Aspects of Resonance: Comparison of High Speed Films and Overtone Measurements

Mette Pedersen MD, Bilal Hussain Akram, Mohammad Shezad Mahmood, Anne Alexius Agersted, Anders Jønsson and Sanilla Mahmood

The Medical Centre, Østergade 18, 3, 1100 Copenhagen, Denmark

Corresponding author: Mette Pedersen MD, The Medical Centre, Østergade 18, 3, 1100 Copenhagen, Denmark, Tel: +4531126184; E-mail: m.f.pedersen@dadinet.dk

Abstract

A comparison has been made simultaneously of high speed films, acoustical analysis and overtone analysis with "VoceVista" and "Sygyt Ltd". The overtones software was compared with acoustical analysis. Till now formants from 1000 Hz-5000 Hz seems to be interesting using fundamental frequency of speech with 110Hz in males and 220 Hz in females. The variation of the 3 formants from 1000 Hz-5000 Hz was between 18-25% for normal persons. Glottal analysis tools from Erlangen showed no evidence of traditional acoustical voice analysis. With the new overtone analyzer (Sygyt Ltd.) we have a tool for measuring pathological voices combined with high speed films.

Keywords: High speed films; Overtone analyses

Introduction

There are new apparatuses on the market-based on voice research-where the patient can benefit from advanced diagnostics of overtones.

A clinical routine with high speed films includes kymography, electroglottography and quantitative measures of the movement of the left and right vocal fold. It was showed that irregularity of the two vocal folds against each other hardly ever was the case. So airflow correction as used traditionally for correction of voices is doubted. Therefore, new approaches had to be made, especially not related to airflow. Documentation of overtones measures up to 20000 Hz (or more) with a stable and well established overtone analyzer leads to new clinical experiences – since it is made in a clinic friendly easy way – even cheap. Next is to optimize the analysis of larynx functions with among others Optical coherence tomography and Narrow Band Imaging to understand the mucosal function.

With high speed films combined with electroglottography, kymography, and overtones, nuanced evaluation of the voice is possible and to some extend evidence based. The software of high speed films includes quantitative measures of the closure of the vocal folds as well as stiffness, a calculation of maximal amplitude versus maximal speed of the vocal folds. In all 345 measurements of the voice can be made with the software that is on the market-"Glottal analyses tools" from Erlangen Germany. Acoustical analysis as well as glottal area waveforms calculations are made online on high speed films. The most used acoustical examination do not give sufficient or reliable information in pathology.

Method

The literature has been reviewed extensively for clinical aspects of resonance/overtones/harmonics: A British Library search of resonance of the human voice included search words: "voice and resonance" for the latest two years found 1610 references out of which some interesting results have been studied especially from the book by Donald Miller.

We have compared the "VoceVista" by Donald Miller with the newer software “Sygyt Ltd” which is easy to use and we have found them comparable.

In pathology there are no answers of: why one singer is better than another. Many demonstrations have been made of average formant analysis during singing of melodies – sentences - by perfect singers (Pavarotti etc.). The challenge is, with one good overtone measuring method, to explain the patient complaints of various kinds of hoarseness, in an evidence based way in the clinic.

We tried to compare the methods of overtone/harmonics analysis of "VoceVista" with "Sygyt Ltd" to find out if the two methods were comparable so that result could be used with both analyses. At least twelve subjects are needed statistically to describe the distribution of a normal material. This was also the case of the analysis of the overtones of the human voice.

In the literature the fundamental frequency (F0), in a scientifically usable way in pathology, was seldom defined in the formant analyses of singers. Therefore, we decided to compare the fundamental frequencies (F0) of voice as used in speech and harmonic overtones hereof, just to have a value for pathology. We also compared the sound analysis of the overtones of the human voice based on the fundamental frequency in speech comparing "VoceVista" with the more flexible “Sygyt Ltd”.

We compared 12 normal voices of persons without voice complaints with measures of “VoceVista” and "Sygyt Ltd". 12 persons were statistically enough in a prospective cohort study to characterize a normal material and it was used to identify major differences between the two systems.

As the voices of patients are very different in nature, it is a challenge to identify the vocal problems. In order to identify the pathological problem in the clinic, a sensitive and specific measurement is needed. In the literature formant measurements over several seconds have been analyzed of famous singers without description of F0. In a clinical situation singing overtones and the measurements of them will vary extremely much and results can be understood in many ways. So to begin with, a clinical approach was made of the mean fundamental
frequency (F0). The formants as defined with "VoceVista" which measures to 5000 Hz were compared with "Sygyt Ltd." that measures to 20,000 Hz. They were made only over 1000 Hz to exclude articulation related variations.

**VoceVista**

**Spectrogram:**

Up to 5000 Hz

Fundamental frequency of speech was chosen to measure overtones:

- Male: 110 Hz
- Female: 220 Hz

**Sygyt (Ltd.)**

**Spectrogram:**

Up to 20000 Hz

Fundamental frequency of speech as chosen for measures to be used in pathology:

- Male: 110 Hz
- Female: 220 Hz

**Formants**

We have defined one formant as bundle of harmonic scale which combined represent the formant.

The harmonic overtone with the highest point in dB reflected in the formant was selected.

The amplitude of the selected harmonic overtones was read from the lowest to the highest point.

The first three formants over 1000 Hz were measured and the associated dB for each formant noted. Some formants less than 1000 Hz were noted in parentheses, even if they are articulation related.

Normal persons were used, 6 males and 6 females were analyzed. The fundamental frequency of speech (F0) during intonation of "ah" was defined, and a prospective cohort study of the formant placements of F0 with "VoceVista" up to 5000 Hz and "Sygyt Ltd" up to 20,000 Hz was made.

From a statistical point of view, measurements of the same millisecond (ms) was necessary to compare the formants of the two systems in a test of intonation, see Figures 1 and 2.

**Results**

"VoceVista" results of formants from 1000 up to 5000 Hz named Fx, Fy, and Fz in a prospective cohort study of 12 normal persons with the F0 nearest to 220 and 110 Hz. (Table 1).

"Sygyt Ltd." results of sound formants from over 1000 Hz up to 5000 Hz with 110 and 220 Hz as basis of overtone measures of the same 12 normal persons (Table 2).
Figure 2: “Sygyt Ltd.” overtone analyzer to 20000 Hz with formant detection to the height at 220 Hz, a normal female's voices.

Table 1: “VoceVista” results 12 normal persons.
### Table 1: "Sygyt" Results

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Name</th>
<th>Gender</th>
<th>Age</th>
<th>ms</th>
<th>F0 (Hz)</th>
<th>F0 (dB)</th>
<th>(Fa) (Hz)</th>
<th>(Fa) (dB)</th>
<th>Fx (Hz)</th>
<th>Fx (dB)</th>
<th>Fy (Hz)</th>
<th>Fy (dB)</th>
<th>Fz (Hz)</th>
<th>Fz (dB)</th>
<th>Lowest</th>
<th>Highest</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MP-B</td>
<td>F</td>
<td>75</td>
<td>2520</td>
<td>220</td>
<td>58</td>
<td>436</td>
<td>55</td>
<td>1545</td>
<td>31</td>
<td>2433</td>
<td>35</td>
<td>4662</td>
<td>20</td>
<td>231</td>
<td>1126</td>
</tr>
<tr>
<td>2</td>
<td>ACA-B</td>
<td>F</td>
<td>25</td>
<td>14150</td>
<td>220</td>
<td>43</td>
<td>436</td>
<td>52</td>
<td>1555</td>
<td>24</td>
<td>2874</td>
<td>33</td>
<td>3881</td>
<td>13</td>
<td>174</td>
<td>610</td>
</tr>
<tr>
<td>3</td>
<td>LTC-B</td>
<td>F</td>
<td>40</td>
<td>2330</td>
<td>220</td>
<td>43</td>
<td>403</td>
<td>48</td>
<td>1114</td>
<td>40</td>
<td>2304</td>
<td>20</td>
<td>3526</td>
<td>16</td>
<td>190</td>
<td>874</td>
</tr>
<tr>
<td>4</td>
<td>KJH-B</td>
<td>F</td>
<td>47</td>
<td>2665</td>
<td>220</td>
<td>58</td>
<td>441</td>
<td>55</td>
<td>1539</td>
<td>42</td>
<td>2201</td>
<td>26</td>
<td>3762</td>
<td>19</td>
<td>151</td>
<td>760</td>
</tr>
<tr>
<td>5</td>
<td>SM-B</td>
<td>F</td>
<td>24</td>
<td>310</td>
<td>220</td>
<td>45</td>
<td>441</td>
<td>56</td>
<td>1765</td>
<td>25</td>
<td>2869</td>
<td>22</td>
<td>4333</td>
<td>12</td>
<td>211</td>
<td>491</td>
</tr>
<tr>
<td>6</td>
<td>NBL-B</td>
<td>F</td>
<td>25</td>
<td>620</td>
<td>220</td>
<td>51</td>
<td>441</td>
<td>49</td>
<td>1765</td>
<td>29</td>
<td>2201</td>
<td>21</td>
<td>3924</td>
<td>25</td>
<td>196</td>
<td>703</td>
</tr>
<tr>
<td>7</td>
<td>AJ-B</td>
<td>M</td>
<td>24</td>
<td>2560</td>
<td>110</td>
<td>48</td>
<td>333</td>
<td>47</td>
<td>1781</td>
<td>22</td>
<td>3321</td>
<td>23</td>
<td>4247</td>
<td>22</td>
<td>88</td>
<td>714</td>
</tr>
<tr>
<td>8</td>
<td>MSM-B</td>
<td>M</td>
<td>23</td>
<td>570</td>
<td>110</td>
<td>38</td>
<td>344</td>
<td>41</td>
<td>1625</td>
<td>21</td>
<td>2470</td>
<td>25</td>
<td>3246</td>
<td>19</td>
<td>98</td>
<td>615</td>
</tr>
<tr>
<td>9</td>
<td>BHA-B</td>
<td>M</td>
<td>22</td>
<td>880</td>
<td>110</td>
<td>44</td>
<td>452</td>
<td>59</td>
<td>1571</td>
<td>25</td>
<td>2255</td>
<td>22</td>
<td>3133</td>
<td>17</td>
<td>99</td>
<td>853</td>
</tr>
<tr>
<td>10</td>
<td>MO-B</td>
<td>M</td>
<td>28</td>
<td>1060</td>
<td>110</td>
<td>48</td>
<td>333</td>
<td>44</td>
<td>1550</td>
<td>27</td>
<td>2131</td>
<td>21</td>
<td>3332</td>
<td>22</td>
<td>102</td>
<td>538</td>
</tr>
<tr>
<td>11</td>
<td>AH-B</td>
<td>M</td>
<td>16</td>
<td>2320</td>
<td>110</td>
<td>48</td>
<td>333</td>
<td>38</td>
<td>1254</td>
<td>20</td>
<td>2217</td>
<td>24</td>
<td>3865</td>
<td>20</td>
<td>110</td>
<td>365</td>
</tr>
<tr>
<td>12</td>
<td>JJ-B</td>
<td>M</td>
<td>33</td>
<td>850</td>
<td>110</td>
<td>42</td>
<td>549</td>
<td>45</td>
<td>1098</td>
<td>24</td>
<td>3305</td>
<td>23</td>
<td>4107</td>
<td>21</td>
<td>107</td>
<td>683</td>
</tr>
</tbody>
</table>

**Mean**

- **Fx (Hz)**: 1513.5
- **Fy (Hz)**: 27.5
- **Fz (Hz)**: 2548.9

**sd**

- **Fx (Hz)**: 24.33
- **Fy (Hz)**: 3834.3
- **Fz (Hz)**: 18.417

**sd**

- **Fx (Hz)**: 5.6159
- **Fy (Hz)**: 5.9461
- **Fz (Hz)**: 6.095

---

Table 2: "Sygyt Ltd" results of 12 normal persons.

Comparison of "Sygyt Ltd" and "VoceVista" are presented in Table 3. The variation in % is presented: Fx (Hz) 18%, Fy(hz) 25%, Fz(Hz) 20%.
near 10000 Hz. With “SAS program 9.4” using “the Spearman Rank
Sygyt Ltd” and the “VoceVista” sound analysis software. Notice that the
be seen for the Fx, Fy and Fz, the mean Hz were 1513, 2548 and 3834,
correlation coefficient of the formants around 1500 Hz and 2550 Hz.

Table 3: Comparing “Sygyt Ltd” and “VoceVista”.

The same sound recording was analyzed of the two different
software. The comparison is based on varying time points of F0 for the
“Sygyt Ltd” and the “VoceVista”. Notice that the coefficient of variation is 18-25%.

The fundamental frequency changes over time during the sound
recording. As seen on table 4 it was essential that the results of the
“Sygyt Ltd” and “VoceVista” was evaluated at the exact same
millisecond for the sound recording.

Table 4 shows the difference in formants of F0 between the “Sygyt
Ltd” and the “VoceVista” sound analysis software over 1000Hz. As can
be seen for the Fx, Fy and Fz, the mean Hz were 1513, 2548 and 3834,
respectively with a difference between the software of 0.75, -10.5 and
-259.5 Hz. The difference in percentage of mean shows that the
difference is 0, 0 and -7%, corresponding to almost identical
assessment of the formants around 1500 Hz and 2550 Hz.

As “Sygyt Ltd” results of up to 20000 Hz can be measured, we tried
to study the normal persons for a start for variation of the formants
near 10000 Hz. With “SAS program 9.4” using “the Spearman Rank
correlation coefficient test”, the variation was 4.3% (Table 5).

Table 4: Comparing “Sygyt Ltd”-“VoceVista” at the same millisecond.

As can be seen in the Fx, Fy and Fz, the mean Hz were 1513, 2548 and 3834,
respectively with a difference between the software of 0.75, -10.5 and
-259.5 Hz. The difference in percentage of mean shows that the
difference is 0, 0 and -7%, corresponding to almost identical
assessments of the formants around 1500 Hz and 2550 Hz.

As “Sygyt Ltd” results of up to 20000 Hz can be measured, we tried
to study the normal persons for a start for variation of the formants
near 10000 Hz. With “SAS program 9.4” using “the Spearman Rank
correlation coefficient test”, the variation was 4.3% (Table 5).

As “Sygyt Ltd” results of up to 20000 Hz can be measured, we tried
to study the normal persons for a start for variation of the formants
near 10000 Hz. With “SAS program 9.4” using “the Spearman Rank
correlation coefficient test”, the variation was 4.3% (Table 5).
Table 5: Variation of the formants near 10000 Hz with “Sygyt Ltd”.

STD 441.4833
MEAN 10293.5
Coefficient of Variation (CV) 4.3%

Table 6: Variation of 3 formants from high speed films were of the same % variation.

Results/Discussion

The results show that formant measures are possible routinely. A normal material has been presented with a variation of the 3 formants over 1000 Hz of 18-25%. Since the literature is calling for more evidence based studies, we refer to measuring the formants between 1000 Hz-5000 Hz in pathology. Formant analysis with the program “Sygyt Ltd.” might be a tool not only for qualified singers, but also in pathology [1,2].

From the literature we have essential documentation that evidence based voice diagnosis in pathology is lacking:

Surgical versus non-surgical interventions for vocal cord nodules:

This is a Cochrane review first published in The Cochrane Library in Issue 2, 2012 and previously updated in 2007 and 2009.

Vocal cord nodules are bilateral, benign, callous-like growths of the mid-portion of the membranous vocal folds. They are of variable size and are characterized histologically by thickening of the epithelium with a variable degree of inflammation in the underlying superficial lamina propria. They characteristically produce hoarseness, discomfort and an unstable voice when speaking or singing.

To assess the effectiveness of surgery versus non-surgical interventions for vocal cord nodules we searched the Cochrane Ear, Nose and Throat Disorders Group Trials Register; the Cochrane Central Register of Controlled Trials (CENTRAL); PubMed; EMBASE;
When searching for randomized and quasi-randomized trials comparing any surgical intervention for vocal cord nodules with non-surgical treatment or no treatment (including acoustical measures) no suitable trials were identified. No studies fulfilled the inclusion criteria. There is a need for high-quality randomized controlled trials to evaluate the effectiveness of surgical and non-surgical treatment of vocal cord nodules [3].

Reliability of objective voice measures of normal speaking voices:

The objective was to determine the reliability of objective voice measures used commonly in clinical practice. 18 healthy volunteers (nine males and nine females) were included. Measures of laryngeal efficiency and perturbation measures of fundamental frequency (F0) for both genders were made. For female cepstral peak prominence (CPP) had moderate reliability, whereas for males, the smooth CPP was reliable. Noise-to-harmonic ratios (NHRs) has the lowest consistency of all measures over the course. Additional research is needed to investigate which factors within the testing protocol and/or changes to the measurement instruments may lead to more consistent test results [4].

Evidence-based Clinical Voice Assessment: A Systematic Review.

To determine what research evidence exists to support the use of voice measures in the clinical assessment of patients with voice disorders.

Literature studies provide measurements results of selected acoustic, laryngeal imaging-based, auditory-perceptual, functional, and aerodynamic measures. There is clearly a pressing need for high-quality research that is specially designed to expand the evidence base for clinical voice assessment [5].

Conclusion

Measures of high speed films combined with overtones/harmonics with “VoceVista” and “Sygyt Ltd.” show that they are comparable up to 5000 Hz. “Sygyt Ltd” has a range up to 20.000 – 30.000 Hz. “Sygyt Ltd.” can therefore be used for formant analysis in clinical praxis for at least 10.000 Hz as it was done in the prospective cohort study of 12 normal patients.

The approach in clinical praxis must be simple therefore we suggest using high speed films simultaneously to F0 around 110 Hz in males and 220 Hz in females analysed with “Sygyt Ltd.” as a basis for voice pathology. Formant analysis and high speed films simultaneously could give much more information of treatment effect in voice pathology.

References