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 9

Digital high-speed imaging related to perceptual evaluations, acoustic analyses, videofluoroscopy, and clinical factors
ABSTRACT

The aim of the present chapter is to give an overview of the relations between the anatomical and morphologic characteristics of the neoglottis studied by means of visual assessments of digital high-speed imaging recordings (described in Chapter 8) at the one hand, and perceptual evaluations (described in Chapter 4) and acoustic analyses (described in Chapter 5) of voice quality at the other hand. Also relations between the anatomical and morphologic characteristics of the neoglottis studied with digital high-speed imaging and these characteristics studied with videofluoroscopy (described in Chapter 6) are investigated and described.

To the best of our knowledge the study presented in Chapter 8 was the first study using digital high-speed imaging of neoglottic characteristics. It is mainly a methodological and descriptive paper on the use of digital high-speed imaging and the anatomical and morphologic characteristics of the neoglottis that can be observed with this novel endoscopic imaging technique. No relations with other parts of the study were investigated at that time. However, there is evidently a need for placing these visual assessments of the digital high-speed imaging recordings in a broader perspective and for relating them to tracheoesophageal voice quality, and other, more common, methods of evaluation of anatomical and morphologic characteristics of the neoglottis, like videofluoroscopy.

Results show relations between the visual assessments of the high-speed recordings and the perceptual evaluations and acoustic analyses of voice quality. Especially, the amount of saliva, the visibility of the origin of the neoglottis, the shape of the neoglottis and the regularity of the vibration are of value in this respect. Assessments of digital high-speed imaging provide information complementary to the videofluoroscopy recordings and only show overlap with the assessments of videofluoroscopy regarding regurgitation of barium/amount of saliva and presence of a neoglottic bar/visibility of the origin of the neoglottis. Additionally, some relations between the visual assessments of the high-speed recordings and the sociodemographic and clinical factors are found.

It can be concluded that digital high-speed imaging provides additional information about anatomical and neoglottic characteristics that cannot be obtained with videofluoroscopy. Also, in relation to voice quality (perceptual evaluation and acoustic analyses) and sociodemographic and clinical factors, relevant information can be obtained from digital high-speed imaging.
9.1 INTRODUCTION

Digital high-speed imaging is at present only in use as a research tool, although developments for clinical implementation are promising. With digital high-speed imaging it is possible, in analogy to stroboscopy, to obtain images of the vibrating neoglottis from above (birds-eye view). These recordings are expected to give valuable information in addition to the information obtained by videofluoroscopy, which provides a lateral view. The results described in Chapter 8 indicate that digital high-speed imaging is a useful tool to study the vibratory behaviour of the neoglottis. Results showed a large variability in shapes and locations of the vibration. Several anatomical and morphologic characteristics of the neoglottis were assessed and described. Prior to the published paper that is described in chapter 8, no studies have been performed using digital high-speed imaging to study characteristics of the neoglottis. Also, nothing is known yet about the relations between anatomical and morphologic characteristics of the neoglottis as studied by means of digital high-speed imaging and tracheoesophageal voice quality assessed with perceptual evaluations and acoustic analyses. Larsson et al. (2000) recently investigated high-speed imaging, kymography and acoustic analysis for diplophonic voices and tremor using an analysis system developed by themselves: the high-speed toolbox. With the high-speed toolbox glottal edge detections and kymograms were made of the high-speed recordings. Relations between glottal vibratory patterns and the sound waveform were found. Lundström and Hammarberg (1999) used the same high-speed toolbox to study the voiced-voiceless distinction in one esophageal and one tracheoesophageal speaker. They conclude that the high-speed recordings together with simultaneously recorded voice signal give answers to important questions about the function of the pharyngoesophageal (PE) segment.

In view of the fact that the use of digital high-speed imaging is at an initial stage, the need for investigation of relations between observations made with this new method and voice quality is evident. Also, comparison between observations of the neoglottis from above with this new method and the more widely, already since decades used method of videofluoroscopy providing a lateral view of the neoglottis is important.

In the present study relations between the anatomical and morphologic characteristics as studied by digital high-speed imaging are related to perceptual evaluation (described in Chapter 4) and acoustic analyses (described in Chapter 5) of voice quality. Also, the anatomical and morphologic characteristics of the neoglottis seen with digital high-speed imaging are related to those seen with videofluoroscopy (described in Chapter 6). Furthermore, the relations between the anatomical and morphologic characteristics of the neoglottis as studied by means of digital high-speed imaging and sociodemographic and clinical factors (see Chapter 3) are described.

9.2 PATIENTS AND METHODS

9.2.1 PATIENTS

The patient group studied by means of digital high-speed imaging consisted of 46 patients; recordings could be obtained from 44 out of these 46 patients. In one patient no recordings could be obtained due to a high gag reflex (patient after standard total laryngectomy), and in one patient no useful recordings could be obtained due to the dispersion of the light in a wide neoglottis (patient with total laryngectomy and full gastric pull-up reconstruction). In the group of the remaining 44 patients, 5 patients who were shortly after surgery were included. These 5 patients did not participate in any other part of the study, leaving a group of 39 patients to be studied in relation to the other parts of the study. For investigation of relations
with the perceptual evaluations 38 patients could be evaluated (one patient died before speech recordings could be obtained), for the relations with acoustic analyses 37 (one patient refused to produce the sustained /a/’s necessary for acoustic analyses) and for the relations with videofluoroscopy 37 patients could be evaluated (two patients refused videofluoroscopy). For the relations with the sociodemographic and clinical factors all 44 patients of whom digital high-speed recordings are obtained could be evaluated. In Chapter 3, concerning the patient group participating in this study) detailed information is given on the patients and the parts of the study in which they participated.

9.2.2 Methods

The perceptual evaluations consisted of an overall judgment of voice quality (good, reasonable, poor) and judgment of 19 bipolar semantic 7-point scales by 20 naive listeners and 4 trained expert listeners. For the investigation of the relations presented in this chapter only the judgments of the trained expert raters were used since they tend to listen more analytically; the primary goal of these judgments was to use them for evaluation of other methods of investigation, while the results of the naive listeners were collected to gain insight in communicative aspects. The perceptual scales and the methods used are described in detail in Chapter 4, the perceptual scales judged by the trained expert listeners are: deviant-normal, unpleasant-pleasant, ugly-beautiful, noise-no noise, monotonous-melodious, expressionless-expressionive, weak-powerful, unsteady-steady, jerking-fluid, slow-quick, low-high, deep-shrill, bubbly-not bubbly, breathy-not breathy, hypertonic-not hypertonic, hypotonic-not hypotonic, and unintelligible-intelligible. The scales are judged on a 7-point scale in which 1 represents the negative scale end and 7 the positive scale end.

The acoustic analyses consisted of acoustic signal typing and acoustic measures of periodicity and harmonicity. The analyses performed and the methods used are described in detail in Chapter 5. The acoustic signal types were: Type I (stable and harmonic); Type II (stable and at least one harmonic); Type III (unstable or partly harmonic); and Type IV (barely harmonic). The acoustic measures collected were: median fundamental frequency (F0-median), standard deviation of fundamental frequency (F0-SD), jitter, percentage of voiced (%Voiced), harmonics-to-noise ratio (HNR), glottal-to-noise ratio (GNE) and band energy difference (BED). In addition to the objective acoustic measures, also maximum phonation time (MPT) was used.

The evaluation of the videofluoroscopy recordings was performed using a newly developed evaluation protocol, consisting of visual assessment of characteristics of the neoglottis and quantitative measures of the neoglottis. The characteristics that were assessed and the measures that were performed, as well as the methods used are described in detail in Chapter 6. The visual assessments, mostly using clear dichotomies (yes/no), consisted of the following parameters: appearance of a neoglottic bar at rest and appearance of a neoglottic bar during phonation, stasis of barium above the neoglottis during phonation, regurgitation of barium during phonation, flattening of the neoglottic bar when swallowing, tonicity of the neoglottis during phonation, and the cervical level of the neoglottis at rest and during phonation. The quantitative measures of the neoglottis were: minimal (neoglottic) distance at rest and during phonation (MINREST and MINPHON), maximal (subneoglottic) distance at rest and during phonation (MAXREST and MAXPHON), surface area of the neoglottic bar at rest and during phonation (SUREST and SURPHON), prominence of the neoglottic bar at rest and during phonation (PROMREST and PROMPHON), and an index of the increase of the maximal (subneoglottic) distance from rest to phonation (INDEX MAXPHON/MAXREST).

The visual assessments of the digital high-speed recordings were performed using an evaluation protocol consisting of three assessments of the overall quality of the high-speed
recording (assessability, brightness, and focus) and seven assessments of anatomical and morphologic characteristics of the neoglottis (amount of saliva, visibility of the origin of the neoglottis, shape of the neoglottis, location of the visible vibration, presence of a mucosal wave, regularity of the vibration, and closure phase). The evaluation form and the results of the visual assessments are described in Chapter 8. Only the seven assessments of the neoglottic characteristics are included in further analyses.

The sociodemographic and clinical factors studied were: sex, age, postoperative follow-up, reconstruction, myotomy, neurectomy, radical neck dissection, and radiotherapy. More information about these parameters and the subgroups used for statistical analyses is given in Chapter 3.

### 9.2.3 STATISTICAL ANALYSES

Statistical analyses were performed using the Statistical Package for Social Sciences, version 10.0 (SPSS Inc. Chicago, Ill.). A p-value below .05 is considered significant.

The results of the visual assessments of the digital high-speed recordings are related to the results of videofluoroscopy (visual assessment and quantitative measures), to the results of perceptual evaluations (overall voice quality judgment and perceptual scale judgments), to the results of acoustic analyses (acoustic signal typing and acoustic measures), and to the sociodemographic and clinical factors.

First, for the statistical analyses subgroups were formed, based on the visual assessments of the high-speed recordings of the neoglottis. The numbers of patients in each of the subgroups mentioned below, are based on the 44 patients of whom high-speed recordings were obtained. Not all characteristics were assessed for all 44 patients, due to the fact that not all parameters were assessable for all patients. The presence of a mucosal wave and the regularity of the vibration were assessed in 44 patients, the location of the visible vibration in 41 patients, the shape of the neoglottis in 36 patients and the closure phase in 30 patients. For saliva the subgroups formed are ‘none or a little’ (n=20), ‘moderate’ (n=15) and ‘much or obstructing’ (n=9); for visibility of the origin of the neoglottis the subgroups formed are ‘yes’ (n=31) and ‘no’ (n=13); for shape of the neoglottis the subgroups formed are ‘circular’ (n=7), ‘split side-to-side (coronal)’ (n=14), ‘split anterior-posterior (sagital)’ (n=5) and ‘irregular’ (n=12), the shape ‘triangular’ was excluded from statistical analysis, since it contained only one patient; for location of the visible vibration the subgroups formed are ‘back wall’ (n=3), ‘front wall’ (n=5), ‘one lateral wall’ (n=3), ‘both lateral walls’ (n=3), ‘anterior and posterior wall’ (n=2), and ‘all walls’ (n=15), in one patient a combination of ‘posterior and one lateral wall’ was seen and in one patient a combination of ‘anterior and both lateral walls’ was seen (not included in statistical analysis), in two patients a combination of ‘posterior and both lateral walls’ was seen (only participated in high-speed part). For the presence of a mucosal wave the subgroups formed are ‘strong or weak’ (n=22) and ‘absent’ (n=22); for the regularity of the vibration the subgroups formed are ‘regular’ (n=12) and ‘irregular’ (n=31); and for the closure phase the subgroups are ‘open’ (n=23) and ‘equal’ (n=7), the subgroup ‘closed’ was excluded from statistical analysis, since it contained only one patient.

Then, the following statistical tests were performed to investigate the relationships:

- For investigation of the relations between digital high-speed imaging and the overall judgment of voice quality, an exact chi-squared test was used for the nominal data, and a chi-squared test for linear-by-linear association was used for the ordinal data.
- Relations between digital high-speed imaging and the perceptual scale judgments were investigated using a t-test for two independent samples for comparison of two subgroups, and using analyses of variance followed by post hoc Tukey tests for comparison of three or more subgroups.
The relations between digital high-speed imaging and acoustic signal typing were investigated using exact chi-squared tests for the nominal data or chi-squared tests for linear-by-linear association for the ordinal data.

Relations between digital high-speed imaging and the acoustic measures and maximum phonation time were investigated using a t-test for two independent samples or analyses of variance followed by post hoc Tukey tests. For the acoustic parameter percentage of voiced, non-parametric tests were used. The acoustic measures median fundamental frequency and standard deviation of fundamental frequency were logarithmically transformed for perceptual relevance; values in text are transformed back into Hz.

The relations between digital high-speed imaging and the visual assessments of videofluoroscopy were investigated with exact chi-squared tests for the nominal data or chi-squared tests for linear-by-linear association for the ordinal data.

For investigation of relations between digital high-speed imaging and the quantitative measures of videofluoroscopy, t-tests for two independent samples, or analyses of variance followed by post hoc Tukey tests were used. The increase of the maximal distance from rest to phonation (index) and the maximal distance at rest needed logarithmic transformation first.

The relations between the visual assessments of digital high-speed imaging and the sociodemographic and clinical factors were investigated with exact chi-squared tests for the nominal data and chi-squared tests for linear-by-linear association for the ordinal data.

In the results presented in the present chapter, only the statistically significant results are given.

9.3 RESULTS

In this section the following results are described:

- digital high-speed imaging versus perceptual evaluations (section 9.3.1)
- digital high-speed imaging versus acoustic analyses (section 9.3.2)
- digital high-speed imaging versus videofluoroscopy (section 9.3.3)
- digital high-speed imaging versus clinical factors (section 9.3.4)

9.3.1 DIGITAL HIGH-SPEED IMAGING VERSUS PERCEPTUAL EVALUATIONS

The relations between the results of digital high-speed imaging and the perceptual evaluations are divided into two different parts, first in section 9.3.1.1 the relations with the overall voice quality judgment as good, reasonable or poor are described, and then, in section 9.3.1.2 the relations with the perceptual evaluations of the 19 semantic bipolar 7-point scales are described.

9.3.1.1 Digital high-speed imaging versus overall voice quality

The relations between the visual assessments of the digital high-speed recordings and the overall judgment of voice quality (good, reasonable, poor) were investigated by means of chi-squared tests. It appeared that overall voice quality was related to the visibility of the origin of the neoglottis (p=.024). In Table 9.1 this relation is shown.
Table 9.1. Relation between the overall judgment of voice quality (column 1) and the visibility of the neoglottis (p=.024). Numbers are number of patients, in total n=38.

<table>
<thead>
<tr>
<th>Overall voice quality</th>
<th>Visibility of the origin of the neoglottis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Good</td>
<td>12</td>
</tr>
<tr>
<td>Reasonable</td>
<td>11</td>
</tr>
<tr>
<td>Poor</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>no</td>
</tr>
<tr>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>6</td>
</tr>
</tbody>
</table>

As can be seen in Table 9.1, when the neoglottis is visible, the speaker is more often judged to have good or reasonable voice quality, while the voice quality is more often poor when the origin of the neoglottis was not seen during digital high-speed imaging. For the videofluoroscopy recordings it has already been shown that the presence of a neoglottic bar is important for good voice quality (see Chapter 7) and this result is thus confirmed here.

No relations with the overall judgment of voice quality were found for the amount of saliva, the shape of the neoglottis, the location of the visible vibration, the presence of a mucosal wave, the regularity of the vibration, and the closure phase. A trend was noted for the amount of saliva to be less in the good speaker group (p=.058).

9.3.1.2 Digital high-speed imaging versus perceptual scale judgments

For investigation of the relations between the visual assessments of the digital high-speed recordings and the perceptual scale judgments of the 17 bipolar semantic 7 point scales (1=most negative scale end, 7=most positive scale end), t-tests for two independent samples were used for the visual assessments with two subgroups, and analyses of variance were used for comparison of three or more subgroups. Results showed relations between the visual assessments of the amount of saliva, the visibility of the origin of the neoglottis, the shape of the neoglottis, and the regularity of the vibration in the digital high-speed recordings and some of the perceptual scales. In Table 9.2 to Table 9.5 the results for these assessments are shown; only the perceptual scales that showed significant relations are shown. Regarding the amount of saliva analyses of variance showed that the perceptual scales deviant-normal, monotonous-melodious, expressionless-expressive, and hypotonic-not hypotonic were related to the subgroups. In Table 9.2 these relations are shown.

The results in Table 9.2 show that when less saliva is seen during high-speed imaging the more 'normal', melodious and expressive the voice sounds. In patients where a large amount of saliva is seen, or where it is even obstructing the neoglottis, the voice sounds more hypotonic. Apparently, saliva interfering with the neoglottic vibration causes a decreased voice quality. Something that is not surprising since the saliva disturbs the voice. Moreover, as already seen with regurgitation during videofluoroscopy (see Chapter 7), the amount of saliva also seems to be related to the tonicity of the neoglottis. When there is no good neoglottic closure the saliva can be driven upwards without any obstruction and interferes with voice production. In a number of cases the saliva itself might even by the major sound-producing element.
Table 9.2. Results of one-way analysis of variance for the amount of saliva in relation to the perceptual scale judgments. In column 1 the perceptual scale is given, in column 2 the p-value of the one-way analysis of variance, and in column 3, 4 and 5 the average values of the perceptual scale judgments for that subgroup. The arrows between the values indicate that there is a significant difference between those subgroups; the box attached to the arrow gives the exact p-value.

<table>
<thead>
<tr>
<th>Perceptual scale</th>
<th>p-value</th>
<th>Amount of saliva</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>none/a little</td>
</tr>
<tr>
<td></td>
<td></td>
<td>moderate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>much/obstructing</td>
</tr>
<tr>
<td>deviant-normal</td>
<td>.027</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.8</td>
</tr>
<tr>
<td>monotonous-melodious</td>
<td>.006</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.5</td>
</tr>
<tr>
<td>expressionless-expressive</td>
<td>&lt;.001</td>
<td>5.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.1</td>
</tr>
<tr>
<td>hypotonic-not hypotonic</td>
<td>.046</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.3</td>
</tr>
</tbody>
</table>

Regarding the visibility of the origin of the neoglottis relations were found with the perceptual scales deviant-normal, monotonous-melodious, expressionless-expressive, and bubbly-not bubbly. These results are shown in Table 9.3.

Table 9.3. Results of t-tests for independent samples for the visibility of the origin of the neoglottis in relation to the perceptual scale judgments. In column 1 the perceptual scale is given, in column 2 the p-value, and in column 3 and 4 the mean values of the judgments of the perceptual scale for that subgroup.

<table>
<thead>
<tr>
<th>Perceptual scale</th>
<th>p-value</th>
<th>Visibility of the neoglottis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>no</td>
</tr>
<tr>
<td>deviant-normal</td>
<td>.019</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.9</td>
</tr>
<tr>
<td>monotonous-melodious</td>
<td>.011</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.3</td>
</tr>
<tr>
<td>expressionless-expressive</td>
<td>.030</td>
<td>5.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4.2</td>
</tr>
<tr>
<td>bubbly-not bubbly</td>
<td>.022</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.2</td>
</tr>
</tbody>
</table>

When the origin of the neoglottis is visible during high-speed imaging, the voice sounds less deviant, more melodious and expressive and less bubbly. The visibility of the origin of the neoglottis as seen from above is thus a good indicator for voice quality. The relation with the perceptual scale bubbly-not bubbly can be explained by the fact that in the cases where the neoglottis was not visible this may also be due to the fact that it is obstructed by saliva, and thus that the saliva interferes with voice quality. The relation with the perceptual scales monotonous-melodious and expressionless-expressive indicates that when the neoglottis was not visible, the possibilities for pitch inflection are decreased.

Analyses of variance show that the shape of the neoglottis is related to the scales expressionless-expressive, low-high, deep-shrill and intelligible-not intelligible. Post hoc Tukey tests show that the group with an irregular neoglottis was rated higher and shriller than the circular shapes. For the perceptual scale expressionless-expressive post hoc Tukey tests revealed no significant differences between the subgroups. The results are given in Table 9.4.
High-speed imaging, perceptual evaluations, acoustic analyses and videofluoroscopy

Table 9.4. Results of one-way analysis of variance for the shape of the neoglottis in relation to the perceptual scale judgments. In column 1 the perceptual scale is given, in column 2 the p-value of the one-way analysis of variance, and in column 3 to 6 the average values of the perceptual scale judgments for that subgroup. The arrows between the values indicate that there is a significant difference between those subgroups; the box attached to the arrow gives the exact p-value.

<table>
<thead>
<tr>
<th>Perceptual scale</th>
<th>p-value</th>
<th>Shape of the neoglottis</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>circular</td>
<td>split side-to-side</td>
<td>split anterior posterior</td>
<td>irregular</td>
<td></td>
</tr>
<tr>
<td>expressionless-expressive</td>
<td>.031</td>
<td>4.1</td>
<td>5.4</td>
<td>4.3</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>low-high</td>
<td>.026</td>
<td>3.4</td>
<td>4.0</td>
<td>3.9</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>deep-shrill</td>
<td>.049</td>
<td>3.3</td>
<td>3.8</td>
<td>3.6</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>unintelligible-intelligible</td>
<td>.034</td>
<td>5.1</td>
<td>6.2</td>
<td>5.4</td>
<td>6.1</td>
<td></td>
</tr>
</tbody>
</table>

The differences in the perceptual scales between the irregular shapes and the circular shapes are difficult to explain. In normal vocal folds, irregularity leads to a decrease in voice quality, whereas in tracheoesophageal voice this is apparently not the case. At present no explanation can be given for this, and more research with digital high-speed imaging is needed to give more insight in this phenomenon. In addition to this, the intelligibility was found to be better in the side-to-side split shapes than in the circular shapes. Also this phenomenon is difficult to explain, although it might be the case that in the side-to-side split shapes in which the vibration is merely expected to be situated at the back wall, the volitional control over the neoglottis might be better than in the circular shapes in which most likely more walls are vibrating.

The regularity of the vibration shows relations with the scales unsteady-steady, low-high, and deep-shrill. In Table 9.5 these results are shown.

Table 9.5. Results of Wests for independent samples for the regularity of the vibration of the neoglottis in relation to the perceptual scale judgments. In column 1 the perceptual scale is given, in column 2 the p-value, and in column 3 and 4 the mean values of the judgments of the perceptual scale for that subgroup.

<table>
<thead>
<tr>
<th>Perceptual scale</th>
<th>p-value</th>
<th>Regularity of vibration</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>irregular</td>
<td>regular</td>
<td></td>
</tr>
<tr>
<td>unsteady-steady</td>
<td>.047</td>
<td>4.9</td>
<td>5.7</td>
<td></td>
</tr>
<tr>
<td>low-high</td>
<td>.025</td>
<td>3.8</td>
<td>4.7</td>
<td></td>
</tr>
<tr>
<td>deep-shrill</td>
<td>.018</td>
<td>3.7</td>
<td>4.4</td>
<td></td>
</tr>
</tbody>
</table>

The subgroup with a regular vibration was rated steadier, higher and shriller compared to the subgroup with irregular vibration. Apparently, the irregularity seen in the high-speed recordings is reflected in a less steady voice. Moreover, the relations with the perceptual scales regarding pitch indicate that the regular vibration results in a perceptually higher pitch. This might also be related to a better periodicity and harmonicity and less noise in case of regular vibration.

The assessments of location of the visible vibration, presence of a mucosal wave and closure phase are not related to any of the perceptual scales.
9.3.2 Digital high-speed imaging versus acoustic analyses and maximum phonation time

In section 9.3.2.1 relations between the results of digital high-speed imaging and acoustic signal typing are described, and in section 9.3.2.2 the relations between the results of digital high-speed imaging and the acoustic measures (including maximum phonation time) described.

9.3.2.1 Digital high-speed imaging versus acoustic signal typing

Relations between the visual assessments of the digital high-speed imaging recordings and the four acoustic signal types were investigated by means of chi-squared tests. A relation was found between the four signal types and the amount of saliva \((p=.007)\) and between the four signal types and the visibility of origin of the neoglottis \((p<.001)\).

In the better signal types, less saliva is seen. In Table 9.6 the relation between the subgroups for the amount of saliva and the four signal types is shown.

Table 9.6. Relation between the amount of saliva seen with digital high-speed imaging (columns) and the four acoustic signal types (rows). Numbers are number of patients, in total \(n=37\).

<table>
<thead>
<tr>
<th>Acoustic signal type</th>
<th>Amount of saliva</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>none/a little</td>
</tr>
<tr>
<td>I (stable &amp; harmonic)</td>
<td>6</td>
</tr>
<tr>
<td>II (stable &amp; at least one harmonic)</td>
<td>7</td>
</tr>
<tr>
<td>III (unstable or partly harmonic)</td>
<td>5</td>
</tr>
<tr>
<td>IV (barely harmonic)</td>
<td>2</td>
</tr>
</tbody>
</table>

In type-I signals, mostly none or a little saliva is seen, a large or obstructing amount of saliva is only seen in the type-III or type-IV signals. This indicates that the amount of saliva, and thereby its interference with the voice production at the level of the neoglottis influences the voice sound in a negative way, the less saliva, the better the acoustic characteristics of the voice quality.

Also, in the better signal types more often the origin of the neoglottis was seen. In Table 9.7 the relation between the visibility of the origin of the neoglottis and the four acoustic signal types is shown.

Table 9.7. Relation between the visibility of the origin of the neoglottis assessed by digital high-speed imaging (columns) and the four acoustic signal types (rows). Numbers are number of patients, in total \(n=37\).

<table>
<thead>
<tr>
<th>Acoustic signal type</th>
<th>Visibility of the origin of the neoglottis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>I (stable &amp; harmonic)</td>
<td>7</td>
</tr>
<tr>
<td>II (stable &amp; at least one harmonic)</td>
<td>12</td>
</tr>
<tr>
<td>III (unstable or partly harmonic)</td>
<td>6</td>
</tr>
<tr>
<td>IV (barely harmonic)</td>
<td>2</td>
</tr>
</tbody>
</table>

When the origin of the neoglottis is visible, type-I, type-II, and type-III signals are seen more often, while in the cases when the origin of the neoglottis was not visible, type-III and type-IV signals are seen more often. This indicates that indeed when the origin of the neoglottis is seen from above, more harmonicity is present in the voice, most probably due to the neoglottic closure that should be present when a neoglottis is seen.

The visual assessments of the shape of the neoglottis, the location of the visible vibration, the presence of a mucosal wave, the regularity of the vibration, and the closure phase were not related to the acoustic signal types.
9.3.2.2 Digital high-speed imaging versus acoustic measures and maximum phonation time

T-tests for two independent samples or analyses of variance (three or more independent samples) showed relations between the visual assessments of the amount of saliva, the shape of the neoglottis, the regularity of the vibration, and the location of the visible vibration in the digital high-speed imaging recordings and some of the acoustic measures. In the Tables 9.8 to 9.11 the results for those four visual assessments are shown.

Regarding the amount of saliva, relations were found with the band energy difference (BED) \((p=.039)\) and percentage of voiced \((\%Voiced)(p=.035)\). These relations are shown in Table 9.8.

Table 9.8. Results of one-way analysis of variance for the amount of saliva in relation to two acoustic measures. In column 1 the acoustic measure is given, in column 2 the p-value of the one-way analysis of variance, and in column 3 to 5 the average values of the acoustic measure for that subgroup. The arrows between the values indicate that there is a significant difference between those subgroups; the box attached to the arrow gives the exact p-value.

<table>
<thead>
<tr>
<th>Acoustic measure</th>
<th>p-value</th>
<th>Amount of saliva</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>none/a little</td>
</tr>
<tr>
<td><strong>BED (dB)</strong></td>
<td>.039</td>
<td>-22.17</td>
</tr>
<tr>
<td><strong>%Voiced</strong></td>
<td>.035</td>
<td>73.9</td>
</tr>
</tbody>
</table>

The band energy difference is largest in the subgroup in which no or only a little saliva is seen, indicating that the high-frequency noise in the long-term average spectrum is lowest in that subgroup. With an increasing amount of saliva, the band energy difference becomes smaller, indicating a larger amount of high-frequency noise. The average percentage of voiced is considerably lower in the patient group in which much or an obstructing amount of saliva is seen. This shows again the negative influence of the saliva on the acoustic characteristics of voice quality, something that was also seen in relation to the perceptual evaluations in section 9.3.1.

Table 9.9. Results of one-way analysis of variance for the shape of the neoglottis in relation to two acoustic measures and maximum phonation time. In column 1 the acoustic measure is given, in column 2 the p-value of the one-way analysis of variance, and in column 3 to 6 the average values of the acoustic measure for the subgroup. The arrows between the values indicate that there is a significant difference between those subgroups; the box attached to the arrow gives the exact p-value.

<table>
<thead>
<tr>
<th>Acoustic measure</th>
<th>p-value</th>
<th>Shape of the neoglottis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>circular</td>
</tr>
<tr>
<td><strong>F0-median (Hz)</strong></td>
<td>.013</td>
<td>78</td>
</tr>
<tr>
<td><strong>HNR (dB)</strong></td>
<td>.038</td>
<td>5.4</td>
</tr>
<tr>
<td><strong>MPT (s)</strong></td>
<td>.034</td>
<td>7.4</td>
</tr>
</tbody>
</table>
Regarding the shape of the neoglottis, relations were found with the acoustic measures median fundamental frequency (F0-median) ($p=.013$) and harmonics-to-noise ratio (HNR) ($p=.038$), and also with maximum phonation time (MPT) ($p=.034$). In Table 9.9 the results are shown.

The median fundamental frequency is found to be higher for the irregular shapes compared to the circular shapes and the anterior-posterior split shapes. This is in concordance with the results of the perceptual evaluations in section 9.3.1 and warrants further study in future. The average harmonics-to-noise ratio is higher (thus better) for the irregular shapes compared to the side-to-side split shapes. This type of relation was not found for the perceptual evaluations. These results suggest, however, that in the subgroup with an irregular shaped neoglottis the voice contains more noise, which might, in analogy to normal vocal folds, be caused by the irregularity of the neoglottis. The maximum phonation time is higher for the side-to-side (coronal) split shapes compared to the circular shapes, suggesting a better neoglottic closure in the side-to-side split subgroup.

Regarding the regularity of the vibration, a relation was found with the harmonics-to-noise ratio (HNR) ($p=.004$) and the percentage of voiced (%Voiced) ($p=.011$). The harmonics-to-noise ratio and percentage of voiced are higher for the group showing regular vibration than for the group showing irregular vibration. The results are shown in Table 9.10.

Table 9.10. Results of the t-test for independent samples for the regularity of the vibration of the neoglottis in relation to two acoustic measures. In column 1 the acoustic measure is given, in column 2 the p-value and in column 3 and 4 the mean values of the acoustic measures for that subgroup.

<table>
<thead>
<tr>
<th>Acoustic measure</th>
<th>p-value</th>
<th>HNR (dB)</th>
<th>%Voiced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regularity of vibration</td>
<td></td>
<td>irregular</td>
<td>regular</td>
</tr>
<tr>
<td>HNR (dB)</td>
<td></td>
<td>3.2</td>
<td>7.8</td>
</tr>
<tr>
<td>%Voiced</td>
<td>.004</td>
<td>44.3</td>
<td>75.9</td>
</tr>
</tbody>
</table>

The harmonics-to-noise ratio is higher in the subgroup where a regular vibrating neoglottis is seen, indicating that the relative amount of noise is lower in that subgroup. Also, the percentage of voiced is higher in the subgroup with regular vibration, a parameter that also points to a better voice quality (see Chapter 5).

For the location of the vibration analyses of variance showed a relation with the harmonics-to-noise ratio (HNR) ($p=.042$). Due to small subgroups no post hoc comparisons could be performed. The harmonics-to-noise ratio was highest in the group in which the vibration was located at the posterior esophageal wall. The results are shown in Table 9.11.

Table 9.11. Results of one-way analysis of variance for the location of the visible vibration of the neoglottis in relation to the harmonics-to-noise ratio. In column 2 the p-value of the one-way analysis of variance is given, and in column 3 to 8 the mean values of the harmonics-to-noise ratio for that subgroup.

<table>
<thead>
<tr>
<th>Acoustic measure</th>
<th>p-value</th>
<th>Location of the visible vibration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1-lateral wall</td>
</tr>
<tr>
<td>HNR (dB)</td>
<td>.042</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Due to small subgroups no post hoc tests could be performed. In addition to the subgroups mentioned, one patient was seen with vibration at the posterior and one lateral wall (HNR -0.05 dB) and one patient was seen with vibration at the anterior wall and both lateral walls (HNR 15.7 dB).

Regarding visibility of the origin of the neoglottis, the presence of a mucosal wave, and closure phase, no significant relations were found.
9.3.3 Digital High-Speed Imaging Versus Videofluoroscopy

In section 9.3.3.1 relations between the visual assessment of the videofluoroscopy recordings and the results of the assessment of the digital high-speed imaging recordings are described and in section 9.3.3.2 the relations between the quantitative measures of the videofluoroscopy recordings and the results of the digital high-speed imaging are described.

9.3.3.1 Digital high-speed imaging versus visual assessment of videofluoroscopy

The relations between the visual assessments of the digital high-speed imaging recordings and the visual assessments of the videofluoroscopy recordings were investigated by means of chi-squared tests. Results showed relations between the amount of saliva and the visibility of the origin of the neoglottis assessed in the digital high-speed (top view) and some of the anatomical and morphologic characteristics of the neoglottis as assessed on videofluoroscopy recordings (lateral view).

Table 9.12. Cross table of the amount of saliva seen with digital high-speed imaging (columns) and the presence of a neoglottic bar at rest, regurgitation of barium during phonation and tonicity of the neoglottis during phonation judged with videofluoroscopy (rows).

<table>
<thead>
<tr>
<th>Videofluoroscopy assessment</th>
<th>Amount of saliva</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>none/a little</td>
</tr>
<tr>
<td>Neoglottic bar at rest</td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>4</td>
</tr>
<tr>
<td>1 or 2</td>
<td>15</td>
</tr>
<tr>
<td>Regurgitation</td>
<td></td>
</tr>
<tr>
<td>yes</td>
<td>3</td>
</tr>
<tr>
<td>no</td>
<td>16</td>
</tr>
<tr>
<td>Tonicity</td>
<td></td>
</tr>
<tr>
<td>hypotonic</td>
<td>3</td>
</tr>
<tr>
<td>&quot;normotonic&quot;</td>
<td>10</td>
</tr>
<tr>
<td>hypertonic</td>
<td>6</td>
</tr>
</tbody>
</table>

The amount of saliva seen during digital high-speed imaging was related to the appearance of a neoglottic bar at rest (p=.027), to regurgitation of barium during phonation (p=.005) and to the tonicity of the neoglottis during phonation (p=.004). In Table 9.12 the relation between the amount of saliva and the presence of a neoglottic bar at rest, regurgitation of barium and tonicity of the neoglottis is shown.

In the patient group with a visible neoglottic bar at rest, there was more often no or a little saliva seen, while in the patient group without a visible neoglottic bar at rest there was more often a large to obstructing amount of saliva. The same trend was seen for the amount of saliva in relation to the presence of a neoglottic bar during phonation, this relation was however not significant below the .05 level (p=.084). In the group where saliva was seen during digital high-speed imaging, often regurgitation was seen during videofluoroscopy. When a larger amount of saliva is seen with digital high-speed imaging, the tonicity of the neoglottis judged with videofluoroscopy is often hypotonic.

The visibility of the origin of the neoglottis was related to the appearance of a neoglottic bar during phonation (p=.036). This is represented in Table 9.13. In the patient group with a visible origin of the neoglottis as seen with digital high-speed imaging, more often appearance of a neoglottic bar during phonation is seen on videofluoroscopy recordings. This shows that, although no one-to-one relation, in the majority of the cases when a neoglottic bar is seen in the lateral view with videofluoroscopy, also a neoglottis is seen in the 'birds-eye' view of endoscopic digital high-speed imaging.
Table 9.13. Relation between the *visibility of the origin of the neoglottis* seen with digital high-speed imaging (columns) and the presence of a neoglottic bar during phonation judged with videofluoroscopy (rows). Numbers represent number of patients, in total n=37.

<table>
<thead>
<tr>
<th>Videofluoroscopy assessment</th>
<th>Visibility of the origin of the neoglottis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>yes</td>
</tr>
<tr>
<td>Neoglottic bar during phonation</td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>3</td>
</tr>
<tr>
<td>1 or 2</td>
<td>22</td>
</tr>
</tbody>
</table>

The *shape of the neoglottis* was related to the *appearance of a neoglottic bar at rest* (p=.046). In Table 9.14 this relation is shown.

Table 9.14 Relation between the *shape of the neoglottis* judged with digital high-speed imaging (columns) and the appearance of a neoglottic bar at rest judged with videofluoroscopy (rows). Numbers are number of patients, in total n=28.

<table>
<thead>
<tr>
<th>Visual assessment</th>
<th>Shape of the neoglottis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>circular</td>
</tr>
<tr>
<td>Neoglottic bar at rest</td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>2</td>
</tr>
<tr>
<td>1 or 2</td>
<td>4</td>
</tr>
</tbody>
</table>

As can be seen side-to-side split shapes and irregular shapes are seen more often when there is a neoglottic bar at rest, while anterior-posterior split shapes are seen more often when there is no visible neoglottic bar at rest seen with videofluoroscopy. Regarding *location of the visible vibration, mucosal wave, regularity of the vibration, and closure phase*, no relations were found with the results of videofluoroscopy.

9.3.3.2 Digital high-speed imaging versus quantitative measures of videofluoroscopy

The relations between the visual assessments of the digital high-speed imaging recordings and the quantitative measures of the videofluoroscopy recordings are investigated with t-tests for two independent samples or analyses of variance (three or more independent samples). Relations were found between the visual assessments of *the amount of saliva* and of the *presence of a mucosal wave* in the digital high-speed recordings and some of the quantitative measures obtained from the videofluoroscopy recordings. In Table 9.15 the results for the amount of saliva are shown.

As can be seen in Table 9.15, the amount of saliva seen with digital high-speed imaging is related to the quantitative measures maximal distance during phonation, minimal distance at rest and the index of the maximal distance during phonation and the maximal distance at rest. The average maximal distance during phonation is larger, the minimal distance is smaller and the index is larger in the patient group with no or a little saliva compared to the patient group with a large or obstructing amount of saliva. This indicates that when there is a better neoglottic closure, the amount of saliva at or above the level of the neoglottis is smaller.

Regarding the *presence of a mucosal wave*, it was shown that the *maximal distance during phonation* was larger in the patient group in which a mucosal wave is seen compared to the patient group in which this was not the case (p=.022; ‘yes/slightly’=12.24; ‘no’=8.67). This could indicate that with a larger subneoglottic distance, pointing to a tighter closure of the neoglottis (see Chapter 6) it is more likely that a mucosal wave of the mucosa overlying the muscles is present.
Table 9.15. Relations between the visual assessments of the amount of saliva in digital high-speed recordings (columns 2-4) and the quantitative measures of the neoglottis performed on videofluoroscopy recordings (column 1). The p-value is given in column 2.

<table>
<thead>
<tr>
<th>Quantitative measure</th>
<th>p-value</th>
<th>Amount of saliva</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>none/a little</td>
</tr>
<tr>
<td>MAXPHON</td>
<td>.018</td>
<td>11.56</td>
</tr>
<tr>
<td>MINREST</td>
<td>&lt;.001</td>
<td>0.55</td>
</tr>
<tr>
<td>INDEX</td>
<td>.020</td>
<td>2.32</td>
</tr>
</tbody>
</table>

Regarding the visibility of the origin of the neoglottis, shape of the neoglottis, location of the visible vibration, regularity of the vibration, and closure phase no relations were found with the quantitative measures.

9.3.4 RELATIONS WITH SOCIODEMOGRAPHIC AND CLINICAL FACTORS

Relations between the visual assessments of the digital high-speed imaging recordings and the sociodemographic and clinical factors are investigated by means of chi-squared tests. Regarding reconstruction, the patient group after standard total laryngectomy differed from the patient group after partial or full pharyngeal reconstruction for the amount of saliva (p=.015) and shape of the neoglottis (p=.029). Separating the reconstruction group into partial (PM-flap) and full pharyngeal reconstruction (Radial Fore Arm Flap) and esophagus resection (tubed gastric pull-up) gives more insight in the differences between the different types of reconstruction. Chi-squared tests on the data with the types of reconstruction separated also showed significant relations for the amount of saliva (p=.029) and shape (p=.021). In Table 9.16 the relation between the different types of surgery and the amount of saliva is shown and in Table 9.17 the relation between the different types of surgery and the different shapes of the neoglottis is shown.

Table 9.16. Relation between the different types of surgery (rows) and the amount of saliva (columns) seen during digital high-speed imaging (p=.021). Numbers are number of patients, in total n=44.

<table>
<thead>
<tr>
<th>Extent of surgery</th>
<th>Amount of saliva</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>none/a little</td>
</tr>
<tr>
<td>Standard total laryngectomy</td>
<td>19</td>
</tr>
<tr>
<td>Pectoralis Major Flap</td>
<td>1</td>
</tr>
<tr>
<td>Free Radial Forearm Flap</td>
<td>0</td>
</tr>
<tr>
<td>Tubed gastric pull-up</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 9.17. Relation between the different types of surgery (rows) and the different shapes of the neoglottis (columns) seen during digital high-speed imaging (p=.029). Numbers are number of patients, in total n=44.

<table>
<thead>
<tr>
<th>Extent of surgery</th>
<th>Shape of the neoglottis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>circular</td>
</tr>
<tr>
<td>Standard total laryngectomy</td>
<td>5</td>
</tr>
<tr>
<td>Pectoralis Major Flap</td>
<td>0</td>
</tr>
<tr>
<td>Free Radial Forearm Flap</td>
<td>1</td>
</tr>
<tr>
<td>Tubed gastric pull-up</td>
<td>1</td>
</tr>
</tbody>
</table>

Since the surgical differences between the patient group with a standard total laryngectomy and the patient group with a reconstruction are large, the clinical parameters sex, age, postoperative follow-up, myotomy, neurectomy, radical neck dissection and radiotherapy are studied within the patient group after standard total laryngectomy only. The reconstruction group is too small to perform statistical analysis on subgroups within it.

Regarding sex, a relation was found with the regularity of the vibration of the neoglottis (p=.024), female patients had more often a regular vibrating neoglottis than male patients. Regarding myotomy a relation was seen with the shape of the neoglottis (p=.017), in the patient group with a myotomy no side-to-side split shapes were seen, and in the group without a myotomy more often an irregular shape was seen. No relations were found with age, postoperative follow-up, neurectomy, radical neck dissection, and radiotherapy.

9.4 DISCUSSION

The aims of the present chapter were to investigate the results of digital high-speed imaging described in Chapter 8, in relation to perceptual evaluations (Chapter 4) and acoustic analysis (Chapter 5) of tracheoesophageal voice quality, to the results of video fluoroscopy (Chapter 6), and to the sociodemographic and clinical factors described in Chapter 3.

Several interesting relations between the high speed imaging observations and the other assessment techniques could be found. When the neoglottis is visible during high-speed imaging, the voice quality is more often judged as ‘good’. Since during high-speed imaging one looks into the esophagus with a ‘birds-eye view’, it might be that the neoglottis cannot be seen when it is situated at a deep level. This relation however shows that in most patients with a good voice the neoglottis can indeed be seen from above, only in two patients from the good speaker group no neoglottis was seen.

The amount of saliva, visibility of the neoglottis, shape of the neoglottis and regularity of the vibration are neoglottic characteristics that are related to more specific aspects of voice quality, as reflected in the relations with the perceptual scales. Regarding the amount of saliva and the visibility of the origin of the neoglottis these relations are not surprising, since saliva interferes with the voice production and the visibility of the neoglottis has already been shown to be of importance for good voice quality in the results of videofluoroscopy (Chapter 7). Shape of the neoglottis and regularity of vibration are however important in this respect, since they cannot be judged by videofluoroscopy or stroboscopy and provide thus new, additional information. Especially, the side-to-side split shape and the irregular shape are judged as more expressive and intelligible, the irregular shapes are also judged as higher and shriller. The positive effect of a side-to-side (coronal) split shape of the neoglottis on voice quality might, in analogy to the normal glottis, point to the preference of a split shape for good voice production. Anterior-posterior (sagital) split shapes might be less favorable since in this respect, it is expected that the musculature activity normally should be located mainly at the posterior esophageal wall. When regular vibration of the neoglottis is seen, relations with the perceptual scales showed that the voice sounds steadier, higher and shriller.
Although one should expect a relation between the regularity of the vibration and scales regarding voice quality (in normal voices short-term irregularities (jitter) are related to hoarseness), the relation with the perceptual scale steady-unsteady is not surprising. Most probably the visual assessment of regularity of the neoglottis merely reflects long-term instability and not the smaller, short-term irregularities. Objective image-analysis of the high-speed recordings is a better tool for investigation of short-term cycle-to-cycle irregularities, as well as for measuring the fundamental frequency of the vibration. These objective analysis methods are unfortunately still very time consuming. They cannot be performed automatically for the majority of the high-speed recordings of the neoglottis, since the starting point of the vibration and the contour of the neoglottis cannot always be defined automatically. For one patient, as an example, the development of the contour of the neoglottis and the extraction of fundamental frequency are calculated (Wittenberg et al., 1999). Further developments of image analyses methods and increased quality of the high-speed recordings might make these objective analyses in future possible for a larger number of recordings.

In addition to the relations with the perceptual evaluations, the visual assessments of the digital high-speed recordings are also related to the results of acoustic analyses. In analogy to the relations with the overall judgment of voice quality, the visibility of the origin of the neoglottis is also related to the four acoustic signal types. These relations are even stronger than the relations shown with the overall perceptual judgment. When the origin of the neoglottis is visible, the resulting signal type is mostly type I (stable & harmonic) or type II (stable & at least one harmonic). Also, the interference of saliva with voice production is reflected in the acoustic signal types; when a large or obstructing amount of saliva is seen, the signal type is III (unstable or partly harmonic), or IV (barely harmonic). Like with perceptual evaluations, here also more specific relations are shown with the separate acoustic parameters. For the amount of saliva, shape of the neoglottis, regularity of the vibration and location of the visible vibration relations were found with the acoustic measures. As can be expected, the acoustic measures were better when less saliva was seen. The relations, regarding pitch, between the perceptual evaluations and the irregular shapes are confirmed by the fundamental frequency measures. The harmonics-to-noise ratio points to a better voice sound in the circular and side-to-side split shapes and maximum phonation time is longer in the split shapes. A regular vibration pattern also results in a better voice quality according to the harmonics-to-noise ratio and the percentage of voiced. The harmonics-to-noise ratio is highest when the vibration is located at the posterior esophageal wall. These relations show that vibration of the posterior wall and a circular or side-to-side split shaped neoglottis result in better voice quality according to the acoustic measures.

Digital high-speed imaging and videofluoroscopy are both used to investigate anatomical and morphologic characteristics of the neoglottis, though both from a different point of view. Digital high-speed imaging provides a 'birds-eye view' (comparable to stroboscopy), while videofluoroscopy enables a lateral view of the neoglottis. Relations between both methods are investigated to gain insight in the overlap between both methods and the complementary value of high-speed imaging. Not surprisingly, the amount of saliva seen during digital high-speed imaging is related to regurgitation of barium seen during videofluoroscopy and the tonicity of the neoglottis seen during videofluoroscopy. The tonicity can at present only be investigated with videofluoroscopy; no information is available yet about the visual characteristics of tonicity in a neoglottis when seen from above. Regurgitation occurs in the majority of the hypotonic cases and is seen as bubbling of saliva during digital high-speed imaging. Also, when a neoglottic bar was seen during videofluoroscopy, in the majority of the cases a neoglottis was visible in the high-speed recordings. More interestingly, when a neoglottic bar is seen at rest during videofluoroscopy, side-to-side split shapes and irregular
shapes are seen more often, while anterior-posterior split shapes are seen more often when no neoglottic bar is seen during videofluoroscopy. For the location of the vibration, presence of a mucosal wave, regularity of the vibration and closure phase, no relations were found with videofluoroscopy. These assessments are thus of additional value to the assessments that can be made on videofluoroscopy recordings. Also, assessment of shape is of importance since there is no direct relation with observations during videofluoroscopy. The regurgitation seen during videofluoroscopy might be influenced by the swallow of barium shortly before phonation, the amount of saliva judged with digital high-speed imaging (where no barium is swallowed) might therefore be a better representative for the normal situation during phonation. On the other hand, it might also be a worse representative for normal voice production since the patient is requested to stick out his tongue to enable the view in the esophagus. In this respect, improvements in digital high-speed imaging, enabling recording with a flexible fiberoptic endoscope (which is introduced through the nose and does not require that the patient sticks his tongue out), would resolve this problem.

Only the visual assessment of the amount of saliva and the presence of a mucosal wave are related to some quantitative measures of the neoglottis during videofluoroscopy. In relation to hypotonicity, the quantitative measures show that the amount of saliva is smaller when the neoglottis is closed with some tension during phonation. Interestingly, the average maximal distance during phonation was larger in the patient group showing a mucosal wave of the esophageal wall(s) during phonation. This can be explained by the tighter neoglottic closure that is seen when the maximal subneoglottic distance is larger. Most probably the muscular closure is good in this patient group and the upper mucosal layers show the mucosal wave.

The sociodemographic and clinical factors that are studied are sex, age, postoperative follow-up, reconstruction, myotomy, neurectomy, radical neck dissection and radiotherapy. As can be expected, the patient group with a partial or full pharyngeal reconstruction showed less favorable characteristics than the patient group with a standard total laryngectomy. More detailed investigation on the small subgroups of reconstruction shows that the partial pharyngeal reconstructions with PM-flap show better results than the full pharyngeal reconstruction with radial forearm flap or tubed gastric pull-up. This might not be surprising, since in the PM-flap reconstruction cases a small strip of normal mucosa, mostly at the posterior wall, remains in situ. Regarding sex, unlike other parts of the study, for the first time a difference is found between the male and female patients, the female patients show relatively more often a regular vibrating neoglottis. However, no meaningful explanations can be formulated for this difference as yet. Furthermore, myotomy was found to be related to the shape of the neoglottis. No explanation can be given for this either, but it should be kept in mind that only a small subgroup of patients underwent a myotomy (n=7).

CONCLUDING REMARKS

In the present chapter a relatively new evaluation tool, endoscopic digital high-speed imaging, still merely used as a research tool and currently under development for use in clinical settings, has been related to perceptual evaluations and acoustic analyses of voice quality. It also has been related to videofluoroscopy recordings, a method that is more widely and also clinically used. In relation to the perceptual evaluations, the amount of saliva, the visibility of the neoglottis, the shape of the neoglottis and the regularity of the vibration are important characteristics. The same assessments were relevant in relation to the acoustic parameters, additionally, the location of the visible vibration showed some relations with the acoustic parameters. Relations with the results of videofluoroscopy show that the judgment of the amount of saliva in digital high-speed imaging and the regurgitation of barium during videofluoroscopy are somewhat overlapping, although it is thought that the judgment of the
amount of saliva in digital high-speed imaging is more representative for daily life. Another overlapping judgment is that of the visibility of the origin of the neoglottis during videofluoroscopy and the appearance of a neoglottic bar in videofluoroscopy, the latter to be preferred since also a deep location of the neoglottis can be identified.

Important characteristics that appeared to be relevant for voice quality and that can be seen during digital high-speed imaging and not, or less good, with videofluoroscopy, are the amount of saliva, the shape of the neoglottis, the regularity of the vibration and the location of the visible vibration.

It can be concluded that digital high-speed imaging can provide valuable extra information about the neoglottic characteristics in relation to tracheoesophageal voice quality. In relation to videofluoroscopy it can be seen partly as additional information and partly as an overlap. Future possibilities for automated objective analyses of these recording would increase its usefulness for research and in clinical practice even more.

REFERENCES


