Biomedical Signal Processing and Control 17 (2015) 3–10

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Correlation between the quantitative video laryngostroboscopic measurements and parameters of multidimensional voice assessment

Virgilijus Uloza[a], Aurelija Vegiene[a,*], Viktoras Saferis[b]

[a] Department of Otolaryngology, Lithuanian University of Health Sciences, Eiveniu 2, LT-50009 Kaunas, Lithuania
[b] Department of Physics, Mathematics and Biophysics, Lithuanian University of Health Sciences, Eiveniu 4, LT-50009 Kaunas, Lithuania

ARTICLE INFO

Article history:
Received 19 March 2014
Received in revised form
15 September 2014
Accepted 7 October 2014
Available online 7 November 2014

Keywords:
Laryngostrobosopy
Acoustic voice parameters
VRP

ABSTRACT

The purpose of this study was to evaluate quantitatively the parameters of the video laryngostroboscopy (VLS) and determine correlations between the VLS parameters and acoustic vocal function parameters. Digital VLS recordings, acoustic voice assessment, calculation of dysphonia severity index (DSI) and registration of voice range profile (VRP) were performed for 206 individuals: 50 healthy and 156 patients with mass lesions of vocal folds and paralysis. 90 of lesions were unilateral, 66 – bilateral. VLS parameters were derived using objective measures made from a single image taken from the VLS recording of a sustained vowel: glottal areas, glottal widths and distances, vocal fold angles. As the result of Fisher's linear discriminant analysis, 11 VLS measurements were identified to be relevant distinguishing between normal and pathological voice groups. Correlations between the VLS parameters and results of acoustic voice analysis parameters, DSI and VRP measurements were tested using Pearson's correlation coefficient. The correlations of VLS variables and acoustic voice measurements were moderate and statistically significant. In pathological voices numerical values of VLS parameters measured reveal significant deviances from these in normal voice; therefore quantification of VLS parameters appears to be sensitive and specific distinguishing normal and pathological voices patients groups. Analysis of correlations between the quantitative measurements obtained via VLS and acoustic voice parameters provides more versatile approach into the pathophysiology of phonation and suggests the documented and measurable evidence of complex mechanisms of vocal function.

1. Introduction

Video laryngostroboscopy (VLS) currently is considered as the most important, popular and clinically feasible imaging method for vocal fold visualization and evaluation of patients with voice disorders [1–4]. Laryngeal examination with VLS has been shown to be essential to indentify the underlying cause of hoarseness increasing the accuracy of diagnostics up to 68.3% [5]. VLS provides a real time information and therefore is suitable for clear imaging of normal or pathological vibrations of vocal folds and for detection of even small mass lesions or infiltrative processes of the vocal folds thus increasing diagnostic possibilities of various laryngeal diseases including early diagnostics of laryngeal carcinoma, despite that stroboscopy cannot be used reliably for classifying laryngeal dysplasia [6,7]. This valuable imaging tool is also the most useful in demand to assess the functional outcomes of therapeutical or surgical treatment of laryngeal diseases.

Development of new objective methods of measurement of vocal fold vibrations, i.e., videoendoscopic high-speed imaging (HSI) and videokymography (VKG) is in promising progress [8–11]. HSI is a frequency-independent visualization technique that allows the observation of previously unavailable information related to vocal fold anatomy during phonation and therefore can be viewed as complementary to VLS, particularly in cases of moderate to severe aperiodicity potentially aiding in clinical decision-making. In voice disorders with aperiodic voices, the HIS showed to be significantly more accurate and interpretable than VLS, allowing for evaluation of phase asymmetry even in cases when VLS was incapable [12].

However, affordable techniques that allow to quantitatively and precise assess the vibratory pattern of the vocal folds have not yet been achieved for actual clinical applications. The high cost of equipment and intensive labor associated with data analysis may

* Corresponding author. Tel.: +370 614 30297; fax: +370 37 326862.
E-mail addresses: virgilijus.ulozas@kaunoklinikos.lt (V. Uloza), v_aurelija@yahoo.com (A. Vegiene), viktoras.saferis@lsmuni.lt (V. Saferis).

http://dx.doi.org/10.1016/j.bspc.2014.10.006
1746-8094/

be the reasons why investigators and clinicians were rather reluctant to adopt these new methods [4]. Moreover, results of the study performed by Mendelsohn et al. showed that HSI appears to be without improved diagnostic capacity as a stand-alone technique apart from VLS [13].

VLS still remains recognized and well-established laryngeal examination modality in the specialized evaluation of the dysphonic patient [3–5]. However, the stroboscopic examination relies on subjective qualitative judgments of physician. This subjective substance background of the interpretation of stroboscopic examination results, which is strongly dependent on the level of experience of the examiner, presents the main limitation of the VLS. To reduce subjectivity of interpretation of VLS findings different VLS rating scales consisting of various parameters, i.e., periodicity and amplitude of vocal fold vibration, mucosal wave, vertical level, glottal closure, phase closure, phase symmetry, presence of non-vibrating portions of the vocal fold have been introduced [2,14]. Nevertheless, subjective and qualitative nature of stroboscopy, differences in rating scales combined with potential pitch tracking errors, necessity of stable phonation frequency and periodic vibrations make data and results of different studies hardly compatible and retard the elaboration of objective measurements and standards to assess vocal fold vibrations and voice disorders using VLS [15,16].

On the other hand, recent studies demonstrate that even small set of stroboscopic ratings is an adequate representation for most of the variance of all laryngostroboscopic characteristics [17,18]. Furthermore, quantification of the basic VLS parameters using visual analogue scales (VAS) revealed a high inter-rater and intra-rater reliability for the majority of basic VLS parameters that showed high sensitivity and specificity distinguishing healthy and pathological voice patient groups [19].

However, despite the clinical availability of methods for objective image analysis there is a scant of information in the literature on objective measurements of VLS parameters and correlations between the data of VLS examination and the results of assessment of laryngeal phonatory function using acoustic measurements [4,20–23].

The aim of this study is to evaluate quantitatively the objective VLS parameters revealing their reliability discriminating normal and pathological voice and to assess correlations between objective VLS parameters and multidimensional vocal function measurements.

2. Materials and methods

2.1. Study group

A study group consisted of 206 individuals examined at the Department of Otolaryngology of Lithuanian University of Health Sciences, Kaunas, Lithuania. The demographic data of the study group is presented in Table 1.

The **normal voice subgroup** was composed of 50 selected healthy volunteer individuals who considered their voice as normal. They had no complaints concerning their voice and no history of chronic laryngeal diseases or other long-lasting voice disorders. The voices of this group of individuals were also evaluated as healthy voices by clinical voice specialists, and no pathological alterations in the larynx found during VLS.

The **pathological voice subgroup** consisted of 156 patients who represented a rather common, clinically discriminative group of laryngeal diseases, that is, mass lesions of vocal folds and paralysis. Mass lesions of vocal folds included in the study consisted of nodules, polyps, cysts, papillomata, keratosis, and carcinoma. 90 of lesions were unilateral, 66 – bilateral.

2.2. VLS quantification

Digital high quality VLS recordings were performed with the XION EndoSTROB DX device (XION GmbH, Berlin, Germany) using a 90° rigid endoscope. The VLS examination and recordings were performed during modal phonation, i.e., each subject was asked to sustain the vowel /ee/ at a steady, comfortable pitch and loudness. Phonation time was kept long enough to allow for registration of a sustained phonation and at least one complete cycle of vibration.

Individual VLS recordings of the tested group were sequenced into separate images using *Kinovea 0.8.15* software (*Kinovea open source Project*, www.kinovea.org). Thus, each individual VLS recording was divided into 300–600 consequent images. The most representative images (maximum lateral amplitude and minimum lateral amplitude, maximum and minimum lateral mucosal wave position) from each recording were selected for analysis and measurements.

The clinically feasible objective measures have been derived from a single image from the VLS recording including measurements of glottal areas and areas of vocal fold lesions, distances (length and width of vocal folds and mass lesions), and angles between vocal folds. It was hypothesized that these measures in the presence of mass lesions of vocal folds would allow assessing and quantifying correlations between deteriorated phonatory pattern and deviation of acoustic voice quality parameters and parameters of VRP.

Contours of glottal area both, during maximum opening and maximum closure were traced manually with a mouse and filled in with color (Fig. 1). Thus, colored areas of the glottis were computed in pixels using the updated "Laringometrija" program [24]. Areas of mass lesions of vocal folds were measured manually in the same manner. Distances (width and length of vocal folds and mass lesions) were measured in pixels.

A correct comparison of images of separate VLS recordings of different larynges requires equal magnifications and view angles in all images. However, in clinical settings the distances between the optical lens of the rigid endoscope and the level of vocal folds are rather different. Therefore, to minimize the influence of these differences the relative measures in pixels were used, i.e., by dividing the areas measured, and dividing length and width of the vocal fold measured. The angles between the vocal folds were measured manually in grades. To calculate angles between the vocal folds separately the line corresponding to the glottis axis was drawn.

Totally, 25 different measurements were performed, including length and width of vocal folds, areas of mass lesions, glottal gap areas during maximum abduction and adduction, etc. However after Fisher's linear discriminant analysis only 11 out of 25 VLS measurements were indentified to be statistically significant and relevant distinguishing between normal and pathological voice groups (Table 2).

Computation of relevant VLS parameters is visualized in Fig. 1.

2.3. Acoustic analysis

Digitized voice recordings of sustained phonation of the vowel sound /a/ (as in the English word "large") were obtained in a sound-proof box by the lingWAVES sound pressure level (SPL) meter microphone placed at a 30.0 cm distance from the mouth and at about 90° microphone-to-mouth angle. The voice recordings were made in the "wav" file format at the rate of 44.100 samples per second. Sixteen bits were allocated to one sample. Segments of at least 2-s duration of the sustained vowel /a:/ were analyzed using Voice Diagnostic Center lingWaves software, Version 2.5 (WEVOSYS, Forchheim, Germany) software. Acoustic voice signal data were obtained for: (1) fundamental frequency (Mean_Fo, Hz), (2) standard deviation of Fo (SD_F0, Hz), (3) maximum Fo (Max_Fo,