ORIGINAL ARTICLE

Acoustic Changes of the Voice as Signs of Vocal Fatigue in Radio Broadcasters: Preliminary Findings

Marco Guzmán, a, * María Celina Malebrán, b Paulina Zavala, c Patricio Saldivar, d Daniel Muñoz e

a Escuela de Fonoaudiología, Universidad de Chile, Santiago, Chile
b Escuela de Fonoaudiología, Universidad del Mar y Universidad de Valparaíso, Valparaíso, Chile
c Escuela de Fonoaudiología, Universidad del Mar, Relhaca, Valparaíso, Chile
d Departamento de Voz y Carrera de actuación teatral de la Universidad de Valparaíso, Valparaíso, Chile
e Facultad de Medicina, Universidad de Chile, Santiago, Chile

Received 4 September 2012; accepted 15 November 2012

KEYWORDS
Vocal fatigue; Broadcasters; Acoustic analysis; Vocal loading; Cepstrum

Abstract

Introduction and objective: Vocal fatigue is one of the most common voice symptoms. It usually refers to the sensation of vocal tiredness after a long period of speaking or singing. The purpose of this study was to compare the acoustic characteristics of the voice before and after a long period of voice use in a group of radio broadcasters.

Methods: Eight radio broadcasters with normal voices were assessed. We used cepstrum, energy ratio, noise to harmonic ratio and soft phonation index as acoustic variables to assess the possible pre–post vocal loading changes objectively.

Results: There were no statistically significant pre–post differences in any of the acoustic parameters. Although cepstrum at high pitch did not show a significant difference, it obtained the greatest difference among the acoustic variables.

Conclusions: The acoustic measurements used in the present study might not be sensitive enough or appropriate for detecting vocal changes after a long period of voice use, whether in reading (as reported in previous research) or speaking tasks. Moreover, a longer period of vocal loading would eventually reveal more evident and consistent acoustic voice changes.

© 2012 Elsevier España, S.L. All rights reserved.

PALABRAS CLAVE
Fatiga vocal; Locutores; Análisis acústico;

Cambios acústicos de la voz como signos de fatiga vocal en locutores de radio: resultados preliminares

Resumen

Introducción y objetivos: En la clínica de voz, unos de los síntomas más frecuentes es la fatiga vocal, la cual usualmente se refiere a la sensación de cansancio de la voz posterior al uso


* Corresponding author.
E-mail address: guzmanvoz@gmail.com (M. Guzmán).
Introduction

Vocal fatigue is one of the most common voice symptoms.1,2 It usually refers to a feeling of tiredness of the voice after its prolonged use during speech tasks.3,4 Subjects presenting such symptoms usually require more effort to continue producing phonation5 and usually present changes in vocal quality (vocal timbre), intensity, fundamental frequency (F0)6 and a feeling of laryngeal discomfort, including pain, foreign body sensation, inflammation, muscle tension in the cervical/pharyngeal area and laryngeal dryness.4 Some voice professionals often report that vocal fatigue is associated to a decrease in vocal projection, tonal range, dynamic range, an increase in the effort required to produce voice and a sensation of laryngeal and pharyngeal constriction.7-9 Several studies agree that the prevalence of vocal fatigue symptoms among certain voice professionals is associated with high rates of vocal disorders.6,10-13 From these observations we can infer that vocal fatigue in certain groups represents a potential for vocal fold lesions.

Precise methods for assessing vocal fatigue are difficult to implement because they require a series of neuromuscular and biomechanical data that can be only obtained through invasive explorations.14 Vocal fatigue has mostly been studied by directed induction of voice fatigue in voice professionals. These studies have employed prolonged periods of reading aloud in order to cause fatigue in the voice of subjects. In most studies, data collection was carried out by acoustic and aerodynamic measurements, before and after vocal loading.4,15-17

Acoustic measurements have been widely used as a method for the assessment of vocal fatigue. However, the results are not conclusive.4,5,15,18,19 For example, jitter and shimmer values have been observed to significantly increase in some subjects following vocal fatigue experimentation, whilst other subjects have not manifested such changes.4,5,18 Several studies have reported that F0 increases with vocal fatigue,4,19,20-22 whilst other studies have suggested the opposite.1

Stroboscopic and aerodynamic measurements have also been used to document vocal fatigue in physiological terms. In general, an increase in the transglottic flow rate, a decrease in the maximum phonation time and incomplete glottal closure have all been associated with vocal fatigue symptoms.4,5,13,20 Netsell indicated that subjects with vocal fatigue may present a breathy voice (also known as murmured voice, soughing or susurration) due to excessive loss of air through the glottis during phonation, or else a tense voice quality due to the compensatory effort made in order to reduce excess transglottic flow.23,24

The pressure threshold of phonation (minimum subglottic pressure required to start the oscillation of the vocal folds) has been investigated as a possible objective marker of presumed changes in the viscosity of vocal tissue during vocal fatigue (biomechanical changes).25-28 Other related studies27-30 have provided evidence suggesting that the hydration state is important in regulating the viscosity of vocal tissues, and that this viscosity is directly proportional to the pressure threshold of phonation.

Most studies of vocal fatigue have been conducted on populations of teachers,2,6,10,31,32 with no studies being conducted among other professional voice users, such as radio broadcasters. Although the latter group of professionals is not among the top sufferers of vocal involvement, as is the case with teachers,6,11,33,34 clinical observations show that vocal fatigue is a common complaints among radio broadcasters. Vocal discomfort reported by these subjects may have various origins, including vocal overuse, habits of vocal abuse and misuse and lack of proper vocal training.35 Malebrán and Saldivar reported that over half of a group of radio broadcasters examined presented signs of vocal fatigue and a lack of adequate vocal technique.36

The purpose of the present study is to compare the vocal acoustic characteristics before and after a prolonged period of vocal use among a population of radio broadcasters. Two of the acoustic markers used in this study have not been previously used.

Method

Subjects

The sample consisted of eight broadcasters aged between 34 and 61 years, with a mean value of 44 years, who had perceptually normal voices. Two of the subjects included in the study were female and six were male. This group of
participants corresponded to the total population of radio hosts working at the radio station selected for data collection in this research. None had a history of vocal treatment prior to the completion of this study and none reported any degree of hearing impairment. The evaluations were carried out with the understanding of all participants and after signing informed consent documents. Importantly, this group of participants was selected taking into account that their vocal complaints were always related to signs of vocal abuse and fatigue after a working shift of speech (1 h).

Recording

The voice of each subject was recorded before and after a session of vocal loading through speech. The recording was done using an AKG® Perception 120 model condenser microphone. The microphone was placed on a pedestal and located 10 cm away from the mouth of each subject. Subjects remained standing during recording. We used an M-Audio® v.1.03 sound interface system for preamplification and digitalisation. The recordings were made with a sampling frequency of 44.1 kHz and 16-bit quantisation, within a soundproofed recording studio. Capturing and recording of the voice signals was done using the Praat® programme. We used a 8810 CE-IEC 651 type II sonometer to control the sound pressure level.

Phonatory Tasks

Participants were asked to perform the following phonatory tasks during the recording process:

- The vowel /a/ sustained in a comfortable tone for each subject (medium range).
- The vowel /a/ sustained using a F0 1 octave higher than the comfortable tone used previously.
- The vowel /a/ sustained using a F0 1 major third below the comfortable tone used previously.
- Reading a phonetically balanced text of 104 words for 1 min using an average and comfortable intensity for each subject.

To control the fundamental frequency used during sustained phonation of the vowel /a/ (three different tones), one of the authors matched the medium spoken tone of the first vowel /a/ for each subject (comfortable tone) to a tone in an electronic keyboard. Subsequently, based on this tone, the experimenter presented a sound 1 octave above the medium spoken tone and then a sound 1 major third below the original tone. Each participant was asked to reproduce each of the tones presented using the vowel /a/. The recording was started after subjects demonstrated the ability to perform the phonatory task in the tone delivered by the examiner. Although this was the intended objective for all subjects tested, in some cases a lack of musical ear did not allow an exact match of the requested tones.

Acoustic Analysis

The acoustic analysis of all voice samples taken before and after the speech task included: cepstrum (dual fast Fourier transform), noise-harmonic ratio (NHR), soft phonation index (SPI) and energy ratio (difference between the highest peak at 0–2000 Hz and the highest peak at 2000–4000 Hz). The latter measurement was used to evaluate changes in the spectral slope. Two of these measurements (cepstrum and energy ratio) have not been used previously as potential acoustic markers of vocal fatigue.

We used the software package Multi-Speech Main Programme®, model 3700, v.3.2.0 (Kay Elemetrics, Lincoln Park, NJ, USA) for the analysis of cepstrum, SPI and NHR. Cepstrum analysis was performed with a 1024 window sample and a Hamming-type window. Energy ratio values were obtained with the software package Wavesurfer® v.1.8.5 (Jonas Beskow and Kåre Sjölander, Sweden) through long-term average spectrum (LTAS) analysis window.

Statistical Analysis

Data were analysed using the statistical software package Stata® 12.1 (StataCorp 2011, College Station, TX: Stata-Corp LP, USA). The acoustic parameters of the voice were expressed as mean and standard deviation. We used the Student t test for paired data when comparing pre- and post-loading values. A value of P<.05 was considered as statistically significant for all hypothesis contrasts.

Results

The results (mean values pre- and post-speech loading, standard deviation and value of P) for the acoustic analysis of cepstrum parameters at different frequencies, long-term average spectrum (LTAS) and noise measurements are presented in Table 1. None of the evaluated parameters showed a statistically significant difference between pre- and post-speech loading values. However, despite not being significant, the parameter which obtained the greatest pre- and post-speech loading difference was the cepstrum during phonation of acute F0 (7.77 ± 9.68) (P=.0574). Fig. 1 shows the effect of the vocal load period on cepstrum values at the various F0 studied in each subject individually. This figure shows that the cepstrum value which obtained a more systematic decrease was in samples produced with high F0. In total, seven of the eight participants presented a decrease in the cepstrum values when comparing pre- and post-loading emissions. Figs. 2 and 3 show cepstrum peaks before and after loading, respectively, obtained for the acute vowel /a/ steadily for subject number 3. Fig. 4 shows the effect of treatment on the energy ratio (0–2/2–4 ratio), HNR and SPI for each subject individually. The SPI was the variable which most systematically showed the pre- and post-loading difference. Specifically, five of the eight subjects obtained an increase in the value of SPI. Additionally, it is important to note that although there were no significant differences, the SPI showed the second lowest P-value in the analysis of statistical significance (P=.1787).
Acoustic Changes of the Voice as Signs of Vocal Fatigue in Radio Broadcasters: Preliminary Findings

Table 1 Pre- and Post-Loading Mean Values, Standard Deviation and Value of P for the Acoustic Analysis of the Cepstrum Parameters at Various Frequencies, LTAS and Noise Measurements.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Pre</th>
<th>Post</th>
<th>Difference</th>
<th>Value of P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cepstrum (acute)</td>
<td>36.32 ± 11.93</td>
<td>28.54 ± 4.70</td>
<td>7.77 ± 9.68</td>
<td>.0574</td>
</tr>
<tr>
<td>Cepstrum (medium)</td>
<td>30.53 ± 13.85</td>
<td>32.41 ± 12.79</td>
<td>1.88 ± 12.41</td>
<td>.6807</td>
</tr>
<tr>
<td>Cepstrum (low)</td>
<td>19.05 ± 7.24</td>
<td>21.05 ± 8.28</td>
<td>2.007 ± 8.00</td>
<td>.5009</td>
</tr>
<tr>
<td>0<del>2/2</del>4 ratio</td>
<td>22.28 ± 4.86</td>
<td>23.77 ± 4.56</td>
<td>1.48 ± 5.86</td>
<td>.4962</td>
</tr>
<tr>
<td>NHR</td>
<td>0.22 ± 0.21</td>
<td>0.13 ± 0.02</td>
<td>0.09 ± 0.22</td>
<td>.2856</td>
</tr>
<tr>
<td>SPI</td>
<td>8.91 ± 5.49</td>
<td>10.67 ± 5.23</td>
<td>1.76 ± 3.34</td>
<td>.1787</td>
</tr>
</tbody>
</table>

LTAS: long-term average spectrum; NHR: noise-harmonic ratio; SPI: soft phonation index.

Discussion

The results of this study reveal that none of the acoustic parameters showed a statistically significant pre- and post-speech loading difference. Despite not being significant, the difference in cepstrum values before (36.33) and after (28.55) the speech task for acute F0 production were the most notable among all the variables. In clinical terms, this decrease in the cepstrum value implies a more unstable voice, with less energy and less periodic signal. The

Figure 1  Effect of the speech loading period on cepstrum values at the different F0 used for each subject in the study individually.

Figure 2  Cepstrum peak prior to the period of speech loading obtained during a sustained, acute vowel /a/.
ceprostrum is defined as a dual fast Fourier transform (FFT)\textsuperscript{38,39}. Voices characterised by a well-defined harmonic structure display high cepstrum peaks (high values). By contrast, breathy or hoarse voices present a poorly defined harmonic structure and, therefore, their cepstrum peak is low (low values).

The cepstrum was chosen as one of the variables for this study because it has not previously been used as a potential acoustic marker of vocal fatigue. Additionally, previous studies have reported that the cepstrum value is the best predictor of dysphonia, compared with the rates of alteration and glottic noise index\textsuperscript{40-43}. In addition, the cepstrum peak has shown a high correlation with the degree of severity of vocal alterations\textsuperscript{44-48}. As an indicator of vocal fatigue or deterioration of the quality of the voice, cepstrum values would show a decrease. In the present study, the difference between pre- and post-loading values was only clinically evident for sustained vowels in high F0 (7.77 ± 9.68). The differences for low and medium F0 were 2.007 ± 8.00 and 1.88 ± 12.41, respectively.

The 0–4/2–4 kHz spectral difference (energy ratio) revealed an increase from 22.28 dB (pre-loading) to 23.77 dB (post-loading). The purpose of this spectral measurement was to calculate the energy difference between the highest spectral peak in the 0–2 kHz range and the highest peak in the 2–4 kHz range. This parameter has been extensively studied among professional singers as an acoustic marker of vocal quality, that is, how resonant a voice is perceived to be. However, this variable has not been used in previous studies as a possible acoustic marker of vocal fatigue.

Although this change (increase of this spectral difference) did not reach statistical significance, from the clinical point of view it suggested a change in the spectral energy distribution following a long period of voice use. The highest values of this variable after the period of vocal use represented an energy decrease in the higher harmonics of the spectrum. In other words, there was a greater difference...
between the energy of low harmonics and the energy of high harmonics (higher spectral slope).

In relation to this spectral change, it is also interesting to note that the SPI experienced an increase after the voice load period (from 8.91 to 10.67). This parameter is also regarded as a measure of the spectral slope. The SPI value is obtained from the ratio of the spectral harmonic energy at low frequencies (70–1600 Hz) and the harmonic energy at high frequencies (1600–4500 Hz). Therefore, from the clinical point of view, an increase in the value of SPI means that the energy of higher harmonics has diminished. That is, higher SPI values are obtained for greater spectral slopes. In our study, we observed an interesting relationship between SPI and energy ratio values after a prolonged use of the voice, could be determined by the vibratory characteristics of the vocal folds. There is a systematic relationship between the rate of closure of the vocal folds and the spectral slope: the faster the closure of the vocal folds (faster return of the folds to the midline), the lower the spectral slope (higher energy in high harmonics). This can be seen in Figs. 5 and 6, which show the spectral slope of subject number 1 for the pre- and post-speech loading samples, respectively. This particular subject showed an increase of the spectral slope (less harmonic energy in the high spectral area).

This change in spectral slope is not only associated with the closing speed of the vocal folds, but also with the contact ratio, that is, the time period during which the vocal folds are closed compared to the total glottic cycle time. From this fact, and remembering that, in our study, the spectral slope showed a decrease, it is possible to speculate that perhaps, after the period of vocal use...
undergone by the broadcasters, the vibration of their vocal folds changed towards a decrease in contact ratio and in closure speed.

From the clinical point of view, it is interesting to note that the average NHR of the group showed a decrease, rather than the expected increase. Furthermore, individually, five of the eight subjects presented a lower NHR after vocal loading. An increase in NHR should occur if there were an increase in glottal noise energy and/or a reduction in harmonic energy. Both cases would be linked to a decrease in fold closure caused by vocal fatigue, as reported by previous studies where transglottic flow velocity increased and maximum phonation time decreased.\(^{4,5,10}\) Furthermore, it is possible that the decrease in NHR values was the result of increased fold closure, as a physiological compensation mechanism following a long period of voice use, as has also been reported by previous studies.\(^ {23}\)

The absence of significant differences between pre- and post-speech loading samples in the present study may be attributed to a number of reasons. The first is the period of time of continuous vocal use. In our study, we used a period of 1 h, as this was the duration of most of the programmes presented by the radio broadcasters in the study. Due to the programming schedule of the radio station, it was not possible to extend these time periods. It is possible that, with prolonged periods of vocal use, acoustic changes may have been more evident. With respect to the time variable, Scherer et al.\(^ {18}\) showed that study subjects showed no significant changes in the jitter and shimmer values after 1 h of continuous reading. Similar results have been reported by more recent studies.\(^ {51}\) However, contrary to these reports, some studies have shown that periods less than 1 h are sufficient to cause acoustic changes. Niebudek-Bogusz et al.\(^ {52}\) observed changes in the acoustic parameters after 30 min of vocal practice, whilst Laukkanen et al.\(^ {53}\) reported the same after just 5 min.

Another possible explanation for the lack of significant pre- and post-loading differences is the type of task employed during the vocal load period. Vocal fatigue has mostly been studied through prolonged and constant periods of reading out loud.\(^ {4,15-17}\) However, it is possible that in our study we used a task (speech) which was less fatiguing. During the time period of radio speech work there are pauses and moments when a radio broadcaster makes a less intense use of the voice. It is possible that the two elements have led to a decrease in overall voice load, compared with an equal period of continuous reading.

It is also possible that the acoustic markers used as an objective measurement of a possible change in voice due to vocal fatigue were not suitable or sensitive enough to detect such changes. Although various studies have used acoustic measurements\(^ {4,5,15,18-22}\) so far, the search for an objective acoustic marker of vocal fatigue has not been successful.

**Conclusion**

Acoustic measurements are not adequate or sensitive enough to detect vocal changes after a prolonged period of voice use, either through reading, as in previous studies, or through speech tasks. For this reason, it would be interesting to replicate this study whilst taking into account other objective variables, such as aerodynamic measurements (transglottic flow, subglottic pressure and glottal resistance) and electroglottographic measurements. Moreover, a longer period of voice loading than that used in this study would reveal more evident and consistent acoustic changes.

**Conflict of Interests**

The authors have no conflict of interests to declare.
Acoustic Changes of the Voice as Signs of Vocal Fatigue in Radio Broadcasters: Preliminary Findings


30. Verdolini-Marston K, Titze IR, Druker DG. Changes in phona


34. Roy N, Merrill RM, Thibeault S, Parsa RA, Gray SD, Smith EM. Prevalence of voice disorders in teachers and the general popu

35. Landazuri E. Prevención vocal: una responsabilidad fonoa

cano de Voz Cantada y Habla. 2010.


40. Heman-Ackah YD, Heuer RJ, Michael DD. Cepstral peak promi


45. Awan SN, Roy N. Toward the development of an objective index of dysphonia severity: a four-factor acoustic model. Clin Linguis
t Phon. 2006;20:35-49.


