Tracheoesophageal Speech. A Multidimensional Assessment of Voice Quality

van As, C.J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
CHAPTER 2

Laryngectomy
CHAPTER 2

ABSTRACT

In this chapter an introduction into the topic of total laryngectomy is given. It starts with an overview of the classification, demographics, etiology, symptoms and treatment of laryngeal cancer. Next, the history and anatomy of total laryngectomy are described, followed by an explanation of the new voice source and the tracheoesophageal voice production mechanism, a description of three different voice rehabilitation methods and an overview of other physical and psychosocial consequences of total laryngectomy.
2.1 INTRODUCTION

Voice production is only one of the three functions of the human larynx. Primary functions of the larynx are breathing and protection of the lungs (closing the airway during swallowing and heavy lifting, coughing for clearance of the airways). The larynx connects the lower airways (lungs) and the upper airways (throat, mouth and nose), and at the same time separates the respiratory and digestive tract.

During breathing the larynx remains open, while during swallowing the true vocal cords, false vocal cords and epiglottis close off the larynx to prevent food from entering the lungs. Furthermore, vocal fold closure is used to generate pressure for activities like lifting and defecation.

Voice production is a secondary, not life-saving, function of the larynx. It is, nevertheless, a significant function of the larynx, since voice production enables, together with the vocal tract and articulators, oral communication with the environment.

After a total laryngectomy all these functions of the larynx are lost and need to be substituted as much as possible. The loss of the normal voice is generally considered to be the most obvious consequences of total laryngectomy, but also pulmonary function problems and the loss of the sense of smell are important consequences that require rehabilitation. This thesis concerns the rehabilitation of voice after total laryngectomy. Therefore, in this chapter, mainly aspects concerning voice production will be introduced, but other consequences of total laryngectomy will be discussed briefly as well. Section 2.2 describes the classification, demographics, etiology, symptoms and treatment options of laryngeal cancer. Although total laryngectomy can be necessary for other malignancies than laryngeal cancer, these are not described any further in this introduction, since they are much less frequent and the end result of the surgical procedure is very much comparable. In section 2.3, the history of total laryngectomy and its anatomy are described, as well as aspects of the surgical reconstruction and the, sometimes necessary, operation for metastatic disease in the neck. In section 2.4, the three types of voice rehabilitation after total laryngectomy are described. In section 2.5, the localization and characteristics of the new sound source (further referred to as neoglottis), the surgical techniques used to influence the neoglottis and the mechanism of tracheoesophageal voice production are described. In section 2.6, other physical and psychosocial consequences of total laryngectomy will be addressed.

2.2 LARYNGEAL CANCER

Laryngeal cancers can present themselves at different localizations and at different stages of the disease. Their classification is shown in section 2.2.1, the demographics of laryngeal cancer in the Netherlands are given in section 2.2.2, the etiology in section 2.2.3, the symptoms in section 2.2.4 and the treatment options of laryngeal cancer are described in section 2.2.5.

2.2.1 CLASSIFICATION OF LARYNGEAL CANCER

For the classification of laryngeal cancer a TNM-staging system is used, as proposed by the International Union against Cancer (UICC) (1997) and the American Joint Committee on Cancer (AJCC) (Ferlito, 1993). This classification is used to describe Tumor stage (T1-4) on the bases of its localization within the larynx, its size, and its invasion in underlying structures, regional metastases (N0-3) on the bases of their absence or presence of lymph Nodes of the neck, and distant Metastases (M0-1) on their absence or presence, for example in the lungs. The UICC TNM-staging is shown in Table 2.1. For the description of the localization of the tumor the larynx has been divided into three regions: supraglottis
(epiglottis, ventricular folds, arytenoids, and aryepiglottic folds), glottis (true vocal folds, anterior commissure, posterior commissure) and subglottis (from below true vocal folds to the first tracheal ring).

Table 2.1. TNM-staging of glottic laryngeal tumors (obtained from UICC (1997)).

<table>
<thead>
<tr>
<th>T-stage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tx</td>
<td>Tumor that cannot be assessed by rules</td>
</tr>
<tr>
<td>T0</td>
<td>No evidence of primary tumor</td>
</tr>
<tr>
<td>T1s</td>
<td>Carcinoma in situ</td>
</tr>
<tr>
<td>T1</td>
<td>Tumor limited to the vocal fold(s) (may involve anterior or posterior commissures) with normal vocal cord mobility</td>
</tr>
<tr>
<td>T1a</td>
<td>Tumor limited to one vocal fold</td>
</tr>
<tr>
<td>T1b</td>
<td>Tumor involves two vocal folds</td>
</tr>
<tr>
<td>T2</td>
<td>Tumor extends to supraglottis and/or subglottis, and/or with impaired vocal cord mobility</td>
</tr>
<tr>
<td>T3</td>
<td>Tumor limited to the larynx with vocal fold fixation</td>
</tr>
<tr>
<td>T4</td>
<td>Tumor invades through thyroid cartilage and/or extends to other tissues beyond the larynx, e.g. to oropharynx or tissues of the neck</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N-stage</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Nx</td>
<td>Lymph node metastasis that cannot be assessed by rules</td>
</tr>
<tr>
<td>N0</td>
<td>No regional lymph node metastasis</td>
</tr>
<tr>
<td>N1</td>
<td>Metastasis in a single ipsilateral lymph node, 3 cm or less in greatest dimension</td>
</tr>
<tr>
<td>N2</td>
<td>Metastasis in a single ipsilateral lymph node, more than 3 cm but not more than 6 cm in greatest dimension; or in multiple ipsilateral lymph nodes, none more than 6 cm in greatest dimension; or in bilateral or contralateral lymph nodes, none more than 6 cm in greatest dimension</td>
</tr>
<tr>
<td>N3</td>
<td>Metastasis in a lymph node more than 6 cm in greatest dimension</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>M-stage</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mx</td>
<td>Metastasis that cannot be assessed by rules</td>
</tr>
<tr>
<td>M0</td>
<td>No distant metastasis</td>
</tr>
<tr>
<td>M1</td>
<td>Distant metastasis</td>
</tr>
</tbody>
</table>

2.2.2 DEMOGRAPHICS

In the Netherlands approximately 700 patients a year are diagnosed with laryngeal cancer. Laryngeal cancer is one of the most frequently detected head and neck cancers in the Netherlands. In total, the head and neck area hosts a wide variety of relatively rare cancers, which all together amount to almost 10% of new patients with cancer in the Netherlands (Visser et al., 1998a). In 1995, in the Netherlands, 622 males and 121 females were diagnosed with laryngeal cancer. Due to a change in consumption pattern and life-style the number of females diagnosed with laryngeal cancer is increasing. For example, in 1995 it occurred in 7.9 out of 100.000 males and 1 out of 100.000 females: ratio 7.9:1, while in 1970 this ratio was 23:1 (Van den Berg et al., 1996). Most of the laryngeal tumors are glottic tumors. Of the laryngeal tumors detected in males between 1989 and 1995, 66% was situated at the glottic level, 30% was situated in the supraglottic region and only 1% in the subglottic region (Visser et al., 1998b). In females the percentage of glottic tumors was 41%, supraglottic tumors 54%, and subglottic tumors 3% (Visser et al., 1998b). The age of diagnosis is mostly situated between 60 and 74 years. For males 4% of the laryngeal tumors is detected between the age of 30 and 44, 27% between the ages of 45 and 59, 52% between the ages of 60 and 74, and 18% above the age of 75 (Visser et al., 1998b). In females 6% of the laryngeal tumors is detected between the ages of 30 and 44, 38% between the ages of 45 and 59, 40% between the ages of 60 and 74, and 16% above the age of 75 (Visser et al., 1998b). The 5-year survival rate of laryngeal cancer is 67% (Visser et al., 1998b).

Apart from laryngeal cancer, a total laryngectomy can also be necessary in hypopharyngeal cancer. This type of cancer occurs less frequent than laryngeal cancer. In the Netherlands, in
1995 133 males were diagnosed with hypopharyngeal cancer and 28 females. In 1995 it occurred in 1.8 of 100,000 males and 0.5 of 100,000 females (Visser et al., 1998b). The age distribution of this type of cancer is comparable to that of laryngeal cancer. The 5-year survival rate is with 39% considerably lower than that for laryngeal cancer (Visser et al., 1998a).

2.2.3 Etiology

Several studies have shown that the risk of developing a malignant change in the squamous epithelium of the larynx is strongly related to smoking and alcohol consumption (Burch et al., 1981; Hinds et al., 1979; Muscat and Wynder, 1992; Trigg et al., 2000; UICC, 1997; Wynder et al., 1976). The combined use of tobacco and alcohol increases the risk of developing laryngeal cancer even more. Wynder et al. (1976) found that the relative risk of developing laryngeal cancer was 7 times greater for smokers who smoked more than 35 cigarettes a day, than for non-smokers. For those who smoked over 35 cigarettes a day and used more than 7 ounces of alcohol daily, this relative risk was 22 times greater than for people who did not smoke or drink.

Other factors, of which relationships with laryngeal cancer are suggested, are passive smoking (Guyatt and Newhouse, 1985; Somerville et al., 1988), environmental contaminants (Wynder et al., 1976), gastroesophageal reflux (Morrison, 1988; Ward and Hanson, 1988), and occupational related co-contaminants like asbestos or other agents (Hinds et al., 1979; Muscat and Wynder, 1992; UICC, 1997).

2.2.4 Symptoms of Laryngeal Cancer

Hoarseness is the most common symptom of laryngeal cancer (Graham, 1983). In tumors of the true vocal folds it appears as an early symptom because it interferes with the vibration of the vocal folds. In tumors of the supraglottic or subglottic region it may be a later symptom of tumor invasion in the true vocal folds. Symptoms of large tumors are discomfort and pain in the throat, referred pain in the ear, dysphagia, fetor (bad smell of the patients’ breath), dyspnoea (difficulty of breathing), neck swelling, sensation of a lump in the throat, and chest infection (as a result of aspiration during swallowing) (Graham, 1983).

2.2.5 Treatment of Laryngeal Cancer

Usually, small laryngeal tumors (T1, T2) are treated with radiotherapy or, sometimes, with laser surgery. Also T3 and the smaller T4 tumors are often treated with radiotherapy or, increasingly, with a combined treatment of radiotherapy and chemotherapy.

A total laryngectomy is necessary in the more extensive T4 tumors, and also in tumors that recur after radiotherapy. For tumors that extend to tissues outside the larynx (pharynx or cervical esophagus) larger surgical resections, often combined with a reconstruction, are necessary (see section 2.3.3).

In the Netherlands approximately 250 patients are laryngectomized every year (Ackerstaff et al., 1990).

In case of the presence of lymph node metastasis, the total laryngectomy is combined with a unilateral or bilateral radical neck dissection (depending on the localization of the lymph nodes) (see also section 2.3.4).

When a total laryngectomy has been performed as a primary treatment, postoperative radiotherapy may be necessary as well, depending on the histopathological aspects of the removed tumor.
2.3 TOTAL LARYNGECTOMY

This section starts with a historical review on total laryngectomy in section 2.3.1. In section 2.3.2 the anatomy before and after total laryngectomy is described. In section 2.3.3 an overview of surgical reconstruction methods is given, and the procedure of a radical neck dissection is described in section 2.3.4.

2.3.1 HISTORICAL REVIEW

In 1873 the first total laryngectomy for a case of laryngeal cancer was performed by the Viennese surgeon Billroth, and described by his fellow Gussenbauer (1874). Already in this first laryngectomy concerns were expressed about rehabilitation of the vocal and pulmonary system. For reasons of surgical safety, the anterior pharyngeal wall was not closed, leaving the patient with a large, unrepaired, defect above the tracheostoma. Gussenbauer (1874) describes an artificial larynx that was used after wound healing. This artificial larynx was connected to the tracheostoma with one tube and to the pharyngeal defect with another tube. A drawing of this first ‘artificial larynx’ is shown in Figure 2.1. With this artificial larynx a metallic reed (g) that is activated by the expired air produced voice sound. Aspiration of saliva and leakage of food was prevented by a flap of the pharyngeal tube (h). The report says the patient was able to speak with a clear voice that was loud enough to be heard at the other side of a large hospital room. It is also reported that the only difference with a normal voice was that it sounded different and monotonous and that much more energy was needed to speak.

After this first laryngectomy, improved versions of this artificial larynx have been made, but this came to an end when in 1894 Gluck and Sorensen succeeded in primary closure of the pharyngeal defect. This ended the ability using artificial larynges using the pharyngeal defect, but enabled esophageal speech. In the following years also air-blown reed-vibrator artificial larynges, and since the 1940’s electronic artificial larynges have been developed. Blom (1979) gives a detailed description of the various artificial larynges that have been developed over the years.

In 1932 Guttman reports about a laryngectomized patient, who performed an operation on himself. The patient used a hot ice pick to create a fistula between the trachea and the hypopharynx. After that, by placing his thumb over the tracheostoma, he was able to force air from his lungs through this small fistula into the mouth. According to the report, he had a decidedly good voice and was able to make himself heard across the extent of a room. Although aspiration must have been a problem in this first report of tracheoesophageal speech, this was not mentioned as such.

In 1958 Conley et al. (1958; 1959) describe two surgical techniques that enable the free flow of air from the trachea into the esophagus without the passage of food or saliva into the trachea. Over the years, several more surgically created shunts between trachea and esophagus, enabling the use of pulmonary air for voice production, have been described (Calcaterra and Jafek, 1971; Arslan, 1972a; Arslan, 1972b; Asai, 1972; Amatsu, 1980; Griffiths, 1980; Hall and Arnesen, 1981; Staffieri, 1981; Brandenburg et al., 1985; Brandenburg et al., 1991). Also surgical voice methods have been described that used a
pharyngocutaneous fistula and a connective voice prosthesis between the tracheostoma and this fistula (Edwards, 1974; Sisson et al., 1975; Taub, 1975). Although voice production was improved by means of those surgical voice restoration methods, complications like leakage through the fistula, pulmonary aspiration and stenosis of the fistula limited their use.

In 1980 Singer and Blom were the first to report about the use of a valved prosthesis, inserted in a surgically created fistula between the trachea and the esophagus (tracheoesophageal (TE) fistula) that enabled voice production by means of pulmonary air and solved the problem of aspiration (Singer and Blom, 1980). After this first voice prosthesis, several others like the Panje Voice Button (Panje, 1981), the Groningen prosthesis (Nijdam et al., 1982), the Nijdam voice prosthesis (Van den Hoogen et al., 1996b) and the Provox® voice prosthesis (Hilgers and Schouwenburg, 1990) and Provox2® voice prosthesis (Hilgers et al., 1997) followed. Nowadays the Blom-Singer and Provox® prostheses are the most widely used voice prostheses. The voice prosthesis can be inserted at the time of surgery.
(primary puncture), but it can also be placed by means of a secondary puncture some time after surgery.

Recently, a voice-producing element, to be inserted into an existing voice prosthesis, has been developed to improve speech quality after laryngectomy (De Vries et al., 2000). The design is based on the lip principle and is especially developed to enable a higher pitched voice for female and a better voice quality for patients with a hypotonic neoglottis. After in vitro testing both for males and females a potentially good functioning prototype could be selected (De Vries et al., 2000). In vivo tests in 6 laryngectomized patients showed that the current prototype seems beneficial in female laryngectomies with a hypotonic pharyngoesophageal (PE) segment only and that pitch regulation of this prosthetic voice is possible, but yet limited (Van der Torn et al., 2001).

Apart from these developments, also attempts have been made to perform transplantation of the larynx. In 1969 Kluyskens et al. (1969; 1970) were able to transplant a human larynx in a patient with laryngeal cancer. Unfortunately this patient died 8 months after surgery because of recurrent tumor; the authors think that this was most probably due to the immunosuppressive treatment that was given to the patient to avoid rejection of the transplant. In an article on the status of laryngeal transplantation in 1974 Tucker mentions that of the four major areas revascularization, reinnervation, prevention of host rejection, and justification, only the first two have been met successfully. He states that until the third criterion is satisfied, one is probably not justified to make further attempts at laryngeal transplantation in humans. However, in 1991 Strome concludes in an article on the future of laryngeal transplantation, that it will be a viable alternative for a select group of patients before the turn of the century; a prediction that he fulfilled in 1998 with his first successful laryngeal transplant in a non-cancer patient.

2.3.2 Anatomy

During a total laryngectomy the following structures are removed:

- Hyoid bone
- Epiglottis
- Thyroid cartilage
- Cricoid cartilage
- First two or three tracheal rings

After removal of these structures the digestive tract is open at the position where the larynx was attached to the pharynx, and the upper and lower airways are disconnected. The digestive tract is re-established by closure of the pharyngeal mucosa and the pharyngeal muscles. The trachea is bowed forward and sutured to the skin of the neck. After this operation, the patient breathes in and out via this opening (tracheostoma) in the neck and can eat via the re-established pharyngeal tract. In Figure 2.2 a schematic drawing of the anatomical situation before total laryngectomy is shown. In Figure 2.3 a schematic drawing of the anatomical situation after total laryngectomy is shown. With the anatomical situation shown in Figure 2.3, only two types of voice rehabilitation are possible: speech with an electrolarynx (see section 2.5.1) and/or esophageal speech (see section 2.5.2).
Figure 2.2. Schematic drawing of the anatomical situation before total laryngectomy.

Figure 2.3. Schematic drawing of the anatomical situation after total laryngectomy. In this case, when there is no voice prosthesis in situ, the possible types of voice rehabilitation are esophageal speech and the use of an electrolarynx.
In the patient group reported on in this thesis, a Provox® or Provox®2 voice prosthesis (Hilgers et al., 1997; Hilgers and Schouwenburg, 1990) (see Figure 2.4 and Figure 2.5) is inserted in a fistula between the trachea and the esophagus (see Figure 2.7) during the surgical procedure of total laryngectomy. This voice prosthesis acts as a one-way valve through which air can be directed into the esophagus by closing the tracheostoma. The one-way hinged valve is molded in one piece with the device to form an integral part with the prosthesis. The valve only opens when the exhaled pulmonary air is diverted through it, and remains closed during swallowing and thereby prevents food or saliva from entering the lungs. This type of voice rehabilitation with a voice prosthesis is called tracheoesophageal speech (see section 2.4.3). The vibrating structure causing the voice sound in this type of speech is situated in the esophagus at the level of the upper esophageal sphincter. In this thesis the term neoglottis will be used when referring to this. More detailed information about the neoglottis is given in section 2.5.1.

![Figure 2.4. Schematic drawing of a Provox® voice prosthesis that is inserted in a fistula between the trachea and the esophagus. When pulmonary air enters the voice prosthesis from the tracheal side, the valve opens and allows the air to pass the esophagus and cause vibration of the neoglottis.](image)

After a total laryngectomy the functions of the nose during breathing (heating, moisturizing and filtering of the inhaled air) are lost, the patient now breathes directly through the tracheostoma. It has been shown that the use of a heat and moisture exchanger (HME) at the entrance of the tracheostoma can substitute for these functions (Hilgers et al., 1991; Ackerstaff et al., 1993; Ackerstaff et al., 1995). The patients in the present study all used a Provox® HME (Hilgers et al., 1996) with a special valve on top so that it can be closed easily by a finger for speaking (see Figure 2.6). Such an HME is held in place at the entrance of the tracheostoma by means of a special adhesive. A schematic drawing of the anatomical situation after total laryngectomy with a Provox®2 voice prosthesis in situ and a Provox® HME in front of the tracheostoma is shown in Figure 2.7).
Figure 2.5. Photograph of the esophageal side of a Provox® voice prosthesis, with the valve opened.

Figure 2.6. Provox® HME (Heat- and Moisture Exchanger) that is used in front of the stoma for heating, moisturizing and filtering of the inhaled air. The valve on top of it can be occluded by a finger during speaking.
2.3.3 RECONSTRUCTIVE SURGERY

In case of a larger extension of the laryngeal tumor into the pharynx or even the cervical part of the esophagus, the total laryngectomy has to be combined with a partial resection of the pharynx (total laryngectomy with partial pharyngectomy), a complete circumferential resection of the pharynx (total laryngopharyngectomy), or a complete circumferential resection of the pharynx and the esophagus (total laryngopharyngoesophagectomy).

When a total laryngopharyngectomy or a total laryngopharyngoesophagectomy has been performed, the sphincteric structures that form the neoglottis after standard total laryngectomy have been removed as well. This has implications for the quality of the voice. Of course, with the large resections, anatomical structures need to be reconstructed. Below the different reconstruction methods are described. When available, the results regarding voice quality after such reconstructions are described too.

Pectoralis Major (PM) Flap

A Pectoralis Major flap is a myocutaneous (muscle and skin) flap of the chest. This flap remains attached to the feeding vessels under the clavicle and is rotated into the surgical defect of the pharynx. The PM flap is an important tool for the repair of specific pharyngeal defects after partial pharyngectomy (Baek et al., 1981). Deschler et al. (1998) found that tracheoesophageal voice quality after this type of reconstruction does not differ significantly from voice after standard laryngectomy for acoustic parameters, but perceptual analysis does
reveal significant differences. Hilgers et al. (1995) found that with primary insertion of a Provox® voice prosthesis after the application of the PM-flap, a relatively good voice was achieved, which was often comparable to that after standard total laryngectomy.

Radial Forearm Flap (RFAF)
A radial forearm flap is a free revascularized fasciocutaneous (fascia and skin) flap of the forearm. This flap is usually tubed and used for reconstruction of circumferential defects of the pharynx after total laryngopharyngectomy (Kelly et al., 1994; Akin et al., 1997). With RFAF reconstruction an acceptable voice can be achieved, but this is found to be significantly different from voice quality after standard total laryngectomy, especially regarding perceptual evaluations (Deschler et al., 1994).

Jejunal graft
A jejunal graft is a free revascularized transplant of the jejunum that can be used to reconstruct circumferential defects of the pharynx after total laryngopharyngectomy (Grasl, 1993; Hoorweg et al., 1994). After this type of reconstruction peristaltic activity of the graft may cause trouble in speech and swallowing. Acceptability and intelligibility of speech after this reconstruction is reduced compared to speech after standard total laryngectomy. The voice is regularly blocked by the autonomous peristalsis of the jejunum segment, and sounds ‘wet’ due to the continuous production of intestinal secretions (Haughey and Forsen, 1992). However, patients preferred the use of speech by means of a voice prosthesis to other forms of communication (Mendelsohn et al., 1993). In a study by Hilgers et al. (1995) approximately half of the patients reconstructed with a free jejunal graft had a good voice. The voice often sounded ‘wet’, which downgraded the voice from good to reasonable, although intelligibility was still good.

Gastric pull-up
This type of reconstruction is used in patients who need a total laryngopharyngoesophagectomy. The stomach is pulled up and replaces the esophagus and pharynx. Two types of reconstruction are possible, a tubed gastric pull up (the stomach is surgically formed into a tube shape and then pulled-up) (Marmuse et al., 1994), or the entire stomach is pulled up. Voice results after tube reconstruction were more often judged to be good than after full stomach transfer; after the latter the voice has more often an amphoric sound with little strength (Hilgers et al., 1995).

2.3.4 RADICAL NECK DISSECTION
When there are regional lymph node metastases in the neck, a total laryngectomy needs to be combined with a unilateral or bilateral neck dissection. With a radical neck dissection, the following structures are removed:

- Sternocleidomastoid muscle
- Omohyoid muscle
- Internal jugular vein
- Spinal accessory nerve (N.XI)
- Submaxillary salivary gland

Depending on the level and size of the lymph nodes, it is sometimes possible to perform a modified radical neck dissection in which the internal jugular vein and/or the spinal accessory nerve can be saved.
CHAPTER 2

2.4 VOICE REHABILITATION

In this section the three best known and presently still widely applied methods of voice rehabilitation after total laryngectomy are described. In section 2.4.1 the electrolaryngeal speech is described, in section 2.4.2 esophageal speech, and in section 2.4.3 tracheoesophageal speech. In the Netherlands, tracheoesophageal voice restoration is the most widely used method, although still most of the patients are also motivated to learn esophageal speech as well (Knegt et al., 1999). In Table 2.2 differences between the three methods of speaking are summarized.

Table 2.2. The initiator, source, resonator and articulators of the three methods of voice production after total laryngectomy.

<table>
<thead>
<tr>
<th></th>
<th>Electrolarynx</th>
<th>Esophageal Speech</th>
<th>Tracheoesophageal speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initiator</td>
<td>Battery</td>
<td>Esophageal air reservoir</td>
<td>Pulmonary air</td>
</tr>
<tr>
<td>Voice Source</td>
<td>Vibrating membrane</td>
<td>Neoglottis</td>
<td>Neoglottis</td>
</tr>
<tr>
<td>Resonator</td>
<td>Vocal tract</td>
<td>Vocal tract</td>
<td>Vocal tract</td>
</tr>
<tr>
<td>Articulator</td>
<td>Articulatory organs</td>
<td>Articulatory organs</td>
<td>Articulatory organs</td>
</tr>
</tbody>
</table>

2.4.1 ELECTROLARYNGEAL SPEECH

An electrolarynx is a hand-held device with an electromagnetically vibrating membrane. When this membrane is held against the skin of the neck or the floor of the mouth, the vibrations, at one fixed frequency, are transmitted through the skin into the vocal tract, where the speaker modulates the sound with the articulatory organs into speech (Norton and Bernstein, 1993). Although this interferes with articulation, it is also possible to direct the sound that is generated by the membrane directly into the mouth by means of a small silicone tube. The Servox Inton® is an example of a well-known electrolarynx; other types are Optivox® and Truetone®. The speech produced by an electrolarynx sounds mechanical, robot-like and monotonous, although the use of an intonation button (Servox Inton®) can make it sound less monotonous (Van Geel, 1983). Also, the sound is produced continuously and thus not interrupted for voiceless consonants. Another drawback is that it is sometimes difficult to position the electrolarynx to the neck. In case of edema or scar tissue of the neck, transmission of the vibrations through the skin is impossible. In the Netherlands, an electrolarynx is only used in case of failure of tracheoesophageal or esophageal speech. In a Dutch study of 1994 Ackerstaff et al. report on 63 laryngectomized patients, only 2 of them (3%) used electrolaryngeal speech. In the U.S.A., on the other hand, 39% of the patients were reported to use an electrolarynx as the major method of alaryngeal speech (Culton and Gerwin, 1998). In Figure 2.8 a schematic drawing of electrolaryngeal speech is shown.

2.4.2 ESOPHAGEAL SPEECH

Esophageal speech can be used as a method of voice rehabilitation when the anatomical situation after surgery is like the one shown in Figure 2.3, and it can also be used when a voice prosthesis is in situ as shown in Figure 2.7. In esophageal speech, air is first brought from the mouth into the esophagus, and then during eructation this air is brought back into the mouth, causing vibrations of the neoglottis. This method can be compared with belching, although the air that creates the voice does not come from the stomach, but from the upper part of the esophagus (just below the neoglottis). In comparison with the several liters of pulmonary air used in normal speech, the amount of air that can be used for voice production
is small, approximately 80 cc (Van den Berg and Moolenaar-Bijl, 1959). Not all laryngectomized patients are able to acquire esophageal speech. Success percentages of 33% (Gates et al., 1982a), but also above 80% (De Beule and Damsté, 1972) are reported. Three methods of air intake have been described: the swallowing method, the inhalation method, and the injection method.

With the swallowing method, swallowing actions are used to fill the esophagus with air after which speech is produced. With this method esophageal air is completely exhausted after a short time, and the intake of air takes a long time (Van den Berg and Moolenaar-Bijl, 1959). Swallowing interrupts speech audibly every 6 to 10 syllables (Moolenaar-Bijl, 1953).

With the inhalation method air is quickly inhaled through the stoma into the lungs. Thereby, an instantaneous negative pressure is created within the thoracic cavity, which is also reflected in the esophagus. This sudden drop in negative pressure in the esophagus causes the air from the mouth and nose to be drawn into the esophagus (Salmon, 1979). According to Moolenaar-Bijl (1953), using the inhalation method breathing is superficial and frequent, and the speech expires in syllables.

With the injection method air is injected into the esophagus by use of the tongue. The tongue is placed against the palate, the soft palate closes the nasal cavities, and the tongue is lifted and retracted in order to compress the air in the pharyngeal cavity (Van den Berg and Moolenaar-Bijl, 1959). Moolenaar-Bijl (1953) describes a method of air injection by which certain consonants are used for air injection. She concludes that the pronunciation of $p$, $t$, and $k$ makes it easier to take air into the esophagus, and that the pronunciation of a $p$, $t$ or $k$ and air-injection into the esophagus perhaps even coincide. The injection method requires less
energy and interrupts speech much less than swallowing does and it makes speech more smooth and continuous than inhaling does (Moolenaar-Bijl, 1953).

Esophageal speech has been described as a harsh voice of low pitch and loudness (Gates et al., 1982b). It is generally lower, with a smaller pitch range and less modulation possibilities than normal speech (Schilling and Binder, 1926). Snidecor and Curry (1959) studied 6 excellent esophageal speakers and found a mean pitch of 63 Hz and a decreased speech rate compared to normal speakers. Better esophageal speech has been related to a faster speaking rate (Hoops and Noll, 1969). Smith et al. (1978) found that esophageal speech was characterized by varying degrees of vocal roughness and that the magnitude of vocal jitter present in vowels was substantially larger than in normal speakers and speakers with laryngeal/vocal disturbance.

2.4.3 TRACHEOESOPHAGEAL SPEECH

This section on tracheoesophageal speech is divided into two sections. First, in section 2.4.3.1 tracheoesophageal voice rehabilitation and results for tracheoesophageal speech are described. In section 2.4.3.2 methods of stoma occlusion are described, followed by their influence on tracheoesophageal speech.

2.4.3.1 Voice rehabilitation and characteristics of tracheoesophageal speech

Two decades after the introduction of the first useful voice prosthesis in 1980 by Singer and Blom (1980), tracheoesophageal speech has become the most widely used method of vocal rehabilitation after total laryngectomy in many countries in the Western world (Hilgers et al., 1999). In Figure 2.7 the anatomical situation after total laryngectomy, with a voice prosthesis that enables tracheoesophageal speech, is shown. As in esophageal speech, the voice source is situated at the level of the pharyngoesophageal segment. In tracheoesophageal speech however, as in normal speech, the pulmonary air is used for voice production. When the tracheostoma is closed, the exhaled air from the lungs is diverted through the voice prosthesis into the esophagus, where it sets the neoglottis into vibration.

The success rates for tracheoesophageal speech are higher than those of esophageal speech, with success rates of up to 90% (Op de Coul et al., 2000). Studies comparing tracheoesophageal and esophageal speech show that tracheoesophageal speech is better than esophageal speech in several respects. Pindzola and Cain (1988) performed perceptual evaluation of fluency, pitch/quality, speech rate, inflection, and overall acceptability. Tracheoesophageal speech was found to be better than esophageal speech on all features. Also Nieboer et al. (1988), investigating both methods of voice rehabilitation by means of perceptual evaluation, found that, on average, tracheoesophageal speakers were rated to have better voice quality than esophageal speakers. They state that the better voice quality is mainly due to a greater speaking rate and fluency, made possible by the voice prosthesis. Baggs and Pine (1983) studied tracheoesophageal, esophageal and normal speakers on acoustic parameters of frequency and duration. The differences found between the groups indicated that the tracheoesophageal speakers are more like normal speakers than the esophageal speakers. Williams and Barber Watson (1987) studied tracheoesophageal, esophageal and electrolaryngeal speech by means of perceptual evaluation of voice quality, pitch, loudness, intelligibility, rate of speaking, visual presentation during speech, extraneous speaking noise, and overall communicative effectiveness. Their results show that tracheoesophageal speech is perceived as superior to esophageal and electrolaryngeal speech.

Robbins et al. (1984; 1984a; 1984b) used acoustic measures of frequency, perturbation, and duration to investigate tracheoesophageal, esophageal and normal speech. They found that tracheoesophageal speech is more similar to normal speech than esophageal speech for frequency and duration variables, and that tracheoesophageal speech is louder than
esophageal speech. Debruyne et al. (1994) performed acoustic measures of tracheoesophageal and esophageal speech. They studied measures applied to the acoustic waveform (fundamental frequency and waveform perturbation) and measures in the frequency spectrum (harmonic prominence, spectral slope). The results showed that in tracheoesophageal speech more often a detectable fundamental frequency was present that was fairly stable. There was also a tendency to more clearly defined harmonics in tracheoesophageal speech. Bertino et al. (1996) studied fundamental frequency and waveform perturbation in tracheoesophageal and esophageal speakers and found that the tracheoesophageal voices are more likely to provide a stable fundamental frequency. They also noted a tendency for tracheoesophageal speech towards more clearly defined harmonics; jitter and shimmer values in tracheoesophageal speech were more similar to normal speech than jitter and shimmer values in esophageal speech. Max et al. (1996) studied tracheoesophageal speech and esophageal speech for the parameters maximum phonation time, maximum number of syllables on one air intake, dynamic range, frequency range, and maximum intensity level. The two groups differed for all parameters except for dynamic loudness range and fundamental frequency range, the maximum performance on the duration and intensity measures was always less for the esophageal speakers than for the tracheoesophageal speakers. Qi and Weinberg (1995) used inverse filtering of flow functions for the investigation of voice source waveforms produced by tracheoesophageal, esophageal and normal speakers. Their results show that the characteristics of the signals appeared to be highly variable for the tracheoesophageal and the esophageal speakers, particularly in comparison with the homogenous pattern of source waves produced by the normal speakers. The source waves of the normal speakers show a homogenous, quasi-periodic, triangular shape. For the esophageal speakers only 3 of the ten speakers (30%) showed waveforms comparable to those of normal speakers, other types of waveforms were single impulse pattern, diplophonic pattern, and long duty cycle pattern. Nine of the twelve tracheoesophageal speakers (75%) showed rounded, fuller-appearing, and/or differentially skewed approximations to normal sound waves, one showed a diplophonic pattern, and two showed a long duty cycle pattern. From the finding that the tracheoesophageal speakers tend to resemble the normal pattern more than the esophageal speakers, the authors conclude that coupling pulmonary volumes and closure of the lower airways might have an influence on the normalizing pattern of pharyngoesophageal source outputs.

Trudeau and Qi (1990) investigated female tracheoesophageal speech. They found that speech of the women closely resembled speech of male tracheoesophageal speakers. The characteristics of tracheoesophageal speech, including fundamental frequency, were highly similar, regardless of speaker gender. In a study by Ackerstaff et al. (1994), the women reported to be less satisfied with their voice than the man, primarily because of the more obvious decrease in pitch.

Besides these studies on voice quality, also studies on intelligibility have been performed. Results for different languages have shown that intelligibility is generally lower for esophageal speech than for tracheoesophageal speech (Cullinan et al., 1986; Hammarberg et al., 1990; Hammarberg et al., 1992; Max et al., 1997). However, also for tracheoesophageal speech decreased intelligibility was noted in comparison to normal speech, especially regarding the voiced-voiceless distinction (Cullinan et al., 1986; Doyle et al., 1988; Hammarberg et al., 1990; Hammarberg et al., 1992; Max et al., 1997; Nord et al., 1992; Polak and Roeleven, 1999). Boon-Kamma (2001) recently finished a Master’s thesis on consonant intelligibility of the same group of laryngectomized speakers participating in the study described in the present thesis. The intelligibility of initial and final Dutch consonants was investigated by transcription of 48 phonetically balanced words and furthermore the intelligibility of spontaneous speech was studied by means of transcription. The transcriptions
were performed by four naive listeners and four experienced listeners (speech-language pathologists). It was found that also in Dutch the intelligibility of tracheoesophageal speech is decreased. The average percentage of correctly identified consonants was 87%, for the initial consonants it was 82% and for the final consonants it was 91%. The final consonants are thus perceived better than the initial consonants. One striking finding was the deletion of the /h/ in 16% of the cases. In concordance with the results of the studies performed in other languages, in Dutch the voiced-voiceless distinction was also found to be a problem. A voiceless consonant was produced as a voiced one in 16.8% of the cases and a voiced consonant as a voiceless one in 15.7%. The transcription of the spontaneous speech showed that 91% of the syllables was correctly transcribed, leading to 90% correct word intelligibility and 73% correct sentence intelligibility. Correlation coefficients showed that the intelligibility of spontaneous speech was related to the intelligibility of the consonants in the words. Apparently, the intelligibility of the different consonants is important for the intelligibility of spontaneous speech; the redundancy in spontaneous speech most probably only compensates for a part of the consonant errors. Not only consonant intelligibility is decreased, but also vowel intelligibility. In 1999 Oubrie finished a Master’s thesis in which the vowels produced by the same speakers as in the present study were studied (Oubrie, 1999). The vowels were studied in h-vowel-t context. Results of transcription by 25 naive listeners showed an average percentage of 50% correctly identified vowels; in addition to these perceptual results, formant frequencies of the first and second formant appeared to be different from normal values in a large number of vowels. This decreased vowel intelligibility is thought to be caused by differences in the vocal tract due to the change in anatomy after total laryngectomy.

Although pulmonary air can be used for tracheoesophageal speech production, the voice quality of tracheoesophageal speech is highly variable and still very deviant from normal voice. In a pilot study preceding this thesis, Van As et al. (1998b) have shown that tracheoesophageal speech of male laryngectomized speakers differed significantly from normal speech of male speakers on all acoustic parameters measured by MDVP (Kay Elemetrics Corp, Lincoln Park, USA) except for the fundamental frequency. Although the standard deviation of the fundamental frequency was much higher in the tracheoesophageal speakers (large inter-speaker variation) the average value was comparable to the average value of the normal speakers. In this respect, however, it should be kept in mind that pitch extraction was not possible in 3 of the 21 patients and that in some of the remaining 18 patients only small pieces of the voice sample could be analyzed for pitch, since the other parts of the signal were too aperiodic for meaningful pitch extraction. This implies that only the ‘good’ parts of the voice sample have been analyzed. The results of the perceptual evaluation performed in that study also showed significant differences between the tracheoesophageal speakers and the normal speakers. Tracheoesophageal male voices were found to be more deviant, ugly, unsteady, weak, dull, breathy, low and deep than normal male voices. Also, the maximum phonation time appeared to be different between tracheoesophageal (13 s on average) and normal speakers (26 s on average).

The use of acoustic parameters would enable more objective analysis of voice quality after total laryngectomy. It is, however, still a problem to acquire a reliable analysis for the highly irregular voices due to problems with the pitch detection in those voices. There is a need for a reliable measure that can be performed for the entire range of tracheoesophageal voice qualities and that is related to the perceptual evaluation of voice quality. Therefore, in the studies described in this thesis expert listeners, specially trained for this purpose, have performed perceptual evaluations and efforts have been made to find acoustic parameters that can be obtained reliably. Relations between those acoustic parameters and the perceptual evaluations will reveal whether the acoustic parameters make sense as an objective measure.
Apart from the perceptual evaluations performed by the trained expert listeners, also perceptual evaluations performed by naïve listeners are part of the studies in this thesis, in order to gain insight in the opinion of the public who meet laryngectomized patients in daily life situations.

2.4.3.2 Stoma occlusion

Stoma occlusion is an important aspect in tracheoesophageal speech. Stoma occlusion can be accomplished by means of:

- Digital occlusion by thumb or finger directly onto the tracheostoma;
- Digital occlusion by thumb or finger on top of an HME with speech valve (Provox® HME, (Hilgers et al., 1996));
- Hands-free speech by means of an automatic speech valve (Blom et al., 1982).

Several studies have investigated the influence of the use of an automatic speech valve (which closes as a result of the increased expiratory airflow when the patient starts to speak) on tracheoesophageal speech. Results have been contradicting. Some studies reported no speech differences between digital occlusion directly on the tracheostoma and occlusion by means of an automatic speech valve (Pauloski et al., 1989; Zanof et al., 1990; Van den Hoogen et al., 1996a; Williams et al., 1989) found that the total pause time and the percentage of total reading time occupied by pauses were longer, and that the maximum phonation time was longer in the case of occlusion by an automatic speech valve. Blakely and Podraza (1987) report that tracheoesophageal speech produced with occlusion by an automatic speech valve was associated with more extraneous sound energy than tracheoesophageal speech produced with direct digital occlusion. Williams et al. (1990) found the same effect regarding extraneous speaking voice, but noted that the visual presentation during speech of a patient was better with the use of an automatic speech valve. Fujimoto et al. (1991) suggest that an automatic speech valve adversely affects the conversational intelligibility of tracheoesophageal speech, a result that is most likely caused by valve noise.

Van As et al. (1998b), studying acoustic and perceptual characteristics of tracheoesophageal speech compared to normal speech, found that the patients in the study using an automatic speech valve (Blom-Singer® ATV) or a heat and moisture exchanger (Freevent®, Provox® HME) had longer maximum phonation times compared to patients using direct digital occlusion of the stoma. It was also noted that the patients using a valve or HME had lower shimmer values during counting. This finding was more intensively investigated in a next study (Van As et al., 1998a) in which direct digital occlusion of the stoma and digital occlusion on top of the valved Provox® HME were studied within the same patient by means of acoustic analysis, maximum phonation time, and dynamic loudness range, all investigated in sustained /a/. Acoustical analyses (only possible in 13 out of the 20 voices) showed no statistically significant differences between both occlusion methods. However, the study showed a positive influence on the maximum phonation time and the dynamic range. The average maximum phonation time was 12.7 s in the digital occlusion group and 16.0 s in the HME occlusion group. The average dynamic loudness range was 23.2 dB in the digital occlusion group and 26.5 dB in the HME occlusion group. It appeared that 75% of the patient group benefited from either one or both of these improvements.

Since it has been shown that stoma occlusion by means of a Provox® HME positively influences some aspects of tracheoesophageal speech, this manner of stoma occlusion was chosen for all patients participating in the present study in order to avoid differences caused by the use or non-use of the Provox® HME.
2.5 THE NEOGLOTTIS

In this section localization and characteristics of the new voice source (neoglottis) are described (section 2.5.1), followed by the techniques used for influencing the voice source (section 2.5.2). Some thoughts about the mechanism of tracheoesophageal voice production are presented here as well (section 2.5.3).

2.5.1 LOCALIZATION AND CHARACTERISTICS OF THE NEOGLOTTIS

The new voice source, referred to as neoglottis, pseudoglottis, or pharyngoesophageal (PE) segment, plays an important role in voice production. After removal of the larynx, the remaining pharyngeal opening is closed, and thereby the alimentary tract is reconstituted. After total laryngectomy, the pharyngoesophageal segment not only serves to receive food from the oral cavity and transport it to the esophagus and stomach, but also to produce the voice sound. Discovering the site or source of vibration in substitute voice production has been the subject of investigation for several years. In Robe et al. (1956) a bibliography of the various investigations is given. At present it is a well-known fact that the source of vibration (neoglottis) is situated at the level of the muscles of the upper esophageal sphincter and/or the middle and inferior constrictor pharyngeus muscles (pharyngoesophageal segment). This upper esophageal sphincter is part of the normal anatomy and thus present in all patients and also in ‘normals’. It prevents gastroesophageal reflux and food from coming up. The belching sound that some ‘normals’ are able to produce at will, (which can be compared to esophageal speech) is caused by this sphincter too. When a reconstruction of the pharynx and/or esophagus is carried out in addition to the total laryngectomy, the pharyngoesophageal segment is mostly removed too (see section 2.3.3). In those cases it is not clear which structure (if any) is vibrating when the air passes the (neo)esophagus. Since patients after these types of reconstructions are included in the present study too, we chose to use the word neoglottis throughout this thesis, since the term pharyngoesophageal segment is not applicable for all patients.

One of the methods frequently used for investigation of the neoglottis is videofluoroscopy. In Figure 2.8 an X-ray photograph of a neoglottis is shown.

Studies on esophageal voice so far were mainly conducted to gain insight in factors influencing the acquisition of esophageal speech and were focused on discovering the site or source of vibration, causing the substitute esophageal voice. Several researchers found that the origin of the sound was situated at the level of the cricopharyngeus muscle (Lindsay et al., 1944; Mooolaar-Bijl, 1951; Robe et al., 1956). It was also thought that it was not the mucus membrane that was set into vibration and caused the sound, but the accumulated mucus above the neoglottis (Brewer et al., 1975). A number of researchers found that the ability to acquire esophageal speech was related to anatomical and morphological characteristics of the neoglottis, such as dilatability of the hypopharynx and shape of the neoglottis (Vrticka and Svoboda, 1961), length and cervical level of the neoglottis (Bentzen et al., 1976), form of the neoglottis (Damsté and Lerman, 1969), extent of surgery (Richardson, 1981), and tonicity of the pharyngoesophageal (PE) segment (McIvor et al., 1990; Sloane et al., 1991). Others could not find any relationship between good speech and variations in the anatomy and morphology of the neoglottis (Kirkner et al., 1963) or thought that the acquisition of voice was merely related to psychological factors (Diedrich and Youngstrom, 1966; Richardson, 1981).

Videofluoroscopic studies regarding tracheoesophageal speech showed that the visual characteristics of the vibratory segment of the tracheoesophageal and esophageal speakers were similar (Wetmore et al., 1985; Isman and O'Brien, 1992). In a study by Van Weissenbruch et al. (2000) it has been shown that also in tracheoesophageal speech tonicity of the PE-segment was related to the quality of the speech.
The three studies mentioned above are the only three studies using videofluoroscopy recordings of tracheoesophageal voice production. Only Van Weissenbruch (2000) mentions a relation between tonicity and voice quality. There is thus a need for a more basic investigation of various anatomical and morphologic characteristics of the neoglottis in tracheoesophageal speech in relation to tracheoesophageal voice quality. Videofluoroscopy recordings, however, cannot provide information on the vibratory characteristics of the neoglottis. In normal laryngeal voices, stroboscopy is used as a tool to create a virtual slow motion image of vocal cord vibration. However, stroboscopy is less suited for visualization of the vibratory characteristics of the neoglottis, since these vibrations are often too irregular to trigger the stroboscopic light pulses and also because of the deformation of the neoglottis that can occur when electroglottographic electrodes or a microphone are used on the neck to detect the fundamental frequency. Information on the vibratory characteristics of the neoglottis and the anatomy and morphology of the neoglottis as seen from above is thus rare. In the present study a solution for this problem was found in the use of a recently developed digital high-speed imaging system that is used for research purposes and is under development for clinical use.
2.5.2 Techniques to influence the tonicity of the neoglottis

Hypertonicity and spasm of the PE-segment are considered to be the major reasons for voice failure after total laryngectomy. Singer and Blom (1981) found that 12% of the tracheoesophageal speakers failed to acquire tracheoesophageal speech because of pharyngoesophageal spasm. They performed a secondary myotomy (=surgically cutting of the muscle so that contraction (and thus hypertonicity) is prohibited) of the cricopharyngeus and pharyngeal constrictor muscles after which all 14 patients eventually achieved fluency, three patients needed a repeated myotomy for residual spasm. That study also reports about 8 patients in whom a myotomy was performed as a primary intervention during secondary tracheoesophageal puncture surgery. Those patients were found to have pharyngoesophageal spasm on the air insufflation test that was performed preoperatively. They all received intelligible speech postoperatively. Good results of myotomy have also been reported by others (Chodosh et al., 1984; Henley and Souliere, 1986). In a study performed by Mahieu et al. (1987) the effects of a myotomy of the inferior constrictor pharyngeal muscles, the cricopharyngeus muscle and a small section of the upper esophageal musculature was investigated. It appeared that fewer patients in the myotomized group had a good tracheoesophageal voice than in the non-myotomized group, more than half of the patients had a certain degree of breathiness during tracheoesophageal voice production. Their results suggest that not only too high, but also too low tonicity has a negative effect on voice production. They suggest to use a less rigorous form of myotomy, and advise that the surgeon performs a myotomy until he feels that the tension of the upper esophageal sphincter is diminished, but not totally absent.

Another method to influence the tonicity of the pharyngeal wall is neurectomy (=surgical cutting of the nerve branches so that contraction of the muscle (and thus hypertonicity) is prohibited) of the pharyngeal nerve plexus (Singer et al., 1986). Other studies propose the use of non-closure of the pharyngeal constrictor muscles (preventing muscular contraction and thus hypertonicity) during surgery (Olson and Callaway, 1990, Clevens et al., 1993), or half-closure of the constrictor muscles (Deschler et al., 2000). In the past few years Botulinum toxin injections (causing paralysis (and thus decreased tension) of the muscle) in tracheoesophageal speech failures with a hypertonic or spastic pharyngoesophageal segment have shown good results (Terrell et al., 1995; Hoffman et al., 1997; Zormeier et al., 1999). This method is less invasive than a secondary myotomy. In a study by Van Weissenbruch (2000) on the differences in tracheoesophageal speech for different methods of additional surgery it has been shown that tracheoesophageal speech was most successful (81%) in the group who underwent a primary myotomy and neurectomy. Functional tracheoesophageal speech was obtained in 95% of the patients with either a single myotomy or a myotomy combined with a neurectomy. Most tracheoesophageal speech failures were observed in the patient group with no additional surgery (26%).

A part of the patients participating in this thesis underwent a myotomy or a neurectomy at the time of primary surgery. In Chapter 3, Table 3.1 this information is shown. Although the patients in the present study were participating irrespective of their surgical history, an attempt is made to study the influence of these surgical techniques.

2.5.3 The mechanism of tracheoesophageal voice production

Although the localization of the neoglottis is not a point of discussion anymore, the mechanism of tracheoesophageal voice production is. Normal voice production is a complex mechanism, which is described as an aerodynamic-myoelastic event (Van den Berg, 1958). This myoelastic-aerodynamic theory implies that the vocal folds are actuated by the stream of air delivered by the lungs (aerodynamic) and that the fundamental frequency of the vibration
depends on the effective mass and the stiffness of the vocal folds that are regulated by the sustained innervation of internal and external laryngeal muscles (myoelastic) and by the activity of the lungs and the associated resonators.

Because of the large aerodynamic differences between esophageal and tracheoesophageal speech, and the scope of this thesis, only results regarding the mechanism of tracheoesophageal voice production are reported here.

The aerodynamic-myoelastic theory of normal voice production cannot be directly applied to the mechanism of tracheoesophageal speech. Since pulmonary air is used as a driving force in tracheoesophageal speech, the aerodynamic principle might be applicable. It is however not very clear whether the neoglottis has an active role (myoelastic) in tracheoesophageal speech production or that it is rather a passive vibration generated by the aerodynamic forces. Also, the abductory and adductory possibilities of the neoglottis are not well understood. Some evidence for tracheoesophageal voice production being an aerodynamic-myoelastic event has been found. According to Moon and Weinberg (1987), tracheoesophageal speakers were capable of adjusting their voicing sources on a myoelastic basis to influence fundamental frequency change. They conclude that this result, together with findings that confirm the aerodynamic contributions to tracheoesophageal phonation suggests that tracheoesophageal voice production should be considered as an aerodynamic-myoelastic event. However, the authors also state that although the tracheoesophageal speakers were capable of adjusting their voice sources to influence fundamental frequency, this active adjustment was not a consistent form of fundamental frequency mediation. The subjects exhibited fundamental frequency variation that could be associated with voice source adjustment, but none of the subjects exhibited source adjustment on a consistent basis. From these findings they conclude that there are important differences between tracheoesophageal and normal voice production. In a study of Mohri et al. (1994) who used electromyography (EMG) to study muscle activity of the neoglottis, it has been shown that the act of phonation was characterized by two bursts of muscle activity. The first burst was obtained during inspiration and the second, more intense, burst during phonation. These results suggest a myoelastic contribution of the neoglottic muscles. A recent study on aerodynamic and myoelastic properties of tracheoesophageal voice production, performed by Deschler et al. (1999), also supports the theory that tracheoesophageal voice production is an active process incorporating the myoelastic properties of the neoglottis as well as the aerodynamic properties of airflow.

Results of intelligibility studies suggest that some patients are and some patients are not able to produce the voiced-voiceless distinction (Hammarberg et al., 1990; Polak and Roeleven, 1999; Lundström and Hammarberg, 1999; Boon-Kamma, 2001). The ability to produce this distinction would suggest the presence of the possibility of abduction and adduction in analogy to the normal vocal folds.

Apparently the aerodynamic-myoelastic function of the neoglottis differs, amongst the speakers. At present it is not clear what causes the differences between the speakers.

Although these types of investigations could provide useful information on the mechanism of tracheoesophageal speech, this falls beyond the scope of this thesis in which the focus is on the quality of the voice as such.

2.6 PHYSICAL AND PSYCHOSOCIAL CONSEQUENCES OF TOTAL LARYNGECTOMY

Although this thesis concerns voice quality after total laryngectomy, this is not the only implication of a total laryngectomy. Due to the disconnection between the upper and lower airways patients breath in and out through the tracheostoma, and the nasal functions during breathing (heating, moisturizing and filtering), as well as the sense of smell, are thus compromised.
Principal complaints of laryngectomized patients are daily sputum production (98%), coughing (68%), the need for frequent forced expectoration (57%), and frequent stoma cleaning (37%) (Hilgers et al., 1990). Other frequently mentioned problems are fatigue (30%) and sleeping problems (24%). Fortunately, however, only few patients exhibit clinically relevant levels of anxiety (5%) and depression (7%). Relations are found between the respiratory symptoms, voice rehabilitation and several aspects of daily life, like fatigue, sleeping problems, lack of social contacts and psychological distress. Efforts have been made to solve the respiratory problems; the use of a heat and moisture exchanger (HME) has shown to improve pulmonary function and thereby also psychosocial functioning of laryngectomized patients (Hilgers et al., 1991; Ackerstaff et al., 1993; Ackerstaff et al., 1995). In these studies, it could be shown that the consistent use of an HME has a significant positive influence on the quality of life of the patient and also improves voice quality by diminishing pulmonary complaints. The more recent availability of the valved Provox® HME has improved patient compliance considerably (Ackerstaff et al., 1998).

In a study by Ackerstaff et al. (1994) 95% of the patients reported deteriorated sense of smell and 44% reported deteriorated sense of taste. Other complaints concerned daily nasal discharge (38%), serious problems swallowing food (25%), and eructation (32%). Recently, at the Netherlands Cancer Institute, efforts have been made to investigate olfactory problems in more detail (Van Dam et al., 1999). In that study olfactory tests showed that approximately one-third of the patients were able to smell. The patients that were able to smell, more often used a variety of methods to smell than the patients that were not able to smell; in most patients the method consisted of active use of the facial muscles. The observations from the study by Van Dam et al. (1999) led to the development of a so-called Nasal Airflow Inducing Maneuver ("Polite Yawning") (Hilgers et al., 2000). With this method, patients learn to create an underpressure in the oral and nasal cavity by lowering the jaw, tongue and floor of mouth. By creating this underpressure in the oral and nasal cavity with the lips closed, aired is pulled into the nose and can thereby reach the olfactory epithelium. This method has been taught to 44 laryngectomized patients; 33 of those patients were non-smellers according to olfactory tests and 15 of them converted to smellers (Hilgers et al., 2000). It can be concluded from this study that it seems possible to rehabilitate the sense of smell in a considerable percentage of the patients.

CONCLUDING REMARKS

In this chapter an overview was given on laryngeal cancer, total laryngectomy and voice rehabilitation after total laryngectomy. Results from the literature were reported and related to the purpose of this thesis. In the next chapter the patient group participating in the present study is described, after which in the following five chapters the results of the different parts of the study are given.

REFERENCES


