Chapter 14
Patterns of vibratory vocal fold behavior in patients with vocal fold paresis and paralysis studied using HSDI

Peak Woo

Abstract

The use of high-speed imaging was pioneered by Drs. Moore and von Leden in the 1960’s. Using today’s high-speed digital imaging (HSDI) technology, we can evaluate pathological states of vocal fold (VF) vibration in patients with VF paresis and paralysis that could not be resolved by stroboscopy. Twelve patients with VF paresis (n=8) and paralysis (n=4) underwent evaluation with both stroboscopy and with HSDI. In patients with paralysis and paresis, HSDI shows many differences in vibratory abnormalities that could not be observed with stroboscopy. The vibratory differences between the two sides were greatest in paralysis. Distinctly different fundamental frequency (F0) between the two sides were only seen in the paralysis group while paresis showed phase lag and delay in onset of vibration between the sides. Subtle differences in vibration in patients with paresis could be differentiated by objective image analysis and objective waveform analysis of the HSDI using signal analysis of the digital kymogram derived VF vibrogram (VFV).

Keywords: HSDI, VF paresis, VF paralysis, F0, sidedness, waveform, digital video kymogram, VF vibrogram (VFV)

Introduction

Voice disturbance secondary to asymmetric VF tension can result in roughness, breathiness, or voice breaks. One model for asymmetric VF tension is the patient with unilateral VF paralysis. Due to the asymmetric or poor VF tension on one side, there is differential mass, tension, and stiffness between the two sides. This may result in incomplete closure and vibratory anomaly that is perceived as dysphonia. Traditionally, the stroboscope has been used to image VF oscillation in the clinic. Kitzing published findings with stroboscopy of phase asymmetry in patients with VF paralysis and its return to normal with reinnervation before gross VF motion [1]. Other stroboscopy findings in patients with VF paralysis include incomplete closure, open phase dominant pattern, and asymmetric amplitude and mucosal wave. These observations depend on a steady state vibration with sustained vibration to be able to utilize the stroboscopy effect.

Despite the advantages of stroboscopy in its ease of application in the clinical setting, the use of the stroboscope has inherent limitations in its ability to evaluate voice breaks, chaotic VF vibratory sources, and different generators of the glottal sound. The use of stroboscopy is especially imperfect in patients with moderate to severe dysphonia. Due to short phonation time, steady state vibration is often not possible. The stroboscopy image is then blurred and cannot track F0. Voice breaks and diplophonic voice quality also cannot be resolved with stroboscopy. In this regard, HSDI may add additional information. The effects of asymmetric VF tension was explored in an animal model using high-
speed cinematography by Moore et al. [2]. They noted side-to-side differences as well as additional waves in the oscillation on the paralyzed side. High-speed video was initially proposed by Hirose [3]. With the use of scanning camera for kymography, applications in the clinical setting confirmed that multiple abnormal vibratory patterns could be identified in patients with VF paresis using the kymography camera. These include phase abnormalities, double vibrations, and transients associated with voice breaks [4-5]. Since then, multiple authors have cited case reports of vibratory asymmetry that adds to the clinical role of detection of vibratory abnormality resulting in dysphonia in patients with VF paralysis or paresis [6-7]. Despite these reports, there are no reports that compare VF behavior from patients with VF paralysis and VF paresis using HSDI versus stroboscopy.

One area that is unique for study by HSDI is the onset of VF oscillation in patients with VF paralysis. This has never been studied, even with high speed cinematography. One can hypothesize that HSDI can add additional information on the initiation of VF vibration in patients with paresis and paralysis. This is especially interesting as there is asymmetric tension in the initiation of phonation between the sides in patients with paresis and/or paralysis. Perhaps such studies can show differences between the two groups of paralysis versus paresis. Furthermore, by using image enhancement techniques to track the VF margin, such an approach may give objective information of each VF vibratory pattern that can be compared within groups.

This preliminary study looks at 12 patients with VF paresis (n=8) or paralysis (n=4). This paper discusses the qualitative differences between videostroboscopy and HSDI for each case and attempts quantitative analysis of the digital kymographic tracing.

Materials and methods

Twelve subjects with VF paralysis and VF paresis underwent both traditional stroboscopy examination followed by HSDI examination. The standard stroboscopy examination was done using a digital image capture and analyzed using standard rating protocol [8].

The digital images were acquired using Kay Pentax HSDI system (Model 9100; Lincoln Park, NJ), which consisted of a 90° rigid endoscope (Model 9106; Lincoln Park, NJ) coupled with a 300-W Olympus cold light source (CLV-U20; Center Valley, PA). The HSDI system acquires gray scale images at a rate of 2000 f/s with a spatial resolution of 240 x 140 pixels. The laryngoscope was performed as in conventional videoendoscopy procedure using a 70° rigid transoral endoscope. The patient was seated with the hips back, the torso leaning forward, and the chin tilted slightly upward. When there was a clear view of the larynx, the patient was asked to produce the phoneme /i/ in short two-second intervals. Modal phonation at comfortable pitch and loudness was requested from the patient. The males typically used 100-150 Hz at 70-75 dB while the females used 190-250 Hz at 70-75 dB. The foot-trigger was subsequently engaged to capture the images and six continuous tokens of the /i/ phoneme utterance were captured. Four seconds of phonation was captured at 2000 f/s. After capture of the token of sustained phonation at comfortable pitch, the segments were reviewed offline and compared to the stroboscopy recordings.

High-speed image analysis and quantification

Kymograph analysis of the VF motions is critically dependent on accurate delineation of the VF edges from HSDI [9]. A ~400 frame video that demonstrated a full view of the VF with minimal movement of the subject was extracted from the HSDI recording to build a kymogram. Kay’s Image Processing Software (KIPS) (Model 9181; Lincoln Park, NJ) was used to generate each kymogram by placing a transversal line across the glottis at the place of maximal VF contact. Edge detection was subsequently applied to the kymogram,
which identified and traced the VF edges of the kymogram. If automatic tracing was imprecise, manual correction functions (e.g., edge restrict, brightening, darkening, and erase) were used to ensure noise was not assimilated into the analysis. When the tracing was proximal to the VF edges, the Kymograph Edge Analysis (KEA) was applied on the kymogram. The resulting values were Kymograph Edge Data (KED), which describes the coordinate values of the left and right edges of the VF presented across time or frequency.

The VF edge from the mid-point of the kymogram was converted to the kymographic edge waveform for each VF and subject to frequency analysis using the Fast Fourier Transformation (FFT) function on the KIPS software. From these spectral plots, the frequency components of each VF was plotted as power versus frequency for each fold. The kymographic patterns from paralysis and paresis were transformed into a frequency spectral plot for further frequency analysis. Because the kymographic tracing traces both right and left VF, the spectrogram plot between the right versus the left side can be compared for frequency and power for that token.

**Quantification of the digital kymography waveform**

Once the video image has been acquired, the cursor is placed on the mid-membranous VF portion. Multiple lines may be placed to obtain multiple digital kymography lines.

Once the videokymography display is done, the image can be converted to the digital kymography waveform using edge detection software. Figure 1 is the line tracing of the digital kymogram. The digital kymogram is the mid-membranous portion of the VF during the onset and stable portion of the token.

![Figure 1](image1.png)

**Figure 1.** Digital Kymogram from the mid-fold of a normal male subject phonating at 118 Hz and 75 dB. Time is in the Y-axis while the excursions show the movement each VF from the midline at the area of interest. Note the tracing by edge detection tracked each fold edge nicely. Note that each VF is almost a mirror image of the other. Image rotated 90° to the left.

The DKG is then transformed into the DKG waveform based on edge detection using the KIPS software. For this DKG, the lines traced out nicely based on the edge detection. The glottogram waveform extracted from the edge detection can then be analyzed as a numeric waveform. One can appreciate two lines for each kymogram with one line going upward representing the right and the other going downward representing the left VF. This is the VF excursion for each VF based on the DKG. This waveform has been termed the VF vibrogram (VFV). The VFV is the edge movement for each VF. The VFV patterns can be transformed into a numerical waveform. Figure 2 is the line tracing of the glottal edge tracing from the edge detection at the mid-membranous VF. The left VF edge goes downward by convention and the right goes up. The midline is at zero. Time is displayed in the x-axis. The VFV displays the edge of each VF over time.

This VFV waveform is suitable for signal analysis using FFT. This is available in the KIPS package. Figure 3 is the FFT transform of a normal VFV. The plot has the power on the y-axis and frequency on the x-axis. The right fold VFV power spectral tracing is plotted in red and the left fold VFV spectrogram is plotted in black. The frequency vs. power plot represents the spectral energy at each frequency. Some features are typical of the nor-
mal power spectral plot: 1) the energy is limited to the F0 and its first few harmonics, 2) there is little sub-harmonic energy and inter-harmonic energy, and 3) there is symmetry of the spectral plot between the left and right VF.

In this way, we can both analyze the vibratory pattern qualitatively and then using the frequency analysis of the VFV, we can get some objective evaluation of functional contribution of VF vibration from each VFV.

Figure 2. Graphic tracing of the VFV. This is derived from the edge detection software that analyzed the DKG tracing at the mid-membranous fold. Time is on the x-axis. The right VF edge goes up while the left tracing goes down from the midline value of 0. This is in a normal subject.

Figure 3. Power spectral plot of right and left VFV in a normal male phonating at 145 Hz and soft phonation of 70 dB. Notice the tracings from the FFT for each VFV waveform are superimposed on each other with spectral energy at F0 and the first three harmonics. There is little inter-harmonic energy and low frequency energy.
Results

The twelve subjects studied here had a variety of severity of dysphonia that varied from aphonia to near normal phonation. The stroboscopy findings in the paresis patients included reduced diadochokinesis, limited abduction or adduction, different height or length differences between the two VF, or a combination of the above. If there was no motion, a diagnosis of paralysis was made. If there was motion, but asymmetry of motion was present, a diagnosis of paresis was made. For this study there were no systematic testing of the laryngeal EMG to verify whether the paresis group had electrical signatures consistent with paresis.

For each case of paresis or paralysis, the HSDI video was reviewed in its entirety from onset to steady phonation to offset of phonation. Selected video features that were unable to be seen on the stroboscopy examination were cut and saved for future image analysis and frequency analysis of the VFV using the KIPS software. This included: onset and offset of VF oscillation, phase differences in the anterior to posterior plane, transient phase shifts in the lateral to medial plane, voice breaks, and diplophonic voice breaks. Table 1 shows the vibratory observations for the cases during onset, steady phonation, and offset.

Table 1. Tabulations of the vibration observations in onset, steady phonation, and offset.

*New observations not otherwise observed by stroboscopy.

<table>
<thead>
<tr>
<th>Dx</th>
<th>Onset*</th>
<th>Steady phonation</th>
<th>Offset*</th>
</tr>
</thead>
<tbody>
<tr>
<td>left paresis</td>
<td>delay onset left, left starts later*</td>
<td>left greater amplitude,</td>
<td>left stops later*</td>
</tr>
<tr>
<td>left paresis</td>
<td>delay onset left*</td>
<td>phase shift, amplitude difference</td>
<td></td>
</tr>
<tr>
<td>right paresis MTD</td>
<td>MTD</td>
<td>MTD no visualization</td>
<td>big wave on right*</td>
</tr>
<tr>
<td>left paresis</td>
<td>normal onset</td>
<td>Phase shift, arytenoid vibration*</td>
<td></td>
</tr>
<tr>
<td>paralysis left</td>
<td>delay onset MTD, on-off on*</td>
<td>floppy irregular vibration, diplophonia *</td>
<td>difference in offset*</td>
</tr>
<tr>
<td>paralysis right</td>
<td>onset delay*</td>
<td>big amplitude, right phase shift</td>
<td>difference in offset*</td>
</tr>
<tr>
<td>left paresis</td>
<td>transient breaks, phase shift</td>
<td></td>
<td></td>
</tr>
<tr>
<td>left paresis</td>
<td>onset delay, MTD*</td>
<td>asymmetric vibration</td>
<td>asymmetric offset*</td>
</tr>
<tr>
<td>left paresis</td>
<td>onset delay*</td>
<td>AP phase shift, breaks</td>
<td></td>
</tr>
<tr>
<td>paralysis right</td>
<td>obscured</td>
<td>asymmetry in vibration, diplophonia, multiple frequency*</td>
<td>asymmetric offset*</td>
</tr>
<tr>
<td>paralysis right</td>
<td>onset chaos*</td>
<td>asymmetry in vibration, diplophonia, multiple frequency*</td>
<td>asymmetric offset*</td>
</tr>
<tr>
<td>right paresis</td>
<td>delay onset*</td>
<td>AP phase shift, phase lag, right vibrates later</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4 represents a typical high speed image montage of four glottal cycles in a patient with VF paralysis. One can clearly see the vibration on the right with each cycle but the right side vibration is not clear. On slow motion viewing, one can appreciate both VF to be oscillating with chaotic motion coming from each side.

Figure 4. A montage of 29 frames of video clipped for analysis. Note the area of interest is limited to the VF vibration area. This facilitates the analysis by the software for edge detection or for kymography analysis.

Figure 5 is a DKG image of the VF at the mid-membranous site during phonation onset. The kymogram is on the left while the right figure shows the line of interest sampled by the digital kymogram software. This is a patient with left VF paralysis. This shows onset of the VF to be quite chaotic with multiple frequencies. The VF have onset of vibration on the paralyzed side (left side) slightly later than the right side. One can see oscillation with multiple beats earlier on the right side. There is a delay in ramping up of the VF oscillation on the left side compared to the right non-paralyzed side. There is a phase lag behind the innervated right side. There is almost a 180° phase shift in some frames while in others there is only brief contact.

The edge of the above DKG can be extracted for spectral analysis (i.e., VFV). The VFV is time-based and has a value for the right and the left fold as it moves from the mid-line. The VFV can be analyzed in frequency domain by signal analysis. Using FFT, a spectrogram can be plotted in the frequency domain with the frequency in the x-axis and the y-axis represented by the power. The FFT results in two lines with one representing the left VF (red line) and the other the right VF (black line). The results of the spectrogram is shown in Figure 6. When this digital kymography line VFV is extracted for analysis, one can appreciate two distinct spectral peaks in the FFT plot. The innervated right VF has a lower fundamental frequency but has larger spectral peak. Paralyzed side appears floppy and has several spectral peaks (red line). Both right and left VF vibrations spectrogram shows a reduction in energy in the first and second harmonics of the F0. There is increased inter-harmonic energy and an increase in the low frequency below the F0. This compares very differently from that of normal (Figure 3), which shows distinct spectral peaks with each VF spectrogram superimposed on the other.
patterns of vibratory vocal fold behavior in patients with vocal fold paresis and paralysis studied using HSDI

Figure 6. Spectrogram of patient with paralysis and diplophonia during phonation onset.

After steady state phonation has been established, the VF vibration is now steady and the vibration now shows quasi-periodic vibration. During steady phonation, there are still irregular changes in vibration. The kymographic signal showing every fourth beat changes in amplitude (Figure 7). These are associated with the patients diplophonia and is shown as low energy peak below the fundamental frequency of VF oscillation.

Figure 8 is the FFT analysis of the kymogram shown in Figure 7. The amplitude difference is seen in the difference in the size of the spectral peak. There is still loss of spectral peaks in the higher harmonics. The spectral peaks are very close but not the same. This in and out of phase interaction is believed to be the cause for diplophonic voice quality. The energy below the fundamental is elevated. This corresponds to the visual reading of the DKG showing irregular beats every six cycles. This small spectral peak at the lower frequency is distinct but it is not clear if this represents the sub-harmonic responsible for the perceived diplophonia.
Figure 7. Kymogram showing regular asymmetric changes in amplitude and phase as the VF come in and out of phase to each other. Every 6th cycle results in an abnormal beat.

Figure 8. Spectral plot of VF in patient with paralysis during steady phonation.

Figure 9 is another patient with unilateral VF paralysis and diplophonia. The right and left VF have different oscillatory frequency of three beats to every two beats. This results in phase asymmetry and incomplete closure. However, every third beat on the left and every fourth beat on the right VF vibrate in phase to the opposite VF. This results in vibration in-phase for that cycle. Subsequent cycles then go out of phase. This creates a sub-harmonic to the F0 as the VF come in and then out of phase to each other.

When this image is subject to image analysis by edge detection (Figure 10), the tracking is robust enough to submit for waveform analysis. Waveform analysis using FFT shows two lines with different spectral peaks (Figure 11). This pattern of multiple frequencies was only seen in the paralysis group in three of the four cases. The paresis group had more subtle findings and did not show different F0 as in the patients with paralysis.

VF that are innervated vibrate better than the side that is paralyzed. Not all paralyzed VF will look floppy, and the side of paralysis may look stiff and more immobile. In such instances, the side that is more pliable will have a greater amplitude, which results in a higher spectral peak. What is unique in patients with VF paralysis is that the patients with
diplophonia due to paralysis have different frequencies between the two sides. This distinct pattern could be better resolved with HSDI analysis, but have not yet been reported. The diplophonic patients with paralysis with multiple frequencies between the right and left sides all had complete paralysis. The VFV and spectral analysis showed the pattern of multiple F0 to be present only in patients with paralysis. We did not see this pattern in paresis and we did not see this pattern in other patients with diplophonia due to mass or stiffness abnormalities. Verification if this is a feature of asymmetry of tension seen more commonly in patients with complete paralysis is needed.

**Figure 9.** DKG of patient with diplophonia and VF paralysis.

**Figure 10.** Successful edge detection for VFV analysis. Image rotated 90° to the left.

**Figure 11.** Spectral analysis of the VFV of patient with VF paralysis.
Paresis

We will now illustrate some typical findings in patients with VF paresis. In patients with VF paresis, the findings may be subtle. Stroboscopy may show open phase dominant pattern or it may show a lateral to medial phase shift. Rarely is there diplophonia, chaotic vibration, or aphonia. Here steady phonation showed a good quasi-periodic vibration with the good side showing greater amplitude. These observations could be appreciated by stroboscopy. However, when one looks at the onset of phonation, there are multiple vibratory anomalies that could be detected only on HSDI. During the onset of phonation there is phonatory adduction but the initial vibrations are often irregular. There is occasional anterior to posterior phase shift as well as prolonged alternating beats before steady state vibration. After steady state has been achieved, there is more quasi-periodic oscillation. Figure 12 is the DKG of a subject with VF paresis. The onset between prephonatory set to the onset of steady oscillation is long and there is a period of adduction followed by VF opening and then steady phonation. This may be a muscle tension dysphonia pattern as the patient compensated for the paresis.

Figure 12. Onset of VF oscillation in a patient with paresis. Note the prolonged delay between VF approximation and onset of VF vibration. The VF are closed and then open and then assume steady oscillation. Image rotated 90° to the left.

During the offset of VF vibration, as the patient is breathing in, the paralyzed side vibrates in the airstream while the innervated side does not. This was observed in several paresis cases. This is assumed to be due to the flaccid VF vibrating in the air stream. Most subjects with paresis were able to sustain phonation and stroboscopy tracking is not a problem. With stroboscopy open-phase dominant pattern and subtle phase shifts may be present. However, HSDI can still detect aperiodicity that otherwise were not obvious. HSDI is effective for picking up aperiodic voice breaks. This is shown in Figure 13. Voice breaks tends to be transient and difficult to pick out except by eye. Signal analysis was not as helpful in identifying transient breaks due to signal averaging.

The most common finding in patients with paresis is the open phase dominant pattern. In this pattern, videostroboscopy can detect the finding as the VF are oscillating in a quasi-periodic fashion. Figure 14 shows a patient with paresis where the open phase seems long compared to the closed phase. The glottal configuration shows a short AP contraction consistent with muscle tension pattern.
Figure 13. A voice break is shown on the DKG as a couplet of VF vibrating with prolonged contact at the top. The VF then assume more periodic oscillation.

Figure 14. DKG of a subject with VF paresis and open phase dominant pattern. Image rotated 90° to the left.

During steady state phonation, the VF in most patients with subtle paresis seems to have excellent in-phase vibration with normal closure. The DKG appears normal and there is little to indicate an abnormal vibratory pattern on analysis of the VFV. However, abnormalities can be appreciated on spectral analysis. Figure 15 shows the VFV tracing during offset and Figure 16 shows the spectrogram showing symmetric spectral peaks. However, there is loss of energy in the higher harmonics of F0. This finding is typical of patients with open phase dominant pattern of VFV. Spectrogram analysis of the DKG shows symmetric spectral peaks with little difference between the two sides.

We will illustrate the added possibilities of VFV analysis in subtle paresis with a last case of a classical singer with subtle voice breaks. The stroboscopy showed no abnormality. The DKG tracing confirmed a low frequency shift every tenth beat (Figure 17). Notice that although the VF are vibrating in phase to the opposite side, the VF goes in and out of open phase dominant pattern. This seemed to correspond to the patient’s complaint of loss of power and fluctuation in control. Additional areas where there was noted voice break that otherwise could not be seen with stroboscopy are shown in Figure 18. On the top three cycles, a clear break has occurred on the third from the top frame. This occurred without phase difference but during that cycle, there is little VF contact.
Figure 15. VFV of VF paresis during offset shows VF in-phase vibration little contact.

Figure 16. FFT of the VFV in a patient with paresis shows spectral peak at F0 but little energy at the first and second harmonic. This corresponds to observations of open phase dominant pattern with loss in the closed phase. Also noted is the large splay of the energy in the first spectral peak, an indication of increased jitter.
Discussion

Because high-speed cinematography was not practical for clinical purposes, the literature regarding HSDI was limited to a few centers. The first report of a digital imaging device based on a charge coupled device (CCD) was reported by Honda [10]. Some of the initial reports using HSDI imaging was by Hirose [3, 11-14]. This was limited to 100 x 100 pixels and had poor detail resolution. Nevertheless, many new observations were now available in terms of vibratory asymmetry as well as sources of vibration that could not be observed using standard videostroboscopy. Other authors have used high-resolution high-speed cameras to make observations regarding periodicity and asymmetry [15].

With the continuing cost reduction of CCD cameras, a commercially available high-speed video system became available in early 2000. This system initially had 256 x 256 pixel resolution in black and white. Today such as system is available in 512 x 512 in color (Kay Pentax, Montvale New Jersey). HSDI is now readily utilized in the clinic with a hand held camera inputting video frames that are quite manageable by using personal computers.

Figure 17. During sustained phonation, the VF goes in and out of open phase dominant pattern in this classical singer complaining of voice instability.

Figure 18. Area of voice instability and voice break is noted on the first four cycles. This is followed by steady oscillation.
HSDI overcomes some of the problems with stroboscopy. HSDI allows thousands of pictures to be taken of the vibrating VF per second. Rigid endoscopes are normally used for HSDI. After clinical investigation of the patient, kymographic analysis from the video can be carried out offline. In one clinical model, up to 8000 grayscale images of 256 x 256 pixels can be stored (Wolf HS Endocam 5560). In other systems a maximum of 4000 f/s can be taken by the high-speed camera [16]. In contrast to stroboscopic investigations, aperiodic movements of the VF can be visualized. The duration of a recording differs from 2-4 seconds depending on the capture speed. The data transfer to the computer for archiving and analyzing the images takes a few minutes.

Clinical applications of HSDI analysis in patients with paralysis and paresis

The use of HSI for assessment of patients with vocal cord paralysis was initially published by von Leden and Moore in 1961. In that article, high-speed cinematography was used to record the vibratory pattern during steady state phonation. Vibratory asymmetry between the innervated VF and the normal VF was noted. Also noted was the presence of chaotic vibration [17]. Hirano also used high-speed cinematography to show the effects of asymmetric tension [18]. The problem of asymmetric tension in patients with unilateral VF paralysis was studied using videostroboscopy in a canine model by Moore et al. in 1987. Results indicated that simulated paralysis of the recurrent laryngeal nerve, or combined paralysis of the recurrent and superior laryngeal nerves, produced a diminished mucosal wave bilaterally with loss of the two-mass system of vibration and diminished lateral excursion of the normal VF. Simulated paralysis of the superior laryngeal nerve was characterized by an abnormally exaggerated vertical movement of the paralyzed VF in relation to the normal cord [2]. Eysholdt was the first to use HSDI to evaluate asymmetric VF vibrations and noted distinct frequency differences between the sides in case of asymmetric tension. Two types of irregularities were noted with one side having frequency difference between left and right VF and the other being vertical asymmetry. It was presumed that laryngeal asymmetry (either in mass or tension) causes these irregular vibrations [19]. In a study of benign VF lesions with videokymography, Kim and co-authors noted irregular VF vibration and incomplete closure of the VF. Much larger asymmetric changes were present for unilateral VF paralysis than in normal controls or for other lesions. The authors noted that asymmetric index may be a good quantitative parameter of vibration in patients with VF paralysis [20].

Since there is often a large gradient of voice disturbances in patients with VF paralysis, the use of HSDI should be considered in patients where the videostroboscopy could not be clinically utilized to resolve the vibratory anomaly adequately. We believe that clinical situations where stroboscopy fails in patients with suspected VF paralysis and paresis is in the diagnosis of voice disturbance such as voice breaks and in the analysis of the severely disturbed voice is that is accompanied by diplophonia.

From evaluation of patients with severe VF paralysis, a very distinct vibratory pattern in patients with VF paralysis have emerged. This pattern is characterized by one VF vibrating at a distinct frequency different from the other. This results in chaotic vibration for which VF independently come in and out of phase relative to the contralateral VF. Such interaction between two VF vibrating at different frequencies results in intermittent changes in VF oscillation where one VF vibrates out of phase relative to the other followed by VF vibrating in phase relative to the other. Such in and out of phase interaction results in a distinct perception of a sub-harmonic of the fundamental frequency. This is perceived as diplophonia.

The lower sub-harmonic frequency is below that of either VF and this is a direct result of the phase shift interaction between the two VF oscillating at distinct frequencies. In all the cases of complete paralysis, the paralyzed VF oscillates at a slightly different
frequencies than the innervated VF. This pattern was only seen in the complete paralysis group. In other examples of diplophonia such as diplophonia due to stiffness, mass lesion, or muscle tension, two distinct VF oscillators with different fundamental frequency between the two folds were not observed. In this way, the finding of two distinct oscillators in a patient with VF paralysis suggest loss in VF symmetry of tension and appears to be uniquely identifiable in patients with complete VF paralysis and not paresis.

Additional findings in patients with complete paralysis using HSDI demonstrate many other observations that were not identified using stroboscopy. Multiple oscillators can be present with arytenoid vibration. False VF vibration and interaction between the true and false VF can occur. In the patient with the paralyzed VF, the paralyzed VF appear to vibrate with greater amplitude and often is accompanied by both lateral to medial movement as well as anterior to posterior movement. This gives the impression of a floppy VF with a single mass of vocal vibration. This contrast with the innervated side during vibration, which shows the typical oscillatory properties of body versus cover.

In patients with subtle voice abnormalities due to paresis, the initial onset of voice by approximation of VF into the phonatory position may offer new insight on the physiology of VF oscillation. During stable onset of phonation, the VF are brought into the pre-phonatory position and stable oscillation occurs usually within 3-5 glottal cycles. In patients with VF paresis, multiple examples of phonatory delay were noted. A typical finding was phonatory onset characterized by excessive muscle tension and delay in the phonatory stability of oscillation. In patients with VF paresis, the patient may show delay in phonatory onset with chaotic vibration followed by steady oscillation. Analysis of the steady oscillation shows that there is brief phonatory contact with open-phase dominant pattern. Using FFT analysis of the vibratory wave, high F0 relative to the energy at the third harmonic suggests an open phase pattern.

Initially, we were quite skeptical that HSDI with all its time and effort would be able to identify vibratory abnormality in patients with voice sources that were stable and periodic. In these patients, stroboscopy is usually very suitable. Indeed, when one examines the HSDI video in subtle cases of VF paresis, many of the subjects have near normal voices. In these patients, the primary observation is open-phase dominant pattern and occasional lateral to medial phase shift. These abnormalities are easily visualized on video stroboscopy. Only with the use of digital kymography and in-depth analysis of prolonged runs of the high-speed data could the more subtle abnormalities be identified. These include observations of voice breaks and phase asymmetries that occur over short time intervals. For example, the identification of voice breaks could not be identified unless there is a systematic reading of the entire vocal gesture from voice onset, steady state phonation, and through voice offset. When this is done, multiple abnormalities in the periodicity of VF oscillation as well as in asymmetric vibration between the two sides can be identified. There remains a lack of data on degree of asymmetry and voice breaks in normal vocal function, but these breaks appear to correlated to subjective complains of voice instability. In patients with subtle paresis, the paralyzed side often vibrates with a more sinusoidal pattern due to thinning of the VF margin. Because VF paresis results in a more sinusoidal pattern of VF oscillation with reduced VF contact, such a pattern results in energy in the F0 but a relative lack of energy in the first, second, and third harmonic. This can be objectively documented by signal analysis of the VFV waveform. We believe with better quantification of the spectral signature of VF oscillation, normograms can be set up for males and females based on frequency and amplitude. This data can be used to objectively document vibratory capability of the VF during sustained phonation based on amplitude and frequency.
Barriers to Clinical Application

Despite these new findings of vibratory abnormality in patients with VF paresis and paralysis using high-speed video, the clinical utility of such findings remains obscure. Mortenson reported on the use of HSDI of a case of patient with diplophonia that showed a distinct vibratory frequency for each VF, suggestive of a paresis pattern. Laryngeal electromyography confirmed the diagnosis of VF paresis. A computed tomography scan of the larynx and chest showed a thymoma. This was not obvious on stroboscopy or other methods of visualization [7]. Other clinical applications using high-speed video have yet to be published in the literature as to how HSDI use in patients with dysphonia adds to clinical utility.

Future investigation will be needed to make HSDI practical and clinically relevant. From HSDI of patients with paralysis and paresis, we can conclude that multiple new observations with regards to VF vibration abnormalities are now identifiable. In patients with subtle paresis, the investigation of voice onset and voice offset will likely be new areas of interest. During steady state VF oscillation, signal analysis of the vibratory waveform using frequency analysis can look at the VF vibratory pattern in a quantitative way. This ability is unique in being able to differentiate vibratory abnormalities between the right versus left VF. The ability to use digital kymography and reduce the data from a complex visual observation to a line-tracing representation of VF behavior is exciting. The ability to analyze the VFV to gain information on vibration-related signatures of each VF for the phonation gesture may also have novel applications in comparing results before and after treatment.

In patients with chaotic VF vibration and in patients where there is short phonation time with severe roughness of voice, high-speed video imaging allows the clinician real insight into VF vibratory behavior. Patients with diplophonia can now be resolved as to the vibratory source contributing to the voice abnormality. Multiple oscillators in the vocal tract can now be routinely identified either qualitatively using visual review of the HSDI in slow motion, or alternatively, using image enhancement and edge detection to do semi-automated image extraction and waveform analysis. In this preliminary project we have identified a unique pattern of VF vibration in patients with high-grade VF paralysis. This is postulated to be due to asymmetric VF tension and results in two different fundamental frequencies that vibrate in and out of phase, resulting in diplophonia.

The difficult problem of analysis of HSDI should be resolved before clinical application can become routine. This may come in semi-automated image analysis routines that screen the large amount of video data for adequate image quality and then extract key areas of VF instability or vibratory abnormality for visual review. Other algorithm for analysis of the vibratory waveform may include waterfall displays of the vibratory capability of the each VF over time from the onset of vibration through the steady state of vibration to the offset VF vibration. In this way, the tedious and time-consuming need to manually view and analyze large amounts of high-speed video data can be reduced.

Despite the obstacles that currently exist in routine capture and analysis of HSDI data, this author believes that HSDI analysis is here to stay. Within the last decade, multiple journal articles have pointed to the new observations in vibratory capability in healthy and disordered VF that are able to be captured and analyzed using HSDI systems. It is now practical and within the budget of clinicians to routinely record and analyze the vibratory behavior of their patients with suspected VF paresis and paralysis using high-speed video. Using systematic qualitative and quantitative methods outlined above, the author believes the clinical utility of high-speed video in understanding patients with VF paralysis should be enhanced.
Conclusions

Despite interest in quantification of the VF vibratory function over the last 50 years, HSDI application in clinical voice disorders is still in its infancy. Routine analysis of both normal and pathological VF vibration before and after treatment using HSDI is now possible. By adding applications of objective analysis of VF vibratory function by waveform analysis of the DKG signal, the clinician has a powerful tool in understanding VF pathology. Better understanding of normal and pathological vibratory function of the VF should be valuable in the development and testing of newer surgical procedures for improvement of vocal function. The availability of low cost high-speed digital video cameras and image processing methods should help to standardize the analysis routines. Collaboration between the clinician and the voice scientist for the development of imaging and analysis software will go far in improving the knowledge base between clinical care and objective results.

References