Perceptual and Acoustic Characteristics of Voice Changes in Reflux Laryngitis Patients

*Ruta Pribuisiene, *Virgilijus Uloza, †Limas Kupcinskas, and †Laimas Jonaitis

Kaunas, Lithuania

Summary: The aim of the study was to outline the multidimensional perceptual, subjective, and instrumental acoustic voice changes in the group of reflux laryngitis (RL) patients. Data of multidimensional voice assessment of 108 RL patients and 90 healthy persons of the control group were subjected to comparative analysis. A slight hoarseness according to the GRB (G-grade, R-rough, B-breathy) scale was prevailing in the RL patients group. Statistically significant difference ($P < 0.001$) between RL patients group and the control group was found of all voice parameters measured, with the patients having worse results—increased mean jitter, shimmer, normalized noise energy, voice handicap index (VHI), and decreased parameters of phonetogram. The results of the study demonstrated that multidimensional voice assessment documented deteriorated voice quality and restricted phonation capabilities in the tested group of RL patients.

Key Words: Reflux laryngitis—Hoarseness—Jitter—Shimmer—Normalized noise energy—Phonetogram—Voice handicap index.

INTRODUCTION

Gastroesophageal reflux disease (GERD) is a widespread affection in the world. Reflux laryngitis (RL) represents one of the most common atypical manifestations of GERD characterized by morphologic and functional changes in the larynx with associated clinical symptoms such as voice disturbances, throat clearing, itching, and chronic cough.1–4 The peculiarities of topographic localization and the physiology of the larynx facilitate the direct negative impact of the pathological gastroesophageal reflux on the laryngeal mucosa manifesting as edema, erythema, and hypertrophy in all parts of the larynx.1–5 This can interfere with the phonation pattern and cause hoarseness, voice fatigue, and hypertonic phonation.1,3,6–8 Vagally mediated reflex can also occur when the lower esophagus is exposed to gastric acid and stimulates abnormal muscle contraction in the upper aerodigestive tract, along with excessive salivation, that has a negative impact on laryngeal phonatory function.9

Recurring hoarseness is supposed to be one of the most common head and neck symptoms of atypical GERD. The incidence of that symptom varies according to the data of different studies from 65% to 88%.9,10 To date, despite numerous diagnostic tests, including radiologic, scintigraphic, acid provocation, manometric, and single- and dual-probe pH
monitoring, there is a lack of studies systematically and completely describing voice changes associated with GERD and RL.

Therefore, the aim of this study was to outline the multidimensional perceptual and instrumental acoustic and aerodynamic voice changes in the group of RL patients, to compare with the control group of healthy persons and to determine whether voice disturbances in RL patients can be objectively described by instrumental means.

PATIENTS AND METHODS

The study was carried out at the Department of Otolaryngology of Kaunas University of Medicine, Kaunas, Lithuania during the years 2000–2003. All in all, 217 consecutive outpatients visiting otorhinolaryngologist because of complaints of permanent, hoarseness, throat itching, cough, “globus” sensation, and heartburn were suspected of having atypical GERD and examined.

After evaluation of their complaints and history, all patients underwent video laryngostroboscopy (VLS), esophageal endoscopy, and biopsy. Patients with evident causes of chronic hoarseness, ie, catarrhal inflammation of the upper respiratory tract within 1 month, signs of allergic or fungal alterations of the larynx, voice abuse, mass lesions of the vocal cords (nodules, polyps, cysts, granuloma), and those who had sustained some mechanical or chemical damage to the laryngeal and pharyngeal mucosa because of operations, burns, or after an intake of stomach-secretion suppressive agents within 2 months, were excluded from the study.

Diagnosis of RL was based on (1) the complaints (permanent hoarseness, throat itching and clearing, cough, heartburn, globus sensation) exceeding 3 months in duration, (2) typical laryngoscopic findings of RL (edema, erythema, roughness, hypertrophy of mucosa of the posterior glottis), and (3) detection of reflux esophagitis as a subsequence of pathological gastroesophageal reflux.1–4,11 Reflux esophagitis was revealed on esophageal endoscopy or biopsy only in 108 patients; the remaining 111 patients were found to have no signs of reflux esophagitis. An erosive form of reflux esophagitis was diagnosed in 36 cases (33.3%). In the other cases, a nonerosive form of reflux esophagitis was revealed.11

We did not use esophageal pH monitoring in this study as it was expensive and inconvenient for the patient, and according to literature statements, it was not definitive in helping to diagnose atypical GERD.5,9 Dual-probe pH monitoring, which is considered by some authors to be the current golden standard, has had a variably reported sensitivity of 17.5% to 78.8% for the diagnostics of atypical forms of GERD, and the reports were found to be controversial.5,8,9 Therefore, esophageal endoscopy and biopsy were preferred in our study for the diagnostics of the consequences of the pathological reflux of both erosive and nonerosive forms of GERD.11,12

Finally, after complex examination RL was proved and documented in 108 patients who had (1) complaints of permanent hoarseness, throat clearing and itching, and heartburn; (2) VLS signs of RL; and (3) reflux esophagitis. These persons (40 men and 68 women, their age ranging between 18 and 64 years; mean 40.1 years, SD 12.9 years) comprised the study group. Eleven patients (10.2%) were cigarette smokers. As other possible causes of voice complaints were excluded in the study group, RL was considered to be the main etiological factor of deterioration of phonatory function of the larynx.

The control group consisted of 90 randomly selected healthy persons who had no chronic laryngeal diseases or other long-lasting voice disorders and who had never seen an otolaryngologist concerning the voice problems. No pathological alterations in the larynx typical for RL were found during indirect optical laryngoscopy. There were 36 men and 54 women. Their age ranging from 18 to 65 years (mean 36.9 years; SD 11.5) in this group. Ten persons (11.1%) were cigarette smokers.

Study and control groups were compared in respect with the number of subjects with and without vocal training.13 Subjects without voice training were prevalent both in study (96 patients, 88.9%) and in control (84 subjects, 93.3%) groups, and there was no statistically significant difference between the number of the subjects with and without vocal training in the study and control groups ($P = 0.33$). Equally, there were no professional singers among the subjects with vocal training. Consequently, the ratio of subjects in the groups studied did not differ
significantly within respect to gender, age, vocal training and usage, social status, and smoking habits ($P < 0.05$).

Vocal function was evaluated with a multidimensional set of VLS, perceptual, acoustic, aerodynamic, and subjective measurements according to the protocol elaborated by the Committee on Phoniatrics of the European Laryngological Society.  

**Video laryngostroboscopy**

Kay Elemetrics RLS (model 9100; Kay Elemetrics Corporation, Lincoln Park, NJ) device with 70° rigid telescope examined the larynx. The following VLS parameters were examined and assessed: (1) longitudinal glottal closure, (2) regularity, (3) symmetry of vibration of vocal folds (VF), and (4) mucosal wave. For rating of each parameter, a 100-mm visual analogue scale (VAS) was indicated (0—no deviation, 100—severe deviation).

**Perceptual voice evaluation**

Digitized voice recordings of a standard phonetically balanced passage read out on habitual pitch and loudness were subjected to perceptual evaluation of hoarseness by two experienced voice specialists serving as judges. A simplified, clinically feasible version of the GRBAS scale, consisting of G (grade), R (rough), and B (breathy) factors was adapted to assess hoarseness (GRB scale). We used a 4-point grading system (0—normal, 1—slight, 2—moderate, 3—extreme) in this study to quantify perceptual assessment of hoarseness. The intrarater reliability was 0.91 for G (grade), 0.79 for R (roughness), and 0.83 for B (breathiness). The interrater reliability (50 voices, time interval 1–2 months) was 0.85 for G, 0.74 for R, and 0.81 for B factors, respectively.

**Acoustic analysis**

Segments of 2-second duration of the sustained phonation on habitual pitch and loudness of vowel /a:/ were recorded in a soundproof 2.5 × 1.5-m room through the D60S Dynamic Vocal (AKG Acoustics, Vienna, Austria) microphone placed at 10.0-cm distance from the mouth coupled to a digitized Sony Mini Disc Recorder MDS-101 (Sony Corporation, Tokyo, Japan). Three separate voice samples were analyzed with Tiger Electronics (Seattle, WA) Dr. Speech software (Voice Assessment, Version 3.0). Acoustic voice-signal parameters measured included fundamental frequency ($F_0$), perturbation of frequency (percent of jitter) and amplitude (percent of shimmer), and turbulent noise as normalized noise energy (NNE) at 44,100-Hz sampling frequency. The very first parts of the phonation sample (0.25 s) were cut off, and measurements were performed during the subsequent 2.0 seconds, thus minimizing variability caused by sampling errors. The remaining parts of the sustained vowel (a) were discarded, which ensured that the beginning and the end of voicing did not influence the final result. Acoustic voice signal data were measured for fundamental frequency ($F_0$), percent of jitter, and shimmer and for NNE at 44,100 Hz sampling frequency. A mean of three consistent samples was calculated for each parameter.

**Phonetogram (PG)**

PGs were recorded according to the recommendations of the Union of European Phoniatricians. Tiger Electronics Dr. Speech software (Phonetogram) and multidirectional microphone D60S Dynamic Vocal (AKG Acoustics) were employed for the registration and assessment of the PG. Seven parameters of PG were evaluated: (1) pitch range (PR) in semitones (smt.), (2) maximum frequency (F-max) in Hertz, (3) minimum frequency (F-min) in Hertz, (4) intensity range (IR) in decibels (A), (5) maximum intensity (I-max) in decibels (A), (6) minimum intensity (I-min) in decibels (A), and (7) PG area (S) in decibels (A)-smt. PGs of men and women were evaluated separately to avoid gender-dependent influence on the following parameters: PR, F-max, F-min, and S.

**Aerodynamics**

Maximum phonation time (MPT) as a simplest aerodynamic parameter of phonation was measured in seconds (sec) for the longest period for phonation of vowel /a:/.

**Subjective self-evaluation of voice**

Subjective self-evaluation of voice by the patient was performed with a 100-mm-long VAS. The patients were asked to rate hoarseness on this scale, ranging from 0 (“no hoarseness”) to 100 (“severe hoarseness”). An impact of possible voice
Table 1. Comparison of the Video Laryngostroboscopic (VLS) Parameters in the RL Patients’ and the Control Groups

<table>
<thead>
<tr>
<th>VLS Parameters</th>
<th>Patients (N = 108)</th>
<th>Controls (N = 90)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incomplete glottal closure</td>
<td>18.1 ± 15.6</td>
<td>1.6 ± 3.8</td>
<td>0.000*</td>
</tr>
<tr>
<td>Irregularity of VF vibrations</td>
<td>9.9 ± 13.9</td>
<td>0.3 ± 1.3</td>
<td>0.000*</td>
</tr>
<tr>
<td>Asymmetry of VF vibrations</td>
<td>3.8 ± 9.3</td>
<td>0.2 ± 1.9</td>
<td>0.000*</td>
</tr>
<tr>
<td>Reduced mucosal wave</td>
<td>7.8 ± 11.6</td>
<td>0.4 ± 1.7</td>
<td>0.000*</td>
</tr>
</tbody>
</table>

*Statistically significant difference.
Abbreviations: VF, vocal folds.

Perceptual voice evaluation

According to the results of perceptual voice assessment on the GRB scale, a slight hoarseness (G1) was detected in 21 (52.5%) men and 45 (66.2%) women, a moderate hoarseness (G2) in 12 (30.0%) men and 21 (30.9%) women, and extreme hoarseness (G3) in 7 (17.5%) men and 2 (2.9%) women, respectively. Consequently, a slight hoarseness (G1) was reliably prevailing in the RL patients group. In the male subgroup, a slight roughness (R1) revealed in 25 patients (62.5%) was prevailing. However, in the female subgroup, slight roughness (R1) was detected in 39 patients (57.4%) and slight breathiness (B1) in 36 patients (52.9%), respectively.

Acoustic analysis

Results of acoustic voice assessment are presented in Table 2. As follows from the data obtained, the mean jitter, shimmer, and NNE in the RL patients group were significantly (α < 0.05, β ≤ 0.14) increased to compare with the corresponding mean values of the control group. However, there was no gender-dependent difference of acoustic voice parameters (P > 0.05) in the RL patients group. The mean fundamental frequency (F0) of the voice of RL patients and healthy persons did not differ statistically significantly.

PG

Because of dependence on gender, the PG parameters of men and women in the RL patients and
TABLE 2. Comparison of the Means of Acoustic Voice Parameters in the RL Patients and Control Groups

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients (N = 40)</td>
<td>Controls (N = 36)</td>
</tr>
<tr>
<td>Jitter (%)</td>
<td>0.25 ± 0.13*</td>
<td>0.18 ± 0.04</td>
</tr>
<tr>
<td>Shimmer (%)</td>
<td>2.39 ± 1.31*</td>
<td>1.57 ± 0.59</td>
</tr>
<tr>
<td>NNE (dB)</td>
<td>−7.89 ± 3.69*</td>
<td>−13.40 ± 2.18</td>
</tr>
<tr>
<td>F₀ (Hz)</td>
<td>142.87 ± 49.99</td>
<td>126.99 ± 26.67</td>
</tr>
</tbody>
</table>

*Statistically significant (P < 0.002) difference between patients and control groups.

Abbreviations: NNE, normalized noise energy; F₀, fundamental frequency.

control groups were compared separately. After statistical analysis, it was established that in the RL patients group, the mean PR, F-max, and S were statistically significantly (α < 0.05, β < 0.05) reduced for both male and female patients as compared with the controls (Table 3).

The PR of the male patients suffering from RL was found to be statistically significantly (α < 0.05, β ≤ 0.04) reduced on the average by 11 smt. that of female patients—by 4 smt. to compare with normals (Figure 1). Consequently, the PR for the male patients decreased in average from F (86.5 Hz) to F# (93.6 Hz) for low frequencies and from f² (683.4 Hz) to g¹# (411.2 Hz) for high frequencies, and for the female patients from e (160.4 Hz) to d# (155.9 Hz) and from a² (891.1 Hz) to f²# (726.7 Hz), respectively.

Aerodynamics

Mean values of MPT in the RL patients and control groups are presented in Table 3. It was established that the mean MPT of RL patients (both men and women) was significantly (α < 0.05, β ≤ 0.01) shorter than that of healthy persons of the control group. Consequently, the mean MPT of male patients was shortened by 8 sec and of female patients by 7 sec to compare with healthy persons.

Subjective self-evaluation

In the RL patients group, the hoarseness evaluated by the patient on the VAS (H_VAS) ranged from 0 to 100 points, mean value was 52.89 (SD 30.76), and median was 52.5 points. In the control group, H_VAS ranged from 0 to 24 points, mean value was 1.01 (SD 4.33), and median was 0 points. The mean values of H_VAS of the patients group were statistically significantly higher (α < 0.01, β < 0.01) than those of the persons of the control group.

VHI

The total VHI scores of the subjects without voice problems and judged by a trained speech pathologist to have a normal voice quality were collected from a control group of 56 men and 69 women. The mean age of the group was 37.0 years (SD 13.5, range 22–72 years). The mean total VHI score in this control group was 5.39 years (SD 7.6, range 0–34 years) with a confidence interval of 1.33. Therefore, the total VHI scores of 4.06 or lower were considered as normal.

A statistically significant difference (P < 0.001) of the mean total VHI score between the patients and control groups was revealed, with the patients group having the worst score. The mean of VHI of the group of RL patients was 28.14 ± 21.91 points, and that of the control group was 4.66 ± 6.31 points (Figure 2). In the patients group, the mean scores were significantly (α < 0.01, β < 0.01) higher for the physical (P) domain compared with the emotional (E) and functional (F) domains. This result suggests that the perception of laryngeal discomfort or negative voice output characteristics by a person (P domain) were much more affected in the RL patients group.

Correlation analysis

Statistical analysis did not show a high correlation of the measured instrumental voice parameters with either subjective assessment of hoarseness or VHI (Table 4). However, a slight to strong (r = 0.27–0.78) significant correlation was observed within
TABLE 3. Comparison of the Means of Phonetograms’ Parameters in the RL Patients’ and Control Groups

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Male Patients (N = 40)</th>
<th>Male Controls (N = 36)</th>
<th>Female Patients (N = 68)</th>
<th>Female Controls (N = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x ± SD</td>
<td>x ± SD</td>
<td>x ± SD</td>
<td>x ± SD</td>
</tr>
<tr>
<td>PR (smt.)</td>
<td>24.6 ± 6.7*</td>
<td>35.7 ± 4.9</td>
<td>26.1 ± 6.5*</td>
<td>30.2 ± 4.1</td>
</tr>
<tr>
<td>F-max (Hz)</td>
<td>411.2 ± 156.2*</td>
<td>683.4 ± 142.5</td>
<td>726.7 ± 212.7*</td>
<td>891.1 ± 157.1</td>
</tr>
<tr>
<td>F-min (Hz)</td>
<td>93.6 ± 17.3</td>
<td>86.5 ± 15.6</td>
<td>155.9 ± 26.6</td>
<td>160.4 ± 48.4</td>
</tr>
<tr>
<td>IR (dB[A])</td>
<td>38.8 ± 9.7*</td>
<td>48.2 ± 6.3</td>
<td>39.3 ± 8.0</td>
<td>42.8 ± 5.8</td>
</tr>
<tr>
<td>I-max (dB[A])</td>
<td>88.4 ± 9.9*</td>
<td>94.7 ± 6.7</td>
<td>88.5 ± 9.5</td>
<td>91.2 ± 7.1</td>
</tr>
<tr>
<td>I-min (dB[A])</td>
<td>49.6 ± 4.5*</td>
<td>46.5 ± 1.2</td>
<td>49.2 ± 4.8</td>
<td>48.4 ± 4.5</td>
</tr>
<tr>
<td>S (smt.dB[A])</td>
<td>514.02 ± 27.4*</td>
<td>915.1 ± 175.1</td>
<td>524.8 ± 179.9*</td>
<td>725.7 ± 160.4</td>
</tr>
<tr>
<td>MPT (sec)</td>
<td>17.8 ± 6.8*</td>
<td>26.1 ± 4.9</td>
<td>14.4 ± 5*</td>
<td>21.1 ± 4.2</td>
</tr>
</tbody>
</table>

*Statistically significant (α < 0.002, β < 0.05) difference between patients’ and control groups.

**Abbreviations:** PR, pitch range; F-max, maximum frequency; F-min, minimum frequency; IR, intensity range; I-max, maximum intensity; I-min, minimum intensity; S, phonetogram area; MPT, maximum phonation time.

We used a multidimensional set of perceptual, subjective, and objective acoustic measurements of voice quality and phonation capabilities in this study. According to the data obtained, a statistically significant difference between the RL patients group and the control group was found in all voice parameters measured, with the patients having worse results. These data documented significant deterioration of voice quality and phonation capabilities in the RL patients group.

However, throughout, systematic studies related to perceptual and instrumental voice assessment in RL patients have been lacking to date and the data of the studies were controversial.

Few studies revealed an increase of perceptual rating of hoarseness in the RL patients group.7,22,23 However, to the best of our knowledge, only one study compared voices of RL patients and control groups and found significantly increased shimmer in the patients group.5 Some studies employed phonotography to assess functional outcomes of antireflux treatment and did not find significant changes after 4 weeks of antireflux therapy neither in PG parameters nor in acoustic voice parameters.6,22 On the contrary, other studies demonstrated significantly decreased jitter, shimmer, and signal-to-noise ratio as a clinical response to antireflux treatment.9,23

A slight hoarseness (G1) revealed according to the GRB scale was prevailing in the RL patients group.7,22,23 We used a multidimensional set of perceptual, subjective, and objective acoustic measurements of voice quality and phonation capabilities in this study. According to the data obtained, a statistically significant difference between the RL patients group and the control group was found in all voice parameters measured, with the patients having worse results. These data documented significant deterioration of voice quality and phonation capabilities in the RL patients group.

As the ratio of subjects in the patients and control groups studied did not differ significantly with respect to age, vocal training and usage, and smoking habits (P < 0.05), and the patients with other possible causes of permanent hoarseness were excluded from the study, one can consider that deterioration of voice quality and phonation capabilities in the RL patients group was related to RL.

Instrumental voice parameters and moderate (r = 0.57) correlation within parameters of subjective (VHI, H_V AS) voice assessment.

**DISCUSSION**

A variety of experimental and clinical studies has confirmed a negative irritative effect of pathological gastroesophageal reflux on the larynx.1,3,8,20 Several studies of the last decade have focused not only on analysis of the morphological and laryngoscopic features of RL, but also they have attempted to assess, analyze, and describe the influence of gastroesophageal reflux on phonatory function of the larynx and voice changes related to RL.4,6,7,9,21 However, throughout, systematic studies related to perceptual and instrumental voice assessment in RL patients have been lacking to date and the data of the studies were controversial.

We used a multidimensional set of perceptual, subjective, and objective acoustic measurements of voice quality and phonation capabilities in this study. According to the data obtained, a statistically significant difference between the RL patients group and the control group was found in all voice parameters measured, with the patients having worse results. These data documented significant deterioration of voice quality and phonation capabilities in the RL patients group.
(jitter) and amplitude (shimmer) as well as an increase of glottal noise (NNE) in the group of RL patients to compare with the control group. Quantitative assessment of PGs and MPT documented significant restriction of phonation capabilities in the RL patients group. The mean values of the parameters of PGs, which reflect voice intensity and pitch (PR, F-max, S) as well as the mean of MPT, were significantly lower in the group of RL patients as compared with the controls.

The most possible negative factors interfering with phonation and altering periodicity of vibration cycle and glottic closure could be oedema of the vibratory margin of the vocal folds, which is caused by potentially noxious materials such as gastric acid, pepsin, and pancreatic enzymes irritation.\(^4,6,21\)

Our data support these presumptions as slight oedema of vocal folds and incomplete glottal closure were the most frequent VLS findings in the RL patients group that could deteriorate the pattern of phonation.

Thus, the results obtained in this study confirm that the voice disturbances of RL patients can be objectively described by acoustic instrumental analysis.

Although there was often generally good agreement of perceptual data with instrumental measurements, statistical analysis did not show a high correlation of the measured acoustic parameters with either perceptual assessments of overall voice abnormality or subjective assessments of a particular voice quality. This result supports the current consensus that voice is a multidimensional phenomenon and cannot be described by one parameter, but

**FIGURE 1.** Reduction of the mean PR of RL patients. Straight line—controls; dotted line—patients.

**FIGURE 2.** Means and SEM of VHI in the RL patients and the control groups. F—functional domain, E—emotional domain, P—physical domain, SEM—standard error of means. *P < .001 patients versus controls, and P domain versus F and E domains.
it should be investigated by means of voice quality and voice function analyses.\textsuperscript{16,24}

Moreover, voice problems in daily life comprise more than just the quality of voice. The VHI as a standardized structured and validated instrument gives additional insight into voice-related problems in the daily life of the RL patients than single state-ment evaluation of voice quality. Therefore, VHI was applied in this study to evaluate the level of handicap of a person resulting from a voice disorder.\textsuperscript{19}

The results of this study demonstrated a statistically significant difference in mean VHI scores between RL patients and control groups, with the patients group having the worst mean score (Figure 2). Furthermore, the mean scores were significantly higher for the physical (P) domain, and the emotional (E) and functional (F) domains did not demonstrate a statistically significant difference between any subgroups. This result suggests that the impact of the voice of a person on the perception of laryngeal discomfort or negative voice output characteristics (P domain) is much more affected in the tested RL patients group. On the other hand, these prevailing scores of P domain of VHI possibly are influenced by the following typical complaints of the RL patients group: hoarseness and throat itching and clearing.

**CONCLUSIONS**

This study demonstrated that perceptual and subjective evaluation as well as objective measures of voice parameters document deteriorated voice quality caused by RL. Deteriorated phonation pattern and voice-related problems lead to negative impact on everyday life and professional activities, as reflected in significantly increased values of VHI. On the other hand, our results suggest that one voice parameter cannot completely describe changes in voice quality cad by RL. In light of these data, multidimensional voice assessment including both subjective and instrumental voice measurements is advocated.

Treatment of RL patients along with special medication and lifestyle modifications should involve care of the voice. Consequently, assessment of the results of the treatment of RL patients should include monitoring of voice quality and VHI.

**REFERENCES**