Chapter 12
Audio-visual analysis of vocal tremor derived from acoustics, laryngovideostroboscopy & HSDI

Julie Barkmeier-Kraemer, Krzysztof Izdebski & Yuling Yan

Abstract

Acoustical, perceptual (sonographic), visual (endoscopic), or physiologic (EMG or EGG) methods are used to support clinical diagnosis of vocal tremor (VoT). The primary auditory test includes subjective clinically graded auditory impression of the generated voice while acoustics include various methods of signal analysis. Visual methods described in the literature encourage characterization of VoT using nasoendoscopy to identify whether VoT is present or absent within specific parts of vocal tract subcomponents or is a part of a more generalized tremor. Because chemical denervation of laryngeal musculature (Botox® injection) is the current VoT treatment approach, making detailed VoT analysis is imperative for determining which laryngeal muscles are involved, hence which muscle(s) are to be targeted for denervation. Consequently, we evaluated gains from a non-invasive VoT analysis using rigid stroboscopic (LVS) and from a HSDI evaluation protocol combined with acoustic patterns of VoT to improve identification of specific laryngeal musculature generating VoT.

Keywords: vocal tremor, stroboscopy, HSDI, acoustics, voice evaluation, visual perception, phonatory function studies, Nyquist plots

Introduction

Vocal output from individuals suffering from VoT is characterized perceptually by a nearly rhythmic modulation of voice pitch and loudness [1-3]. VoT is classified as a neurologic movement disorder and although it may occur in isolation, VoT often occurs as part of such conditions as essential tremor, Parkinson’s Disease, dystonia, or cerebellar degeneration [4]. Current clinical methods for diagnosing this voice disorder rely heavily on perceptual evaluation [5-8] or visual imaging using nasoendoscopy [6-7, 9-11]. For example, Bové and colleagues [9] published a scoring protocol for evaluating VoT using nasoendoscopic visualization of the pharynx. Their scoring method requires detection of the presence/absence and severity of tremor in visible pharyngeal structures as a way of quantification of the severity of VoT. The total score is considered reflective of VoT severity and can be used for pre- and post-treatment comparisons.

Structures identified for rating in their scoring system include the soft palate, base of tongue, pharyngeal wall, and larynx (globally) with additional scoring of tremor affecting the supraglottal region and vocal folds (VF). Although this tool is helpful for evaluating the number of visible portions of the pharynx and larynx affected by tremor that contribute to VoT, the rank order of the individual contribution of the affected structures to the overall acoustic pattern of VoT remains unclear. Although their scoring system is not a completely valid method for scoring severity of VoT per se, they demonstrated that the
concept of using visual observation to evaluate severity of VoT during laryngeal imaging was a reliable method for evaluating VoT and in determining whether reduced signs of tremor occurred in structures injected with neurotoxin.

Currently, the most common medical treatment for vocal tremor is to inject a neurotoxin (i.e., Botox®) into laryngeal musculature [5-6, 9, 12-16]. These studies reported benefits to individuals with VoT after chemical denervation with drug varied dosage levels ranging from 0.625-15 μg ipsilaterally or bilaterally targeting thyroarytenoid (TA), cricothyroid (CT), or both. Bové and colleagues [9] also reported injection of 10 μg of Botox® into the posterior cricoarytenoid in two patients. Of interest are reports that many individuals with VoT reject chemical denervation treatment due to poor outcomes and persistence of such side effects as severe breathiness, difficulty breathing, and choking on food and liquids [12, 14-15].

Although nasoendoscopy and perceptual methods of evaluation are commonly used clinical methods for diagnosing VoT, electromyography (EMG) is the most accurate method for clearly determining which laryngeal muscles exhibit VoT. Several studies have used this method to describe tremor patterns within the larynx in individuals with VoT [17-18]. Although EMG is considered to be the standard referent for identifying musculature exhibiting tremor, it may not always be practical or available to individuals receiving VoT treatment. Thus, a standard and clinically simpler protocol for evaluating VoT in patients is vital for determining the laryngeal musculature affected by tremor [19]. Hence, we investigated here the value of using laryngovideostroboscopy (LVS) and the newest generation of visual technology, High Speed Digital Imaging (HSDI). When possible, visual findings were synchronized with other signals recorded simultaneously.

The purpose of this paper is to propose a structured visual imaging protocol using rigid transoral visualization to non-invasively determine the contribution of intrinsic laryngeal muscles to VoT. If demonstrated to have predictive value, such a clinical protocol should improve treatment outcomes by targeting those laryngeal muscles identified as exhibiting oscillation associated with VoT patterns.

**Methodology**

An LVS protocol was developed using a range of pitch and loudness tasks to elucidate the intrinsic laryngeal muscles most affected by tremor. The contribution of the intrinsic laryngeal muscles to phonatory patterns has been well described in the literature [19-26]. Known intrinsic laryngeal muscle contributions to laryngeal structure movement patterns can be used to develop a protocol to determine the muscle source(s) of laryngeal-based tremor (see Figure 1).

**Pitch modulation**

Two laryngeal muscles contribute antagonistically to increase and decrease pitch as the VF lengthen and shorten. Contraction of the TA muscle causes shortening of the VF resulting in increased effective vibratory mass and slower VF vibrations. Thus, VoT affecting the TA muscle would be expected to result in increased severity during production of lower pitches. In contrast, the CT muscle elongates the VF during contraction resulting in thinning of the VF and reduction in the effective vibratory mass and faster VF vibrations. Thus, oscillations affecting the CT muscle would result in increased severity of VoT during production of higher pitch productions.
Figure 1. Representation of intrinsic (TA & CT) and extrinsic (TH & ST) laryngeal muscle activity in a male subject producing range glissando. Recorded with bipolar hooked-wire electrodes and plotted with solid line against pitch (F0 in kHz) and vertical laryngeal position (VLP) in mm trajectories. Horizontal bar represents 100 ms time-frame.

Intensity modulation

Three laryngeal muscles contribute antagonistically to adduction and abduction of the VF position during phonation associated with laryngeal adjustments of intensity. The interarytenoid (IA) muscles and lateral cricoarytenoid (LCA) muscles contribute primarily to adduction of the VF. The IA muscles typically close the cartilaginous portion of the VF whereas the LCA is attributed with adducting the membranous VF. The posterior cricoarytenoid muscle (PCA) is the sole abductor of the VF.

During phonation, the VF are first approximated and set into vibration during exhalation. Maintaining this position of the VF during phonation requires a consistent balance between the VF abductor (i.e., PCA) and adductors of the VF (i.e., IA and LCA). During increased loudness, the VF must be held in position against increased tracheal pressures requiring increased contraction of the adductory muscles so that the VF do not get blown apart. Likewise, during softer phonations, the glottal width may be allowed to increase, associated with increased activation of the abductory muscle. Thus, during loud phonations, increased vocal tremor would be expected to occur in the muscles of adduction. During soft phonation, increased tremor would be associated with the PCA.

The larynx may also exhibit vertical oscillations associated with tremor involving the extrinsic laryngeal musculature such as the thyrohyoid (TH) or sternothyroid (ST). As the larynx moves in the semi-vertical plane, the length of the vocal tract is affected (shortens and lengthens) due to these oscillations. Tremor affecting the vertical larynx position (VLP) may result in modulation of intensity associated with shifting of the vocal tract formant frequencies (FF) as the length of the vocal tract modulates. It is crucial to note that the VLP movement may exceed 6 cm [19].
Audio-visual recordings

An individual diagnosed with essential vocal tremor was recorded using a 70° rigid scope and Kay Elemetrics Stroboscopic light source (Kay Elemetrics Computer Integrated System Model RLS 9100, Lincoln Park, NJ, USA). The individual was seated and positioned for transoral placement of the rigid scope while the examiner held the patient’s tongue with a gauze pad to prevent tongue retraction during recordings. The condenser microphone recorded voice during the examination and the audio recordings were later analyzed for rate and magnitude of fundamental frequency and intensity modulation during each examination task using PRAAT software (version 5.2.35, 2011) [27]. Recordings comprised the following voicing tasks executed in sustained phonation mode of the vowel, /i/ during:

1. Habitual pitch at comfortable loudness
2. High pitch
3. Low pitch
4. Comfortable pitch loudly
5. Comfortable pitch softly
6. Quiet breathing

Acoustic measures were completed by an expert naïve to the purpose of this study. Visual assessment of laryngeal movements was evaluated with relation to movements occurring during productions at the different pitches and loudness levels. Based on visual observations, hypotheses were formulated regarding affected intrinsic laryngeal musculature and resulting acoustic patterns in tremor modulation on fundamental frequency and intensity modulation magnitude, or extent. Visual observations and hypothesized contributions from specified laryngeal musculature were compared to acoustic measures. Figure 2 shows the expected motion of laryngeal structures during VoT.

In addition, HSDI (KayPENTAX Model 9710, NJ, USA) of simulated vocal tremor with vocal arrests were obtained from a second female. HSDI were then processed using KIPS program (KayPENTAX) and with custom driven programs (Vocalizer®) derived from visual and auditory signal [28-30].
Results

Clinical Vocal Tremor

Spectrographic analysis of the patient’s voice showed various magnitude and pattern of voice signal modulations. Though visual inspection shows discriminant features, auditory perception of these signals of VoT is not as categorical as the visual representation. This observation was derived from another study [6].

Visual observation of the larynx during the LVS protocol showed higher VoT magnitudes at high pitch and during loud phonations at all pitches. Laryngeal movements associated with worsened tremor included VF length change indicating CT involvement. In addition, the arytenoid cartilages exhibited tremor associated with abduction and adduction during loud phonations indicating involvement of laryngeal adductors (IA and LCA). Finally, VLP movements were also observed indicating tremor involvement of extrinsic and supplemental laryngeal muscles resulting in elevation and depression of the larynx during phonation.

Based on visual evaluation of laryngeal tremor patterns, adductory/abductory movement of the VF and VLP movements would both be hypothesized to contribute to intensity modulation. This was hypothesized due to resulting changes in glottal width and vocal tract length. In addition, tremor involving the CT muscle was hypothesized to cause F0 modulation with pitch increases. Thus, during habitual pitch phonations (speaking pitch), it was predicted that the individual would demonstrate greater magnitude of intensity modulation than pitch modulation. As shown in bold font in Table 1, intensity modulation magnitude, or extent, was shown to be greater than the extent of F0 modulation during habitual pitch phonation.

Given these findings, it was hypothesized that muscles associated with adduction/abduction should be targeted for denervation treatment. It may also be beneficial to consider treatment of extrinsic and supplementary laryngeal muscles. However, it is recommended that EMG recordings be obtained from these muscles to determine specific extrinsic and supplementary muscles to target.

Table 1. Acoustic measures of the individual with vocal tremor.

<table>
<thead>
<tr>
<th>Rate of F0 Modulation</th>
<th>Rate of Intensity Modulation</th>
<th>Average F0</th>
<th>Average Intensity</th>
<th>Average Extent of F0 Modulation</th>
<th>Average Extent of Intensity Modulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.0</td>
<td>5.0</td>
<td>156.5</td>
<td>84.7</td>
<td>3.3</td>
<td>21.4</td>
</tr>
</tbody>
</table>

Simulated Vocal Tremor

HSDI showed essentially an equal contribution of the left versus the right VF to VoT generation (Figure 3). Therefore either the left or the right VF (TA) can be targeted for denervation.

Nyquist plots form a visual observations based on glottic wave characteristics (Figure 4) showed a “spider web” pattern. Such pattern is seen in severe familial type VoT while a mild Parkinsonian type tremor demonstrated less pronounced F0 dispersions (Figure 5). Nyquist plots of F0 dispersions as shown in Figures 4 & 5 are acquired from the acoustic signal only.
Figure 3. Analysis of the left and right VF using the Vocalizer® system. Note that a separation of the VF occurs at the point of vocal arrest, yet laterality of vibration is not evident. Reprinted with permission [28].

Figure 4. Nyquist plot of severe familial type VoT.

Figure 5. Nyquist plot of mild Parkinsonian type tremor.

Discussion

Here, we proposed a non-invasive clinical method for analyzing audiovisual recordings representing VoT to improve identification of laryngeal intrinsic muscles involved in VoT generation. Patterns of tremor observed within the larynx associated with voicing pat-
terns of increased or decreased VoT may be associated with increased and decreased acoustic modulation of F0 and dB SPL (intensity). One clinical case example was used to test hypotheses proposed, and one control case producing simulated VoT was studied as well.

Additional investigation of this method is necessary before it can be recommended as a method for evaluating individuals with VoT to discern epicenter of laryngeal muscles oscillations in generation of VoT.

References