpen met het Engelse redigeerwerk. Veel dank daarvoor.

Mijn lieve Marco, vanaf het tweede jaar van dit traject kwam jij in mijn leven en wat hebben we het goed samen. Je maakt me aan het lachen, je laat me tot rust komen en samen trekken we de wijde wereld in. Ik kijk uit naar nog veel meer met jou.