Institute of Phonetic Sciences,
University of Amsterdam,
Proceedings 23 (1999), 37- 57

VOICE QUALITY OF TRACHEOESOPHAGEAL SPEAKERS RELATED TO VIDEOFLUOROSCOPIC OBSERVATIONS AND MEASUREMENTS.

Corina van As[1,2], *Bas op de Coul*[3], *Frank van den Hoogen*[3], *Florien Koopmans-van Beinum*[1], *and Frans Hilgers*[2]

[1] Institute of Phonetic Sciences, University of Amsterdam
[2] Department of Otolaryngology-Head&Neck Surgery, The Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital, Amsterdam
[3]Department of Otolaryngology/Head&Neck Surgery, University Hospital St. Radboud, Nijmegen

Abstract

This paper concerns the relationship between tracheoesophageal voice quality and anatomical and morphologic characteristics of the neoglottis as studied by means of visual assessment and quantitative measures of videofluoroscopic recordings of the neoglottis during phonation. Subjects were 39 laryngectomized speakers, all using tracheoesophageal speech by means of a Provox voice prosthesis. Voice quality was judged by 4 speech-language pathologists, who are experienced in voice treatment of laryngectomized patients, and were trained for this purpose. These judgments of voice quality were used to divide the speaker group into three subgroups with a 'good', 'reasonable', or 'poor' voice. Results from the visual assessment of the video-fluoroscopic recordings showed that the voice quality is significantly related to the tonicity of the neoglottis, and the appearance of a neoglottic bar during phonation. Results from the quantitative measures of digitized images of the videofluoroscopic recordings also showed significant relationships with the voice quality. The 'minimal distance between the neoglottic bar and the anterior esophageal wall at rest' and the 'minimal distance distance between the neoglottic bar and the anterior esophageal wall during phonation' were significantly smaller in the 'good' voice group, and the ratio between the relative increase in the maximal subneoglottic distance from rest to phonation was significantly larger in the 'good' speaker group. From this study it can be concluded that anatomical and morphologic characteristics of the neoglottis, as well as quantitative measures of the neoglottis, are significantly related to voice quality. These quantitative measures are promising for use as a more standardized manner of evaluating the neoglottic characteristics.

1. Introduction

Since the introduction of the voice prosthesis (Singer and Blom, 1980), two decades ago, tracheoesophageal voice has nowadays become a widely used and successful method of voice restoration after total laryngectomy (Hilgers et al., 1999). The main advantage of this type of voice rehabilitation above conventional esophageal voice is that it is pulmonary driven. In conventional esophageal speech the esophagus serves as an air reservoir and only small amounts of air that are trapped into the esophagus are available for speech. Tracheoesophageal voice is proven to be closer to normal laryngeal voice than esophageal voice regarding acoustical characteristics (Debruyne et al., 1994; Robbins, 1984; Robbins et al., 1984; Robbins et al., 1984), perceptual characteristics (Williams and Watson, 1987), and intelligibility (Max et al., 1997).

With the shift of tracheoesophageal voice rehabilitation becoming the method of choice after total laryngectomy (Hilgers et al., 1999), the main issue in voice rehabilitation is not anymore the ability to acquire speech, as it was in esophageal speech. In tracheoesophageal voice rehabilitation around 88% of the patients acquires a fair to excellent voice (Op de Coul et al., submitted), while for esophageal speech only 25% of the patients acquires voice (Gates et al., 1982). However, the tracheoesophageal voice still shows a large variability in its quality (Van As et al., 1998).

The new sound source, being the so-called neoglottis, pseudoglottis, or pharyngoesophageal (PE) segment, is considered to play an important role in the voice production. Throughout this paper the term neoglottis will be used. One of the methods most frequently used for investigation of the neoglottis is videofluoroscopy. Videofluoroscopic studies of the characteristics of the neoglottis are frequently performed for esophageal voice (Bentzen et al., 1976; Damsté and Lerman, 1969; Daou et al., 1992; Kirchner et al., 1963; Lindsay et al., 1944; McIvor et al., 1990; Richardson, 1981; Robe et al., 1956; Sloane et al., 1991; Smith et al., 1966; Vrticka and Svoboda, 1961). Studies on esophageal voice were mainly conducted to get insight in factors influencing the possibility to acquire esophageal speech. Early studies of esophageal speech were focused on discovering the site or source of vibration, which was causing the substitute esophageal voice. Several researchers found that the origin of the sound was situated at the level of the cricopharyngeus muscle, which serves as a neoglottis (Lindsay et al., 1944; Moolenaar-Bijl, 1951; Robe et al., 1956). It was also thought that it was not the mucous membrane which was set into vibration and caused the sound, but the accumulated mucous above the neoglottis (Brewer et al., 1975). A number of researchers found that the ability to acquire esophageal speech was related to 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 PE-segment (McIvor et al., 1990; Sloane et al., 1991). Others researchers could not find any relationship between good speech and variations in the anatomy of neoglottis (Kirchner et al., 1963), or even thought that the acquisition of a voice was merely related to psychological factors (Diedrich and Youngstrom, 1966; Richardson, 1981). Also age of the patient was found to be an important factor for the ability to learn esophageal speech (Bentzen et al., 1976; Smith et al., 1966). Regarding tracheoesophageal speech, videofluoroscopic studies are rare. Probably since tracheoesophageal speech has a shorter history. Wetmore et al. (1985) and Isman and O'Brien (1992) found that the visual characteristics of the vibratory segment of the tracheoesophageal speakers were similar to the visual characteristics of the vibratory segment of esophageal speakers. Both these studies lack assessment of voice quality in tracheoesophageal speech in relation to the characteristics of the neoglottis. However, concluding that the vibratory segments in both types of alaryngeal speech are similar is not enough evidence to conclude that voice quality in both types of alaryngeal speech is similar. The difference in air supply, causes a large difference between tracheoesophageal and esophageal voice, which is reflected in the acoustical and perceptual characteristics. Therefore, the most favorable conditions of the neoglottis to acquire good esophageal speech, may differ from the most favorable conditions of the neoglottis to acquire good tracheoesophageal speech. Thus, there is a need for an extensive study of the relation between voice quality and characteristics of the neoglottis in tracheoesophageal voice.

In the present study voice quality of tracheoesophageal voice is related to visual assessment of anatomical and morphologic characteristics of the neoglottis and to quantitative measures of the neoglottis in videofluoroscopic recordings, by means of a newly developed evaluation protocol. Apart from the relationship between voice quality and the characteristics of the neoglottis, also the relationship between the visual assessment of the neoglottis and the quantitative measures of the neoglottis will be studied. Quantitative measures would enable a more standardized manner of evaluation. Therefore, it is important to investigate this relationships, since relationships between both evaluation methods could provide evidence for substitution of the more subjective visual assessment by standardized quantitative measures.

2. Subjects and methods

2.1 Subjects and clinical parameters

This paper describes part of a larger project in which alaryngeal speech is investigated. The speakers participating in this study are also described in a study in which even more patients participated (Van As et al., 1999). In the present study, subjects were 39 laryngectomized patients, who all used tracheoesophageal speech by means of a Provox®2 voice prosthesis (Hilgers et al., 1997). Special care was taken to obtain a sample of patients, in which all variations of the whole speaker group are represented. Therefore, also female patients were studied, patients with a poorer voice quality and patients with a reconstruction of their pharynx and/or esophagus were included in the study. Patients were selected from a group of 173 laryngectomized speakers that are currently in follow-up in the Netherlands Cancer Institute (Op de Coul et al., submitted). All patients, who were asked to participate, agreed to take part and received written information about the study. There were 29 males and 10 females. Ages varied from 47 to 82 years, with a mean of 67 years. The post-operative follow-up varied from 11 months to 18 years, with a mean of 6 years. In 9 patients a partial (myocutaneous pectoralis major flap (n=4)) or full (tubed gastric pull up (n=2), full gastric pull up (n=1), or free radial forearm flap (n=2)) pharyngeal reconstruction was performed. The remaining 30 patients had a standard wide-field total laryngectomy. From now on, in this paper, the group of 30 patients with the standard total laryngectomy will be referred to as the 'standard' group and the 9 patients with the partial or full pharyngeal reconstruction will be referred to as the 'reconstruction' group. In 22 patients of the 'standard' group, during surgery an attempt was made to influence the tonicity (i.e. muscular tension) of the neoglottis, by performing a myotomy of the cricopharyngeus muscle (Mahieu et al., 1987; Singer and Blom, 1981), or a neurectomy of the pharyngeal nerve plexus (Singer et al., 1986). In 5 patients a myotomy of the cricopharyngeus muscle combined with a neurectomy of the pharyngeal plexus was performed, and in 17 patients only a neurectomy of the pharyngeal plexus was performed. A unilateral neck dissection was performed in 13 patients, a bilateral neck dissection in 5 patients. Sixteen patients received primary radiotherapy to treat their laryngeal cancer (these patients were laryngectomized because of recurrence of the tumor), 21 patients received radiotherapy after their total laryngectomy, and 2 patients received no radiotherapy at all. In Table 1 an overview of the clinical information of the patients is given.

Table 1. Clinical parameters of the patient group used in this study.

Parameter	*Subgroups*	*Number (n=39)*
Surgery	Standard total laryngectomy	30
	Reconstruction	9
Sex	Male	29
	Female	10
Tonicity control (standard group, n=30)	Myotomy + Neurectomy	5
	Neurectomy	17
	No	8
Neck dissection	No	21
	Unilateral	13
	Bilateral	5
Radiotherapy	No	2
	Primary	16
	Post-operative	21

2.2 Methods

Recordings for the perceptual evaluation were made of one fixed read-aloud text. The recordings were made with use of the Computerized Speech Lab of Kay Elemetrics (Lincoln Park, NJ, USA). A head-set microphone (AKG-c410) which comes standard with this equipment was used, and via the hardware of this system, with use of the CSL software, the speech data were directly recorded on a DAT-tape by means of a Sony TCD-8 DAT recorder. For the perceptual evaluation the read-aloud texts of all speakers were randomly recorded on another DAT tape, for each speaker the text was repeated until 2,5 minutes were recorded on the listening tape.

The videofluoroscopy recordings were obtained with a Philips Diagnost 92 together with a Panasonic NV-HD650 videorecorder. Off all patients video-fluoroscopic recordings were made of two phonations of the sustained vowel /a/ at a comfortable pitch and loudness level. All recordings were made in lateral view. The patient was asked to take a swallow of Barium and to phonate afterwards. A reference coin was adhered to each patient, in order to enable later quantification of the different dimensions.

2.3 Perceptual evaluation of voice quality

Four speech-language pathologists who were experienced regarding treatment of laryngectomized patients, were trained for doing perceptual evaluations in this patient group. The training session took 4 hours. A DAT tape with read-aloud text of ten tracheoesophageal speakers served as training tape. The training concerned the consistent use of 19 bipolar semantic 7-point scales, and one overall judgment of voice quality in which the voice was judged as 'good', 'reasonable' or 'poor'.

A 'good' voice was defined as 'most similar to normal voice', a 'poor' voice was defined as 'least similar to normal voice', and 'reasonable' was used for the group in between both extremes. During the training the speakers were judged separately, results were compared and interpretation of the scales was discussed. The training ended when the 4 listeners judged three subsequent speakers consistently, within one scale point difference. One week after this training session the 4 speech-language pathologists, independently of each other evaluated the voice quality. In the present study only the results of the overall judgment will be used. The inter-rater reliability calculated by Cronbach's alpha, was 0,88, indicating that the 4 listeners performed a reliable judgment. For the comparison with the videofluoroscopy recordings the speakers were divided into three subgroups on the basis of the evaluation of the overall voice quality. A voice was considered 'good' when at least two out of the 4 listeners evaluated it as such, 'poor' when at least two out of the 4 listeners evaluated it as such. Voices that were judged as 'good' by one listener and 'reasonable' by three listeners, or voices that were judged as 'poor' by one listener and 'reasonable' by three listeners were considered 'reasonable'. It never occurred that a voice was judged as 'good' by one listener and as 'poor' by another listener.

2.4 Visual assessment of anatomical and morphologic characteristics of the neoglottis

First, all videofluoroscopy recordings were jointly viewed by three judges (one ENT specialist, one ENT-resident, and one speech-language pathologist). In this session of two hours the characteristics to be assessed were discussed and chosen. During this first assessment an attempt was made to use the definitions of tonicity as proposed by McIvor et al.(1990) and Van Weissenbruch (1996). These definitions of tonicity do not only include the tonicity during phonation, but also the tonicity during swallowing and at rest. At first, it appeared to be difficult to reach consensus of opinion among the three judges using these definitions. A large number of our patients could not be categorized in one tonicity group, since they did not meet all criteria for this particular group, or they met criteria of different groups. Therefore, we chose to judge the flattening of the neoglottic bar during swallowing, and the appearance of a neoglottic bar at rest, separately from the tonicity of the neoglottis during phonation. For the assessment of the tonicity during phonation, the following definitions were composed. The tonicity of the neoglottis was judged as *'normotonic'* when there was a full or almost full closure of the neoglottis during phonation. The tonicity of the neoglottis was judged as *hypotonic* when there was no closure of the neoglottis during phonation, and the tonicity was judged as *hypertonic* when the neoglottis was fully closed during phonation combined with extreme dilation of the esophagus below the neoglottis during phonation. *Stricture* was defined as a narrowing of the esophagus, which was present equally at rest, during phonation and during swallowing. *Spasm* was defined as complete closure of the neoglottis, with extreme dilation of the esophagus below the neoglottis during attempted phonation with no passage of air through the neoglottis. At the end of this pilot evaluation session a new evaluation protocol was formed, containing both visual assessment of the anatomical and morphologic characteristics of the neoglottis, as well as quantitative measures of the neoglottis.

In a final judgment of three sessions of one hour each the visual assessment of all recording was performed by means of the evaluation protocol. On the same day, all recordings were judged in a consensus judgment of three judges. Two of them were the same as in the first session (ENT-resident and speech-language pathologist) and another ENT-specialist was included to have one judge who was not involved in the development of the assessment form. In this final judgment, with the separate assessment of the neoglottis during phonation, swallowing, and at rest, consensus of opinion was reached easily for all videofluoroscopic recordings. The definitions of tonicity appeared to be very helpful in making a decision. Also the separate judgment of the neoglottis at rest, during phonation, and during swallowing, appeared to be helpful for easy and fast judgment. In Table 2 the characteristics chosen are summarized.

Table 2. Anatomical and morphologic aspects of the neoglottis as used in the consensus judgment by the three judges.

Characteristic	*Visual assessment*
Number of neoglottic bars at rest	0 1 2
Number of neoglottic bars during phonation	0 1 2
Cervical level of the neoglottis at rest	C3-C7
Cervical level of the neoglottis during phonation	C3-C7
Tonicity of the neoglottis during phonation	Hypotonic Normotonic Hypertonic Stricture Spasm
Regurgitation of barium during phonation	Yes No
Stasis of barium on the neoglottic bar during phonation	Yes No
Flattening of the neoglottic bar during swallowing	Yes No

2.5. Quantitative measures of the neoglottis

In addition to the visual assessment of the anatomical and morphologic characteristics, also metrical measures were part of the evaluation protocol. These quantitative measures of the neoglottis were obtained by means of the program Drawer (developed by M.B. van Herk, physicist at the NKI/AvL), initially designed to measure tumor volume for radiotherapy purposes. Two of the three judges (ENT-resident and speech-language pathologist) jointly chose a relevant frame out of the recordings for both phonation and rest. These frames were then digitized with a frame grabber and all files were saved as an image file. Then the two judges jointly calculated the quantitative measures.

In Figure 1 a schematical drawing of the neoglottis is given, together with indications of the measures performed. All measures were in pixels and then transformed to millimeters or square millimeters by use of a coin with a known diameter, which was adhered to the patients' skin during videofluoroscopy.

The following parameters were measured:

Minimal distance during phonation in mm (MINPHON), this is the distance during phonation between the neoglottic bar and the anterior wall of the esophagus (i.e. the width of the neoglottis). In Figure 1 this distance is indicated by the letter 'a';

Minimal distance at rest in mm (MINREST), definitions are the same as for the measurement during phonation;

Maximal subneoglottic distance during phonation in mm (MAXPHON), this is the maximal subneoglottic distance during phonation. This place was defined as the level of maximum distance, the line on which the measurements were made was placed perpendicular on the anterior wall. In Figure 1 the letter 'b' indicates this distance;

Maximal subneoglottic distance at rest in mm (MAXREST), definitions are the same as the situation during phonation;

Size of the neoglottic bar during phonation in mm² (SIZEPHON), this is the area size in square millimeters of the neoglottis in lateral view, in Figure 1 indicated as the gray colored surface;

Size of the neoglottic bar at rest in mm² (SIZEREST), definitions are the same as for the situation during phonation;

Prominence of the neoglottic bar during phonation in mm (PROMPHON), this measures the prominence of the neoglottic bar towards the anterior wall. The line on which the measurement is based is placed perpendicularly to the anterior wall at the most prominent point of the neoglottic bar. In Figure 1 the letter 'c' indicates this line;

Prominence of the neoglottic bar at rest in mm (PROMREST), definitions are the same as for the situation at rest.

In addition to these quantitative measures, also one ratio is calculated. This ratio is chosen in order to reflect the relative increase of the maximal subneoglottic distance during phonation compared to the distance at rest: MAXPHON/MAXREST. This ratio is thought to give an impression of the tension of the closure of the neoglottis. A more tight closure of the neoglottis might be reflected in a relatively larger increase in subneoglottal distance, although this might also be dependent on the rigidity of the subneoglottal tissues.

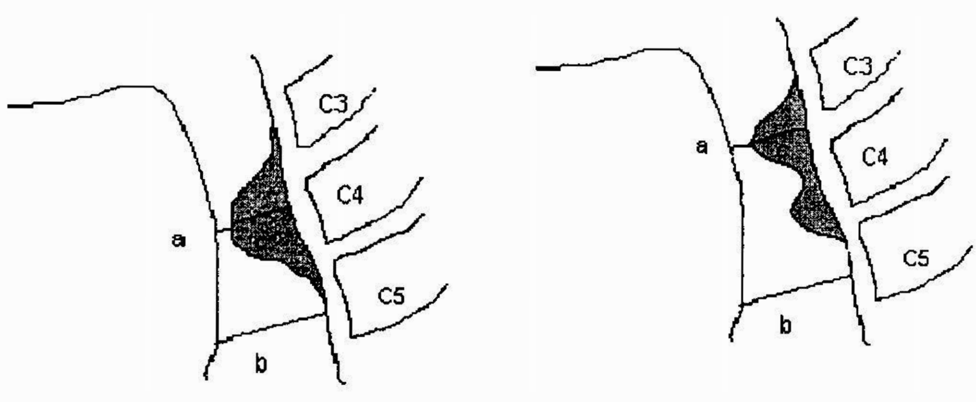

Figure 1. Schematical drawing of the neoglottis after total laryngectomy. On the left hand side an example with one neoglottic bar and on the right hand side an example of a double neoglottic bar. *a* = minimal (neoglottic) distance, *b* = maximal (subneoglottic) distance, *c* = prominence of the neoglottic bar. C3-C4-C5 =cervical vertebrae C3, C4 and C5, gray surface = size of the neoglottic bar.

Later on in this paper the abbreviations of the quantitative measures will be used.

3. Results

The results of the assessment of the anatomical and morphologic characteristics of the neoglottis and the quantitative measures of the neoglottis will be given separately for the 'standard' group and the 'reconstruction' group. Results of the relations between the assessment and the quantitative measures, and between voice quality and assessment, and between voice quality and quantitative measures will be given for the entire patient group (i.e. 'standard' and 'reconstruction' group together).
In this paragraph, first the results of the separate parts of this study will be described: voice quality (3.1), visual assessment (3.2), and quantitative measures (3.3). Then, the relationships between those separate parts and the clinical parameters will be described: clinical parameters versus voice quality (3.4), clinical parameters versus visual assessment (3.5), and clinical parameters versus quantitative measures (3.6). Finally, the relationship between the visual assessment and the quantitative measures will be described (3.7), and the relationship between the voice quality and the visual assessment (3.8) and the relationship between the voice quality and the quantitative measures will be described (3.9).

3.1 Voice quality

The voice quality, judged as 'good', 'reasonable', or 'poor' by the 4 experienced, trained listeners resulted in three subgroups. Voice quality could be judged for 38 of the 39 patients, since one patient died before the speech sample could be obtained. The 'good' group, consisted of 13 patients who are considered to have the 'most similar to normal voice quality'. The 'poor' group consisted of 11 patients who are considered to have the 'least similar to normal voice quality'. The 'reasonable' group consisted of the 14 patients, who fell in between both extremes.

3.2 Visual assessment of the anatomical and morphologic characteristics of the neoglottis

The results of the visual assessment of the anatomical and morphologic characteristics of the neoglottis for both speaker groups is given in Table 3. The cervical levels of the neoglottis is given in a separate table, Table 4.

As can be seen in Table 3, at rest 19 patients within the "standard"group showed one neoglottic bar and 4 patients two neoglottic bars. During phonation 21 patients showed one neoglottic bar and 3 patients two neoglottic bars. The cervical levels where these neoglottic bars were seen are given in Table 4, separately for each patient within the 'standard' group and the 'reconstruction' group. The *cervical level of the neoglottis* differed among the patients, and tended to move upwards for approximately half a vertebra during phonation in most patients.

Table 3. Results of the assessment of the anatomical and morphologic characteristics for the 'standard' group (n=30) and the 'reconstruction' group (n=9) separately.

Parameter	*Assessment*	*Standard (n=30)*	*Reconstruction (n=9)*
Number of neoglottic bars at rest	0 1 2	7 19 4	7 2
Number of neoglottic bars during phonation	0 1 2	6 21 3	4 5
Tonicity of the neoglottic bar during phonation	Hypotonic 'Normotonic' Hypertonic Stricture	9 13 8	6 2 1
Regurgitation of barium during phonation	Yes No	10 20	4 5
Stasis of barium on the neoglottic bar during phonation	Yes No	5 25	4 5
Flattening of the neoglottic bar during swallowing*	Yes No	15 8	1 1

*Only relevant in those patients in whom a neoglottis is present at rest, therefore the sum of the numbers equals the number of patients with the appearance of a neoglottis at rest.

Table 4. Cervical level of the neoglottis of the 'standard' group and the 'reconstruction' group at rest and during phonation. Numbers in this table represent patient numbers. For example patient number 14 has a neoglottic bar at rest at the level of C4, which shifts upwards to the level of C3-C4 during phonation. In the case of two neoglottic bars the letter a or b is added to the patient number.

	Standard (n=30)		***Reconstruction (n=9)***	
Cervical vertebrae	Rest	Phonation	Rest	Phonation
C3		1ab,35	18	18, 37
C3-C4	1a,25	14,20,21b,25		
C4	1b,14,15,20,35,38	6a,7,8,9,15,21a,24,26,30		16
C4-C5	7,8,9,21,24	10,12,22,28	16	
C5	6ab,10,12,22,30,33a	23		
C5-C6	23,34	33b,36		
C6	33b,36	2ab,13		19
C6-C7	2ab,13			
C7		5		17

In the 'standard' group the neoglottis at rest is mostly situated around cervical vertebras C4 and C5. During phonation the neoglottis is situated somewhat higher in most speakers, with an upward shift from in between C4 and C5 to in between C3 and C4.

The number of patients with two neoglottic bars at rest (n=4) or during phonation (n=3) is too small for separate evaluation in statistical analysis. Since the voice quality of the groups with one and two neoglottic bars is thought to be comparable, these groups will be taken together in further analysis. The one patient with stricture will be left out of analysis regarding tonicity, since no meaningful statistical analysis is possible with only one patient in a subgroup.

3.3 Quantitative measures of the neoglottis

In Table 5 the results for the measures MINREST, MINPHON, MAXREST, MAXPHON, SIZEREST, SIZEPHON, PROMREST, PROMPHON, and the ratio MAXPHON/MAXREST are given. Results are separated for both speaker groups.

Table 5. Means and standard deviations (between brackets) of the various measures of the neoglottis for both the 'standard' speaker group and the 'reconstruction' group.

Parameter	***Standard***	***Reconstruction***
	Mean (s.d.)	Mean (s.d.)
MINREST	1.2 mm (1,4 mm)	5.5 mm (4.2 mm)
MINPHON	1.5 mm (2.1 mm)	4.2 mm (3.1 mm)
MAXREST	6.4 mm (3.6 mm)	8.1 mm (4.0 mm)
MAXPHON	10.6 mm (4.7 mm)	9.1 mm (4.6 mm)
SIZEREST	81 mm^2 (80 mm^2)	14 mm^2 (23 mm^2)
SIZEPHON	88 mm^2 (67 mm^2)	28 mm^2 (26 mm^2)
PROMREST	5.2 mm (3.8 mm)	1.8 mm (1.9 mm)
PROMPHON	8.0 mm (5.6 mm)	4.6 mm (5.37 mm)
MAXPHON/MAXREST	2.05 (1.39)	.61 (.30)

Paired t-tests between the measures at rest and during phonation showed statistically significant differences between rest and phonation in the 'standard' group for two measurements. The maximum distance during phonation (MAXPHON) was significantly larger than that at rest (MAXREST) ($p<.05$), and the prominence of the neoglottis (PROMPHON) was larger during phonation than at rest (PROMREST)($p<.05$). For the 'reconstruction' group, no statistically significant differences were found between the quantitative measures at rest and during phonation. These results indicate that there is a dynamic change of the neoglottis between rest and phonation in the 'standard' group. In the 'reconstruction' group the neoglottis is more rigid and does not alter between rest and phonation.

3.4 Relation between the clinical parameters and the voice quality

Chi-square tests for linear-to-linear association did not reveal any significant relationships between the voice quality and the clinical parameters *reconstruction, myotomy, neurectomy, radiotherapy, neck dissection, age, post-operative follow-up, and sex.*

3.5 Relation between the clinical parameters and the visual assessment of anatomical and morphologic characteristics of the neoglottis

In this section results are given about the relationship between the visual assessment and the clinical parameters myotomy, neurectomy, radiotherapy, neck dissection, reconstruction, age, post-operative follow-up, and sex. First, differences between the 'standard' and the 'reconstruction' group on the basis of the assessment are investigated, then the other clinical parameters are evaluated for the 'standard' group only. The number of speakers in the 'reconstruction' group is too small for meaningful statistical analysis in this respect. Statistical analysis was performed by means of Linear-by-Linear Association tests.

Regarding *reconstruction of the pharynx* some statistically significant differences were observed. It was found that the patients who have had a reconstruction, more often show no neoglottic bar at rest, while in the standard group there was more often a visible neoglottic bar at rest ($p<.003$). Patients with a standard total laryngectomy had more often a 'normotonic' or hypertonic neoglottis during phonation than the patients with reconstruction of the pharynx ($p=.019$).

Regarding the relation between the clinical parameters *myotomy, neurectomy, radiotherapy, neck dissection, age, post-operative follow-up, and sex* of the 'standard' group and the visual assessment no statistically significant relations were found.

3.6 Relation between the clinical parameters and the quantitative measures of the neoglottis

In this section results are given about the relationship between the quantitative measures of the neoglottis and the clinical parameters. Based on the assumptions of normality two different statistical methods are used: a t-test for independent samples for the measures MAXREST, MAXPHON, SIZEREST, SIZEPHON, PROMREST, PROMPHON and MAXPHON/MAXREST, and the Mann-Whitney test for the measures MINREST and MINPHON.

As in the former section, first results between the 'standard' and the 'reconstruction' group will be given, and then relations between the quantitative measures and the clinical parameters for the 'standard' group will be given.

Regarding *reconstruction* various statistically significant differences were found. The two-dimensional size of the neoglottis was larger both at rest (SIZEREST) and during phonation (SIZEPHON) in the 'standard' group ($p<.05$). The prominence of the neoglottis (PROMREST and PROMPHON) was smaller in the 'reconstruction' group ($p<.05$). The minimal distance during phonation (MINPHON) and at rest (MINREST) were smaller in the standard group ($p<.05$).

Results for the 'standard' group revealed some relationships between the clinical parameters and the quantitative measures. Whether or not the patient had *a radical neck dissection* appeared to influence some measurements of the neoglottis. The MINREST was smaller in the group without neck dissection ($p<.05$), indicating a more closed neoglottis at rest in the group without a neck dissection. Another factor which showed some differences in the measurements was *age*, it appeared that MINREST was smaller in the younger patient group (<70 years) ($p<.05$). This indicates a narrower neoglottis in the younger patient group and a looser neoglottis in the older patient group. The clinical parameters *myotomy, neurectomy, post-operative follow-up, radiotherapy, and sex* did not show any statistically significant differences for the measurements.

3.7 Relation between the visual assessment and the quantitative measures of the neoglottis

In this section results are given about the relation between the visual assessment of the anatomical and morphologic characteristics of the neoglottis and the quantitative measures of the neoglottis. These relations are based on the results for the whole speaker group together, since the distinction between type of surgery is not relevant for this purpose. Concerning the assumption of normality for the measures MINREST and MINPHON non-parametric Mann-Whitney test are used, and for the measures MAXREST, MAXPHON, PROMREST, PROMPHON, SIZEREST, SIZEPHON, and MAXPHON/MAXREST t-tests for independent samples are used.

The assessment of the *appearance of a neoglottic bar at rest* was related to five measures. A t-test showed that in the subgroup with appearance of a neoglottic bar at rest SIZEREST and SIZEPHON, PROMREST and PROMPHON, and MAXPHON were significantly larger ($p<.05$) than in the subgroup without appearance of a neoglottic bar at rest. According to Mann-Whitney tests, in the subgroup with appearance of a neoglottic bar at rest, MINREST was significantly smaller ($p<.05$) than in the subgroup without appearance of a neoglottic bar at rest.

The assessment of *the appearance of a neoglottic bar during phonation* showed several relations with the quantitative measures. A t-test revealed that in the subgroup with a neoglottic bar during phonation, SIZEPHON and SIZEREST, PROMPHON and PROMREST, and MAXPHON and MAXPHON/MAXREST were significantly larger ($p<.05$) than in the subgroup without a neoglottic bar during phonation. The results of a non-parametric Mann-Whitney test showed that in the group in which a neoglottic bar during phonation was seen, MINREST and MINPHON were significantly smaller ($p<.001$) than in the subgroup in which no neoglottic bar during phonation was seen. These results indicate that the presence of a neoglottic bar during phonation is related to a smaller distance between the neoglottic bar and the anterior wall of the esophagus, a larger subneoglottic distance during phonation, a larger size and prominence of the neoglottic bar during phonation, and a relatively larger increase of the maximal subneoglottic distance during phonation.

The *tonicity of the neoglottic bar during phonation,* divided into the subgroups hypotonicity, normotonicity, and hypertonicity, showed relationships with the quantitative measures as well. These data are given in Table 6. For the parametric measures ANOVA's were used and for the non-parametric measures Kruskal-Wallis tests were used to investigate whether there is a relationship. Then, the ANOVA test was followed by a Post Hoc Tukey test, in order to analyze the differences between the three tonicity-groups. The Kruskal-Wallis test was followed by a Mann-Whitney test for each possible pair of tonicity groups. P-values were then corrected by means of a Bonferroni correction. In Table 6 mean values and standard deviations of the quantitative measures are given for each subgroup of tonicity. Significance levels are given for the group as a whole as well as for the separate subgroups. In this Table a clear distinction between the three levels of tonicity on the basis of the measurements can be seen. For instance the size of the neoglottis during phonation (SIZEPHON), the prominence of the neoglottic bar during phonation (PROMPHON), the minimal distance at rest (MINREST), and the minimal distance during phonation (MINPHON) are distinctive between hypotonicity and normotonicity and between hypotonicity and hypertonicity. While on the other hand the measures of maximal subneoglottic distance (MAXREST and MAXPHON) are distinctive between hypotonicity and hypertonicity and between normotonicity and hypertonicity.

The assessment of *regurgitation of barium during phonation* was also related to the quantitative measures. A t-test showed that in the subgroup in which regurgitation of barium was seen during phonation, SIZEREST and SIZEPHON, PROMREST and PROMPHON, and MAXPHON were significantly smaller ($p<.05$) than in the subgroup in which no regurgitation of barium was seen during phonation. The non-parametric Mann-Whitney test showed that in the group with regurgitation of barium during phonation MINREST and MINPHON were significantly larger ($p<.05$) than in the subgroup without regurgitation of barium during phonation. These results indicate that regurgitation occurs when there is no or only a small neoglottic bar and when the minimal distance between the neoglottic bar and the anterior esophageal wall is large.

Regarding the *cervical level of the neoglottis, stasis of barium above the neoglottic bar during phonation* and *flattening of the neoglottic bar during swallowing* no statistically significant relations with the quantitative measures were found.

Table 6. Relations between the quantitative measures (left hand column) and the tonicity groups (columns 3-5), expressed as p-value for the subgroups together (column 2), and for the subgroups separately (right hand column).

Measure	*P*	*Hypotonic*	*Normotonic*	*Hypertonic*	*Separate P-values*
MIN-REST[2]	.000	4.1 mm (3.4 mm)	.79 mm (1.1 mm)	.49 mm (.94 mm)	Hypo-Normo (.000) Hypo-Hyper (.003) Normo-Hyper (.999)
MIN-PHON[2]	.000	4.6 mm (2.3 mm)	.66 mm (.98 mm)	0 mm (0 mm)	Hypo-Normo (.000) Hypo-Hyper (.000) Normo-Hyper (.393)
MAX-REST[1]	.025	7.0 mm (3.3 mm)	5.0 mm (2.7 mm)	9.3 mm (4.7 mm)	Hypo-Normo (.274) Hypo-Hyper (.286) Normo-Hyper (.020)
MAX-PHON[1]	.000	7.7 mm (2.6 mm)	9.8 mm (4.1 mm)	16.2 mm (3.9 mm)	Hypo-Normo (.240) Hypo-Hyper (.000) Normo-Hyper (.001)
SIZE-REST[1]	.045	31 mm^2 (38 mm2)	84 mm^2 (95 mm^2)	106 mm^2 (72 mm^2)	Hypo-Normo (.127) Hypo-Hyper (.061) Normo-Hyper (.771)
SIZE-PHON[1]	.001	32 mm^2 (45 mm^2)	97 mm^2 (64 mm^2)	118 mm^2 (54 mm^2)	Hypo-Normo (.009) Hypo-Hyper (.003) Normo-Hyper (.655)
PROM-REST[1]	.006	2.5 mm (2.6 mm)	5.2 mm (3.1 mm)	7.2 mm (4.6 mm)	Hypo-Normo (.079) Hypo-Hyper (.006) Normo-Hyper (.330)
PROM-PHON[1]	.000	2.7 mm (3.2 mm)	9.2 mm (3.7mm)	12.9 mm (5.8 mm)	Hypo-Normo (.000) Hypo-Hyper (.000) Normo-Hyper (.107)

[1] parametric measures (ANOVA + Post Hoc Tukey)

[2] non-parametric measures (Kruskal-Wallis + Mann-Whitney & Bonferroni correction)

3.8 Voice quality related to the visual assessment of the anatomical and morphologic characteristics of the neoglottis

Linear-by-Linear Association tests on the three voice quality groups and the visual assessment, revealed that the assessment of the *appearance of a neoglottis during* *phonation* and the *tonicity of the neoglottis during phonation* showed significant relations with voice quality ($p<.05$). A 'good' voice was related to the appearance of a neoglottic bar during phonation, while a 'poor' voice was related to no appearance of a neoglottic bar during phonation. Regarding tonicity a 'good' voice was related to a 'normotonic' or 'hypertonic' neoglottis, and in the 'good' voices there was never a 'hypotonic' neoglottis observed. An overview of the results for both parameters is given in Tables 7 and 8.

Table 7. The relation between the appearance of a neoglottis during phonation and voice quality ($p<.05$).

		Voice Quality			*Total*
		Poor	Reasonable	Good	
Neoglottis during phonation	None	4	5		9
	1 or 2	7	9	13	29
	Total	11	14	13	38

The results from Table 7 clearly show that a 'good' voice always has at least one neoglottic bar during phonation, and never no visible neoglottic bar during phonation. A 'poor' or 'reasonable' voice may have either one, or two, or no apparent neoglottis during phonation.

Table 8. The relation between the tonicity of the neoglottis during phonation and voice quality ($p<.05$).

		Voice Quality			*Total*
		Poor	Reasonable	Good	
Tonicity	Hypotonic	6	8		14
	'Normotonic'	2	4	9	15
	Hypertonic	2	2	4	8
	Total	10	14	13	37

The number of speakers in this analysis is 37 since one speaker with stricture also fell in the 'poor' group, this speaker was not included in statistical analysis. From Table 8 it becomes clear that a 'good' voice is always normotonic or hypertonic and never hypotonic. Hypotonicity is seen in both poor and reasonable voices. Hypertonicity is seen in all speaker groups, but more often in the good group.

The assessment of the *appearance of a neoglottis at rest, regurgitation of barium during phonation, stasis of barium on the neoglottis during phonation,* and *flattening of the neoglottic bar during swallowing* showed no statistically significant relations with the voice quality.

3.9 Voice quality related to the quantitative measures of the neoglottis

As in the former section the results for the 'standard' group and the 'reconstruction' group will be taken together and related to the three different voice quality groups. For the non-parametric measures MINREST and MINPHON Kruskal-Wallis tests were used, followed by Mann-Whitney tests with a Bonferroni correction. For the analysis of the quantitative measures MAXREST, MAXPHON, PROMREST, PROMPHON, SIZEREST, SIZEPHON, and MAXPHON/MAXREST, ANOVA's were used followed by Post-Hoc Tukey tests.

ANOVA's showed that the ratio MAXPHON/MAXREST was significantly different between the speaker groups (p<.05). Kruskal-Wallis test showed statistically significant differences between the voice quality subgroups for MINPHON (p<.05) and MINREST (p<.05). Results for these three quantitative measures are given in Table 9. For the other quantitative measures no relations with the overall judgment of voice quality were found.

Table 9. Means and standard deviations, p-values and p-values between the subgroups of MINPHON, MINREST, and MAXPHON/MAXREST for the three voice groups.

Parameter	*P*	*Poor (n=11)*	*Reasonable (n=14)*	*Good (n=13)*	*Separate p-values*
MIN-PHON	.009	3.6 mm (3.6 mm)	2.5 mm (1.8 mm)	.51 mm (1.0 mm)	Poor-Reasonable (.999) Poor-Good (.021) Reasonable-Good (.021)
MIN-REST	.001	3.8 mm (4.0 mm)	2.5 mm (2.6 mm)	.33 mm (.69 mm)	Poor-Reasonable (.999) Poor-Good (.000) Reasonable-Good (.024)
MAX-PHON/ MAX-REST	.005	1.4 (.61)	1.7 (1.5)	2.9 (1.6)	Poor-Reasonable (.999) Poor-Good (.006) Reasonable-Good (.021)

Results in Table 9 show that on the basis of these three quantitative measures a distinction can be made between a poor and a good voice, or between a reasonable and a good voice. The measures indicate that in a good voice the minimal distance at rest (MINREST) and during phonation is (MINPHON) smaller than in a poor voice, and that the better speakers showed a relatively larger increase of the maximal distance during phonation compared to the maximal distance at rest (MAXREST/MAXPHON).

4. Discussion

In the present study anatomical and morphologic characteristics of the neoglottis and quantitative measures of the neoglottis are described. These results are related to each other and to voice quality, whereas the influence of clinical parameters was investigated as well. The discussion will be ordered in the same way as the results in paragraph 3: voice quality (4.1), visual assessment (4.2), quantitative measures (4.3), clinical parameters versus voice quality (4.4), clinical parameters versus visual assessment (4.5), clinical parameters versus quantitative measures (4.6), visual assessment versus quantitative measures (4.7), voice quality versus visual assessment (4.8), and voice quality versus quantitative measures (4.9).

4.1 Voice quality

On the basis of the perceptual evaluation of the voice quality by 4 trained judges, three subgroups of voice quality were formed. This resulted in 11 speakers with 'poor' voice quality, 14 speakers with 'reasonable' voice quality, and 13 speakers with a 'good' voice quality. Reasons to judge a voice as 'good', 'reasonable', or 'poor', are not easily defined. In this study it was chosen to judge them in relation to 'normal' voice quality. Thus, a voice that was considered as close to normal voice was called 'good', and a voice that was very deviant from normal was judged as 'poor'. There may be several reasons to judge a voice as 'poor', for instance a very bubbly, or a very breathy, or a very tensed voice quality. This is also the case for the 'reasonable' group, there may be several reasons why a voice is not very close to normal, but also not very deviant, for instance slightly bubbly, or rough. A 'good' voice however, is better defined, since it is only judged as 'good' when it has no specific characteristics that are obviously different from normal voices (of course we should keep in mind that a tracheoesophageal voice never really sounds like a normal voice, but it can for instance sound like a 'normal hoarse voice'). The variability of perceptual characteristics that may lead to a 'poor' or 'reasonable' voice quality indicates that a more precise perceptual evaluation is necessary to investigate the relation with specific aspects of tracheoesophageal voice quality. The definition of voice quality is also important in respect to esophageal voice quality, since the definitions of a 'good' esophageal voice are very different from the definitions of 'good' tracheoesophageal voice. Judgments of esophageal voice are based on the ability to produce sound and the number of syllables that can be produced. In tracheoesophageal voice however, there is almost always a voice, and the voice quality is judged on several aspects, like in normal voices.

4.2 Visual assessment of the anatomical and morphologic characteristics of the neoglottis

Visual assessment of the anatomical and morphologic characteristics of the neoglottis already gives an impression of the variations among the speakers. It is also shown that the appearance of a neoglottic bar at rest is not always related to the appearance of a neoglottic bar during phonation, and the other way around. Only a few patients showed a double neoglottic bar. In the 'standard' group all three grades of tonicity were found, in the 'reconstruction' group no hypertonicity was found. The neoglottis was located at the level of C4-C5 in the majority of the patients. This is somewhat higher than reported in studies of esophageal speech in which the level was at C5-C6 in the majority of the patients (Bentzen et al., 1976; Damsté and Lerman, 1969; Diedrich and Youngstrom, 1966). The level tended to rise by approximately one half vertebra from rest to phonation in the majority of the patients, something which was not found in esophageal speech (Diedrich and Youngstrom, 1966). Most probably this upward shift is larger in tracheoesophageal voice due to the pulmonary power. In contrast to earlier studies (McIvor et al., 1990; Van Weissenbruch, 1996), the assessments were separated for the situation at rest, during swallowing, and during phonation. During consensus judgments it appeared to be more efficient to judge these characteristics separately. First, it is often not possible to group a patient in one of the tonicity groups since they do not always meet all the criteria of that subgroup. Second, the relevance of the situation at rest and during swallowing in relation to voice quality is not clear.

4.3 Quantitative measures of the neoglottis

Results of the quantitative measures show a large variability in the speaker group, which is reflected, in the relatively high standard deviations. Statistically significant differences were found between rest and phonation for the maximal subneoglottic distance and the prominence of the neoglottis in the 'standard' group, indicating a dynamic change of the neoglottis from rest to phonation. This phenomenon was not found in the 'reconstruction' group, indicating a more rigid neoglottis.

4.4 Clinical parameters versus voice quality

In this study no relationships were found between the clinical parameters and the voice quality. Earlier studies of tracheoesophageal speech did not study relationships with clinical parameters. In the studies of esophageal speech results differ. Gates et al. (1982) found that less post-operative radiotherapy was related to successful esophageal speech. Bentzen et al. (1976) conclude that the esophageal speech quality does depend on operative factors, like whether the operation is followed by infections and fistula. Smith et al. (1966) found that the size of the tumor did not influence esophageal speech results, but that the patient with a laryngectomy only, had better results than the patients with an additive neck dissection. Richardson et al. (1981) however did not find an influence of neckdissection and operative factors.

4.5 Clinical parameters versus visual assessment of the anatomical and morphologic characteristics of the neoglottis

In the 'standard' group a neoglottic bar at rest was seen more often than in the 'reconstruction' group, and in the 'standard' group more often a 'normotonic' or hypertonic neoglottis was found. These results can be explained by the anatomy, which is more normal in the 'standard' group, through which this group shows muscular activity represented in the appearance of a neoglottic bar at rest, and a higher tonicity of the neoglottis.

4.6 Clinical parameters versus quantitative measures of the neoglottis

For the quantitative measures SIZEREST, SIZEPHON, PROMREST, PROMPHON, MINREST, and MINPHON a statistically significant difference was found between the 'standard' and the 'reconstruction' group. This of course can be explained by the larger extent of surgery in the 'reconstruction' group, leading to less favorable conditions of the neoglottis. The quantitative measures were also influenced by the clinical parameters neck dissection and age. The MINREST was smaller in the group without a neck dissection, and in the patient group below 70 years of age. These results indicate that in the subgroup without a neck dissection and in the subgroup of younger age the neoglottic closure is better.

4.7 Quantitative measures versus visual assessment of the anatomical and morphologic characteristics of the neoglottis: evaluation of the protocol

Investigating the relationship between the results of the visual assessment and the quantitative measures was one of the main goals of this study. Replacement of the more subjective visual assessments of the neoglottic characteristics by more objective and precise quantitative measures, could give the possibility to evaluate videofluoroscopic recordings in a more objective and standardized manner.

Relations are found for all parameters, except for flattening of the neoglottic bar during swallowing and stasis of barium above the neoglottic bar during phonation.

This means that the majority of the visual assessments can be replaced by quantitative measures, and, more interesting, all visual assessment parameters that appeared to be relevant in voice quality can be replaced by quantitative measures.

4.8 Voice quality versus visual assessment of the anatomical and morphologic characteristics of the neoglottis

The most important part of this study was to investigate possible relations between the neoglottis and the voice quality. This particular type of study has to the best of our knowledge, not been performed before for tracheoesophageal speech. Results showed that voice quality was significantly related to the appearance of a neoglottic bar during phonation and to tonicity of the PE-segment. The differences between a 'poor' or a 'reasonable' voice are not that clear in this respect, however the 'good' group always has a neoglottic bar during phonation and is never hypotonic. These results differ from results for esophageal speech, since in esophageal speech a hypotonic neoglottis leads to no voice (McIvor et al., 1990), while in tracheoesophageal voice a poor or reasonable voice exists. McIvor et al. (1990) who used voice quality as a parameter to judge the tonicity of the neoglottis in esophageal speech, found a variable voice with a hypertonic neoglottis. This is in agreement with our results, hypertonicity is found in 'poor' as well as in 'reasonable' and 'good' voices. These findings suggest that tonicity is not easily assessed on the visual characteristics and different grades of hypertonicity may give rise to different voice qualities. However, these results make clear that hypotonicity is an unfavorable condition of the neoglottis regarding voice quality, and should be avoided. Hypertonicity and spasm of the neoglottis can relatively easily be solved by surgical intervention, i.e. myotomy (Mahieu et al., 1987), or medical intervention, i.e. Botox injection (Hoffman et al., 1997) in order to decrease the tonicity. Hypotonicity can only be solved by exerting some digital pressure on the external neck and thereby enabling approximation of the esophageal tissues during phonation, other forms of surgical intervention or phonosurgery of the neoglottis are not yet available for this kind of problems. Since it is nowadays very common to perform a neurectomy of the pharyngeal plexus and a myotomy of the cricopharyngeus muscle during the total laryngectomy, it should be stressed that care should be taken to avoid 'overcorrection' and thereby hypotonicity.

4.9 Voice quality versus quantitative measures of the neoglottis

Results of the investigation of the relations between voice quality and quantitative measures were 'clear-cut', but surprising. The most important factor was the minimal distance during phonation: the closer the neoglottis the better the voice quality. This shows the relevance of closure for the sound production, something that is already well known for normal laryngeal voices. Surprisingly enough this measure was not used in earlier studies regarding esophageal speech. However, some quantitative measures were performed in some. They were always performed during phonation, and consisted of the length of the neoglottis (Isman and O'Brien, 1992), the prominence of the neoglottis (Damsté and Lerman, 1969), the dilatability of the esophagus below the neoglottis (Diedrich and Youngstrom, 1966), and the width of the hypopharynx (Vrticka and Svoboda, 1961). The only measure comparable to our study is the dilatability of the esophagus, which appeared to be of influence on the voice quality (Diedrich and Youngstrom, 1966). In the present study it was found that the ratio MAXPHON/MAXREST differed between the voice quality groups. The 'good' speakers showed a relatively larger increase in maximal subneoglottic distance during phonation compared to rest. Results of Diedrich and Youngstrom (1966) showed a relation between the acquisition of esophageal speech and the width of the esophagus, indicating that a wider subneoglottic distance was related to a better voice. They presume that a greater amount of air would be obtained within the lumen, providing a sufficient amount of air for voice production. Something, that may be important in esophageal speech, but not in tracheoesophageal speech. Presumably, in tracheoesophageal speech the increase in maximal subneoglottic distance is related to the tension of the neoglottic closure together with the flexibility of the tissues of the neck. It should be noted, in this respect, that a too high tension of the neoglottis (i.e. extreme hypertonicity or spasm) also leads to a relatively large increase in the maximal subneoglottic distance, while the voice is very poor or even absent.

The finding that the visual assessments and the quantitative measures are distinctive between the 'good' voice quality group and both other voice quality groups, but not between the 'bad' and the 'reasonable' voice quality group can be explained by the fact that the perceptual characteristics leading to a 'poor' or 'reasonable' voice quality are divers and can not be uniformly described by one specific characteristic.

5. Conclusions

In conclusion, this study gives a first impression about relationships between tracheoesophageal voice quality and anatomical and morphologic characteristics, as well as between tracheoesophageal voice quality and quantitative measures of the neoglottis as assessed on videofluoroscopic recordings. Some earlier studies mentioned that the vibratory segment in esophageal and tracheoesophageal voice is similar (Isman and O'Brien, 1992; Wetmore et al., 1985), they thereby more or less concluded that investigations of the vibratory segment in esophageal voice also hold true for tracheoesophageal voice. This study however, provides evidence that although the vibratory segment may look similar in both type of voice restoration, the resulting voice quality is not. The requirements for a good voice quality differ between both methods, since the difference in air supply causes differences between both methods of voice restoration. Optimal conditions of the neoglottis for voice production may even differ between the two types of speech. A 'good' esophageal voice does not necessarily correspond with a 'good' tracheoesophageal voice.

This study gives a first impression of the characteristics of the neoglottis related to voice quality in tracheoesophageal voice. More precise investigations of several perceptual aspects of voice quality separately will provide more information about relations between particular aspects of the neoglottis related to particular perceptual aspects of the voice quality. These relations are currently investigated and will be reported later.

Furthermore, the relations between the more subjective visual assessment and the more objective quantitative measures show that in future it might be possible to use the more standardized quantitative measures as a method to evaluate anatomical and morphologic characteristics of the neoglottis in relation to voice quality.

6. Acknowledgments

The authors wish to thank prof. dr. ir. L.C.W. Pols for his critical review of this manuscript. They also wish to thank the patients participating is this study. A.A.M.

Hart, MSc, is acknowledged for his help with the statistics. M.B. van Herk, PhD, is acknowledged for providing and adapting the program Drawer for the quantitative measures. M.F. Polak, B.E.G.M. Scholtens and A.B. Kamma, speech-language pathologists are acknowledged for their assistance during the videofluoroscopy recordings and their participation as experienced listeners.

The Maurits and Anna de Kock Stichting has financially supported the equipment needed for the speech recordings and the listening experiment.

7. References

Bentzen, N., Guld, A. and Rasmussen, H. (1976), "X-ray video-tape studies of laryngectomized patients", *Journal of Laryngology and Otology*, **90**: 655-666.

Brewer, D.W., Gould, L.V. and Casper, J. (1975), "Fiber-optic study of post-laryngectomized voice", *Laryngoscope*, **85**: 666-670.

Damsté, P.H. and Lerman, J.W. (1969), "Configuration of the neoglottis: An X-ray study", *Folia Phoniatrica*, **21**: 347-358.

Daou, R.A., Robillard Schultz, J., Remy, H., Turner Chan, N. and Attia, E.L. (1992), "Laryngectomee study: Clinical and radiologic correlates of esophageal voice", *Otolaryngology -- Head and Neck Surgery*, **92**: 628-634.

Debruyne, F., Delaere, P., Wouters, J. and Uwents, P. (1994), "Acoustic analysis of tracheo-oesophageal versus oesophageal speech", *Journal of Laryngology and Otology*, **108**: 325-328.

Diedrich, W.M. and Youngstrom, K.A. (1966). *Alaryngeal speech.* Charles C Thomas. Springfield, Illinois (USA).

Gates, G.A., Ryan, W., Cooper, J.C., Lawlis, F., Cantu, E., Hayashi, T., Lauder, E., Welch, R.W. and Hearne, E. (1982), "Current status of laryngectomee rehabilitation: I. Results of therapy", *Am J Otolaryngol*, **3**: 1-7.

Hilgers, F.J.M., Ackerstaff, A.H., Balm, A.J.M., Tan, I.B., Aaronson, N.K. and Persson, J.-O. (1997), "Development and clinical evaluation of a second-generation voice prosthesis (Provox®2), designed for anterograde and retrograde insertion", *Acta Otolaryngol (Stockh)*, **117**: 889-896.

Hilgers, F.J.M., Ackerstaff, A.H. and Van As, C.J. (1999), "Tracheoesophageal puncture: prosthetic voice management.", *Current Opinion in Otolaryngology &Head and Neck Surgery*, **7**: 112-118.

Hoffman, H.T., Fischer, H., VanDenmark, D., Peterson, K.L., McCulloch, T.M., Hynds Karnell, L. and Funk, G.F. (1997), "Botulinum neurotoxin injection after total laryngectomy", *Head & Neck*, **3**: 92-97.

Isman, K.A. and O'Brien, C.J. (1992), "Videofluoroscopy of the pharyngoesophageal segment during tracheoesophageal and esophageal speech", *Head & Neck*, **14**: 352-358.

Kirchner, J.A., Scatliff, J.H., Dey, F.L. and Shedd, D.P. (1963). "The pharynx after laryngectomy. Changes in its structure function", *Laryngoscope*, **73**: 18-33.

Lindsay, J.R., Morgan, R.H. and Wepman, J.M. (1944), "The cricopharyngeus muscle in esphageal speech", *The Laryngoscope*, **14**: 55-65.

Mahieu, H.F., Annyas, A.A., Schutte, H.K. and Van der Jagt, E.J. (1987), "Pharyngoesophageal myotomy for vocal rehabilitation of laryngectomees", *Laryngoscope*, **97**: 451-457.

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

McIvor, J., Evans, P.F., Perry, A. and Cheesman, D. (1990), "Radiological assessment of post laryngectomy speech", *Clinical Radiology*, **41**: 312-316.

Moolenaar-Bijl, A. (1951), "Some data on speech without a larynx", *Folia Phoniatrica*, **3**: 21-26.

Op de Coul, B.M.R., Hilgers, F.J.M., Balm, A.J.M., Tan, I.B., Van den Hoogen, F.J.A. and Van Tinteren, H. (submitted). "A decade of postlaryngectomy vocal rehabilitation in 318 patients: a single institutions experience with consistent application of indwelling voice prostheses (Provox)",

Richardson, J.L. (1981), "Surgical and radiological effects upon the development of speech after total laryngectomy", *Annals of Otology*, **90**: 294-297.

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

Robbins, J., Fisher, H.B., Blom, E.D. and Singer, M.I. (1984), "A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production", *Journal of Speech and Hearing Disorders*, **49**: 202-210.

Robbins, J., Fisher, H.B., Blom, E.D. and Singer, M.I. (1984), "Selected acoustic features of tracheoesophageal, esophageal, and laryngeal speech", *Archives of Otolaryngology*, **110**: 670-672.

Robe, E.Y., Moore, P., Andrews, A.H. and Holinger, P.H. (1956), "A study of the role of certain factors in the development of speech after laryngectomy. Part 2: Site of pseudoglottis", *Laryngoscope*, **66**: 382-401.

Singer, M.I. and Blom, E.D. (1980),"An endoscopic technique for restoration of voice after laryngectomy.", *Ann Otol Rhinol Laryngol*, **89**: 529-533.

Singer, M.I. and Blom, E.D. (1981). "Selective myotomy for voice restoration after total laryngectomy", *Archives of Otolaryngology*, **107**: 670-673.

Singer, M.I., Blom, E.D. and Hamaker, R.C. (1986), "Pharyngeal plexus neurectomy for alaryngeal speech rehabilitation", *Laryngoscope*, **96**: 50-53.

Sloane, P.M., Griffin, J.M. and O'Dwyer, T.P. (1991), "Esophageal insufflation and videofluoroscopy for evaluation of esophageal speech in laryngectomy patients: Clinical Implications", *Radiology*, **181**: 433-437.

Smith, J.K., Rise, E.N. and Gralnek, D.E. (1966), "Speech recovery in laryngectomized patients", *Laryngoscope*, **76**: 1540-1546.

Van As, C.J., Hilgers, F.J.M., Verdonck-de Leeuw, I.M. and Koopmans-van Beinum, F.J. (1998), "Acoustical analysis and perceptual evaluation of tracheoesophageal prosthetic voice", *Journal of Voice*, **12**: 239-248.

Van As, C.J., Tigges, M., Wittenberg, T., Op de Coul, B.M.R., Eysholdt, U. and Hilgers, F.J.M. (1999),"High-speed digital imaging of neoglottic vibration after total laryngectomy", *Archives of Otolaryngology Head and Neck Surgery*, in press.

Van Weissenbruch, R. (1996), "Voice restoration after total laryngectomy", *University of Groningen,*

Wetmore, S.J., Ryan, S.P., Montague, J.C., Krueger, K., Wesson, K., Tirman, R. and Diner, W. (1985), "Location of the vibratory segment in tracheoesophageal speakers", *Otolaryngology -- Head and Neck Surgery*, **93**: 355-360.

Williams, S.E. and Watson, B. (1987), "Speaking proficiency variations according to method of alaryngeal voicing", *Laryngoscope*, **97**: 737-739.

Vrticka, K. and Svoboda, M. (1961), "A clinical and X-ray study of 100 laryngectomized speakers", *Folia Phoniatrica*, **13**: 174-186.