Objective evaluation of TE-substitute voice and speech

U. Eysholdt, M. Schuster, J. Lohscheller, F. Rosanowski

Department of Phoniatrics and Pediatric Audiology, University Hospital of Erlangen, Germany
ulrich.eysholdt@uk-erlangen.de

Abstract
Three studies are presented which aim at rater-independent methods to evaluate voice quality, speech intelligibility and the benefit of rehabilitation of the patients: (1) real time endoscopy of the vibrating PE segment, (2) quantification of speech intelligibility by means of a computerized speech recognition system and (3) rating of the laryngectomees’ quality-of-life in relation to his substitute voice.

1. Real time endoscopy of the vibrating PE segment

Objectives: Evaluation of the vibration pattern of the substitute voice generator of patients who have undergone laryngectomy. For automatic quantification of the oscillations of the pharyngoesophageal (PE) segments, image processing of digital high-speed video sequences is applied.
Patients and methods: Endoscopic recordings were taken of 10 men who underwent laryngectomy (mean +/- SD age, 61.5 +/- 5.2 years) during sustained phonation of a vowel using a 90 degree endoscope coupled to a high-speed camera. An image-processing algorithm was developed to automatically define the pseudoglottis in each recording and track its movements.

Results: In a first step, the high-speed recordings of the substitute voice generator have been replayed in slow motion and visually categorized. The forms and oscillation characteristics of the pseudoglottides varied considerably: 3 pseudoglottides were circular, 6 were split shaped, and 1 was triangle shaped. Quasi-periodic openings and closings were observed and automatically detected by the described algorithm in each recording independently from quality of the recording and from morphologic and oscillation characteristics of the PE segment. The frequencies of the extracted oscillations of the pseudoglottides correspond to the structure of the acoustic signals.

Conclusion: Automatic image processing of PE segments derived from high-speed endoscopic recordings enables the detection and quantification of the substitute voice generator’s oscillations in high temporal resolution. These data directly prove that the detected pseudoglottis is the source of the substitute voice. Close relations between substitute voice and functional properties of the PE
segment exist. In the future, these data will be interpreted by applying biomechanical models of the PE segment. Presumably, results may help to optimize surgical and adaptive procedures for specific substitute voice restoration.

2. Quantification of speech intelligibility by means of a computerized speech recognition system

Objectives: Substitute speech after laryngectomy is characterized by restricted aero-acoustic properties in comparison with laryngeal speech and has therefore lower intelligibility. Until now, an objective means to determine and quantify the intelligibility does not exist although the intelligibility can serve as a global outcome parameter of voice restoration after laryngectomy.

Patients and methods: An automatic speech recognition system was applied on recordings of a standard text read by 18 German male laryngectomees with tracheoesophageal substitute speech. The system was trained with normal laryngeal speakers and not adapted to severely disturbed voices. Substitute speech was compared to laryngeal speech of a control group. Subjective evaluation of intelligibility was performed by a panel of 5 experts and confronted to automatic speech evaluation.

Results: Substitute speech showed lower syllables/sec and lower word accuracy than laryngeal speech. Automatic speech recognition for substitute speech yielded word accuracy between 10.0% and 50% (28.7% ± 12.1%) with sufficient discrimination. It complied with experts’ subjective evaluation of intelligibility. The multi-rater kappa of the experts alone differed not from the multi-rater kappa of experts and recognizer.

Conclusion: Automatic speech recognition serves as a good means to objectify and quantify global speech outcome of laryngectomees. For clinical use, the speech recognition system will be adapted to disturbed voices and can also be applied in other languages.

3. Rating of the laryngectomees’ quality-of-life in relation to his substitute voice

Objectives: Health-related quality of life (QoL) and subjective health have become popular constructs for the evaluation of both efficacy and efficiency of diagnostic and therapeutic procedures in medicine. QoL is considered a multidimensional construct encompassing physical, mental and social facets of life. It is an accepted outcome parameter not only in international classification systems such as ICIDH and ICF, but as well in clinical guidelines and disease management programs. Measuring quality of life allows for comparison of different diseases though it certainly lacks disease specific aspects. Thus, it has to be assumed that in patients with distinct functional deficits QoL cannot cover
all aspects that are important for the individual patient. This study focusses on laryngectomees and their self-evaluation of post-laryngectomy speech. It is well known that these patients experience a decreased QoL compared to patients after partial laryngectomy or healthy persons. In this study, the impact of voice restoration on the laryngectomees’ QoL was evaluated.

Patients and methods: In 20 male laryngectomees aged 62 ± 8 years, relations between QoL and voice handicap were evaluated using two instruments as proposed in the international literature, i.e. the Short-Form Health Survey (SF-36) and the Voice Handicap Index (VHI). All patients had successfully been using tracheoesophageal substitute voice for at least one year. Complete data sets were available from all patients. Data were analysed using Microsoft Excel® and Sigma Plot®, Jandel Corp. software packages.

Results: Results of both the SF-36 and the VHI reveal wide interindividual ranges. There are statistically significant correlations (p < 0.05) between the SF-36’s scales General Health and Vitality and the VHI, whereas no correlations were found between social and psychological scales and the voice handicap.

Conclusion: Wide ranges of the data obtained reveal that obviously both health related quality of life and voice handicap are not affected in a group specific way. So, both tests are of clinical value to depict individual aspects of wellbeing after laryngectomy. The combination of VHI and SF-36 illuminates correlations between general and special subjective aspects. Strikingly, the laryngectomees’ social and psychological status is not related to their voice handicap. Data allow for conclusion that laryngectomees’ substitute voice problems do not affect social aspects of quality of life in a disease specific way, at least as long as voice restoration was successful at all.

Key words: laryngectomy, high speed endoscopy, substitute speech, automatic speech recognition, speech intelligibility, Quality of life, voice handicap index

4. References


