Chapter 6

Physiological structure of the vocal folds: Historical perspectives as a basis for HSDI/HSDP

Harry Hollien

Abstract

This chapter is designed to provide certain baseline information and perspectives for how physiological structure of the vocal folds (VF) was addressed in the literature. This information will take the form of 1) a brief history of some of the concepts and relationships which have led to modern HSDI research, 2) a short review of phonatory theory, and 3) baseline data for a model of VF activity and laryngeal functions. In turn, the model can be used to assist the HSDI investigators in interpreting their observations and findings. The review will start with data/concepts from the 20\textsuperscript{th} century. It will do so because much of the baseline research to be discussed was carried out during that period.

Keywords: VF, neurochronaxic, myoelastic and aerodynamic theories, lateral soft-tissue x-rays, VF length, VF thickness, laminagraphy, Stroboscopic Laminagraphy (STROL), F0, Ps

Some history

The early 1900’s can be noted primarily for controversies and competing ideas about voice and laryngeal function. A number of proposals about how the VF operated to create “voice” were advanced and debated. Of course many of these disputes actually did not start during that period, but rather had been going on for many years. They resulted from the efforts of musicians, voice teachers, singers, medical personnel (of many types), and phoneticians. All of these groups were interested in laryngeal functions and a substantial number of them wrote about it. In doing so, they contributed both information and misinformation to the field. Unfortunately, the extent of the misinformation was greater than that desirable.

An example, singers and their teachers sought to describe the registers and ranges of the singing voice by developing such “concepts” as “head” and “chest” register. They based these descriptions primarily on the “feelings” that were created within their body (head) when they sang at high frequencies and in their chest when they sang at lower ones. It was also during this period that voice teachers attempted to improve their teaching by developing a number of systems, learning tasks, and especially rules. Again, many of these efforts actually thwarted an orderly development of the field.

Physicians also participated. They often studied cadavers in an effort to gain useful information. For a perspective here, consult Kahane’s work [1-2], plus his critique as to their success—or lack of it. They also attempted to peer down a patient’s throat in order to observe how the vocal folds appeared to operate and/or find evidence of pathology (by means of a procedure initiated by Garcia [3]. In any event, the field teemed with competing descriptions, explanations, theories, and rules—many of them vague, and others inaccurate. This turmoil continued up to, and past, World War II.
Yet, it also was during that period that a series of advances occurred. First, the Scientific Method [4-5] was proposed and adopted (over time anyway). It resulted in both the upgrading of descriptive research and the development of much more precise experiments. The second shift involved a sharp improvement in technology—especially with respect to acoustic processing equipment (both recording devices and analysis systems). They provided more sophisticated approaches for the assessment and evaluation of many of the unanswered questions about voice. It also was during this period that some of the old descriptions about phonation were discarded and new ones established. While there would be little purpose in reviewing them all, it should be useful to consider the three theories of phonation that were dominant during that period. Specifically, they were the 1) neurochronaxic, 2) myoeastic, and 3) aerodynamic theories.

In its simplest form, the Neurochronaxic Theory [6-7] specified that the vibratory cycles (of voice) resulted from neural pulses serially activating the VF muscles, thereby making them move over and over to (and from) the laryngeal midline. Of course, movement away from midline contact was aided partly by muscle relaxation and subglottal air pressure. To reiterate, each cycle occurred only when the contracting process was repeated—with the entire operation replicated again and again in order to create a vocal tone. Fundamental frequency (F₀) was determined simply by the number of neural activations (of the relevant VF muscles) which occurred each second. To produce phonation with an F₀ of 100 Hz, the neural system had to fire, and the VF muscles contract, 100 times each second. For phonation at 1000 Hz, 1000 contractions a second were required.

When neurophysiologists suggested that these very high rates of muscle contraction were physiologically impossible, the proponents of the Neurochronaxic Theory modified it somewhat. They indicated that the neural pulse would occur in “relays” with a portion of the available muscles (say a quarter of them) contracting for the first cycle, a different but parallel set contracting for the second cycle, and so on. They also argued that while airflow was necessary to create sound, the cyclical movement could occur without it. In this latter instance, of course, no acoustic signal would be created but the VF still would move rapidly. Unfortunately, the proponents of this theory were unable to provide very much data to support their arguments.

The second of the three concepts was the Myoeastic Theory. Its proponents suggest that phonation occurs when the VF are adducted by a general contraction of the relevant laryngeal muscles and then blown apart by subglottal air pressure. The next cycle occurs when the folds are adducted again, but this movement is created essentially by elastic recoil. The process is repeated over and over to create sound. Proponents of the theory also proposed that different F₀ could be produced by changes in the mass and stiffness of the adducted VF, and variation of voice quality and intensity could be accomplished by other VF manipulations (e.g., size, structure, “tension”, etc.)—all of which would result in phonation when coupled to aerodynamic support. Early on, they backed their arguments with descriptions based on anatomical drawings and observations made after manipulating cadaver larynges. These explanations were later supplemented by observations of phonation using laryngeal mirrors (indirect laryngoscopy) wherein they could view the VF in live subjects. Again, appropriate subglottal airflow was integral to the creation of sound. While these relationships were reasonably well documented, the theory’s proponents did experience some difficulty explaining just how certain sounds were made or why air pressure appeared to systematically vary for different types of phonation.

The third, the Aerodynamic Theory, resulted (in part, anyway) from observations made of airflow through tubes. Its proponents indicated that their theory was based on the well-known fact that the velocity of airflow is increased—but its pressure reduced—when air streaming through a tube was impeded by the narrowing of channel diameter—i.e., by the Bernoulli effect [8-9]. In turn, these relationships lead to the proposal that
increase in laryngeal airflow and reduction in pressure—resulting from the VF projecting into the laryngeal tube—would produce a sucking action relative to the folds. This action, in turn, would pull them toward each other. Thus, this theory suggested (at first anyway) that all that was needed for phonation was airflow over the VF as they projected into the airway. They would be sucked together (adducted) until closed and then pushed away from the midline by the increase in subglottic pressure (which occurred once the airway was blocked). The problem here was that this theory did not explain 1) how many of the frequency and quality changes associated with phonation occurred, 2) why different VF lengths, shapes, etc. had been observed, or 3) why changes in VF configuration appeared to be directly associated with different types of voice.

The competition among these three theories, while intense, lasted only a short time. Indeed, data quickly emerged supporting (in part, anyway) the second two theories but not the first. A break in this controversy occurred when Moore [10], and later others, demonstrated that a person could not move their vocal folds in a vibratory pattern unless airflow was present. He did this by making ultra-high-speed motion pictures of subjects with both normal vocal folds and a stoma (i.e., an opening in the trachea) just below them. Using this approach, he demonstrated that phonation would occur if the vocal folds were moved to the laryngeal midline and airflow passed through them. He then demonstrated that sound would be eliminated when the stoma was opened—functionally removing the airflow past the VF by shunting it out through the opening in the neck. That is, his demonstrations showed that when the VF were adducted but the stoma was left open, no movement occurred but when the stoma was covered, VF vibration (and phonation) immediately resulted—and continued until the stoma was again reopened.

These informal studies, plus other evidence, mounted until van den Berg [11-12] summarized the case against the Neurochronaxic Theory of phonation. He did so most effectively, putting an end to any speculation in that direction.

But what of the other two theories? Experiments carried out during that same period supported both the Myoelastic and Aerodynamic theories. Yet, still other findings suggested that phonation could not be explained by either alone. In short, it became obvious that they were simply parts of a single theory. That is, phonation was not possible unless both systems operated and in reasonable synchrony. Of course, even to the present day, some physiologists appear to believe that the Myoelastic system is dominant while, on the other hand, some engineers appear not to be aware of the many physiological operations necessary (by the larynx and VF) if the plethora of different sounds made by the system are to be created. Nonetheless, the great preponderance of thought—as well as nearly all of the relevant data—can be used to argue that the Aerodynamic-Myoelastic Theory (ADMET) provides an appropriate foundation for human voice. Accordingly, it will be employed to provide the framework for the models to follow as well as a sound basis for the other experiments found in this book.

A Perspective

As has been explained, the purpose of this chapter is to provide a model, within the ADMET framework, which can be used as an aid in interpreting the very important findings found in other chapters. The difficulty in developing such a model (or models) is that, since there are many, many dimensions to “voice,” it would be futile to attempt to include them all. For example, normal voice, vocal training (of all types) and voice disorders (due to disease, behaviors, etc.) create but one of the many complex continua that exist; so too do the different registers and ranges of voice as well as the extent and nature of the vocal intensity and quality domains. Indeed, the complexity of phonation and its varying dimensions extends in many directions.
Accordingly, and in order to provide a reasonably comprehensive—and useful—model, certain limitations will be observed here. Specifically, the focus will be on a “core” model rather than a huge all-inclusive one. It will be limited to 1) the normal voice exclusive of any kind of specialty training—or disorders, disease or damage, 2) the adult voice, exclusive of that of children, youths and the elderly, and 3) the modal register only [13]; no attempt will be made to include loft or falsetto on the high frequency end or vocal fry, pulse or mixed voice associated with very low frequencies.

The primary reason for these limitations is that healthy adults phonating in the modal register create, by far, the most common type of phonation which occurs. Of even greater importance is that more research has been carried out here than on any of the other potential continua. Thus, the type of baseline needed for other data sets will be provided. Protocol will be initiated with data on laryngeal size followed by consideration of VF length, VF thickness, and air flow/pressure variation. Many of the experiments in the first three of these segments were carried out by the present author.

**Laryngeal size**

As implied, information about laryngeal size, as related to gender and voice frequency range, is useful in providing a baseline for investigations of VF length, thickness, and mass—especially when they are considered within the complexities of phonation. Accordingly, an investigation [14] was carried out in order to provide such information. As has been suggested, up until the initiation of that project, most attempts to quantitatively relate laryngeal dimensions to gender, age, voice level, etc. had been based on either simple observations or by dissection of cadaver preparations [15-17]. Although the male/female measurements suggested significant gender differences in laryngeal size, little actual data had been provided to show if size variations within each sex also were directly related to the phonatory “dimensions” of voice. Moreover, even those data which were available proved vulnerable to the criticism that may be leveled at virtually all measurements of cadaver preparations: i.e., that they do not adequately represent living tissue. And, certainly, there was/is almost no way by which to discover how even a few of that type of “subject” sounded when producing controlled phonation.

The approach used in this first investigation was to study live subjects by means of lateral soft-tissue x-rays. Its advantages were many; they included use of large—and living—populations, obtaining quantitative measures, providing for comparison of vocal behaviors, and so on. Four groups of six subjects each were selected from healthy volunteers aged 18-29 years (Figure 1). Group LM were six males with low pitched voices; group HM, six with high pitched voices; group LF, six females with low pitched voices and group HF, six with high pitched voices [18-19]. Selection criteria required that subjects be free of speech or voice defects, had received no formal training in singing, and exhibited the ability to perform the phonatory tasks necessary for this, plus other experiments. Also, and as can be seen from Figure 1, the HM and LF subjects had very similar phonatory ranges.

A standard lateral x-ray procedure was adapted to permit accurate laryngeal size measurements. The equipment was a Picker unit; padding and cork blocks were used to position the subject. Standard corrections were employed to compensate for enlargement differences. The size measurements chosen for analysis can be best understood by reference to Figure 2, which is a tracing of the borders of the laryngeal tract as seen on the (lateral) x-ray photographs. The walls of the laryngeal pharynx were outlined as seen and a variety of size measurements made. Differences (Table 1) among the subject groups were statistically significant (re: ANOVA’s) with all but two critical differences also significant.
It could be concluded from this research that individuals with low pitched voices can be expected to have larger laryngeal tracts and more massive vocal folds than individuals exhibiting higher pitch ranges. It also was shown that gender differences constitute an
even more powerful variable than does pitch level; i.e., males will have larger laryngeal structures than females even though they may have nearly identical pitch ranges. Data were not provided, however, about VF size/activity, or how F0s usage (“habitual” pitch level, for example) was accomplished. However, data on many of those relationships will be found in subsequent sections.

**VF length**

The next step was to assess the dimensions of the VF (as opposed to general laryngeal size) as a function of gender, modal pitch range and phonatory activity. To do so, the present author [20] first carried out relevant research on the same group of subjects studied in the preceding experiment. Subsequently, other groups and other relationships were addressed.

**First investigation of VF length**

This experiment involved the same population as that cited immediately above. In this instance, however, subjects had their VF photographed when they were abducting them and then when phonating at low (10%), medium (25%), and high (50%) F0 levels within the modal register—plus falsetto at 85%. However, since falsetto is outside of the purview of this research, it will not be considered in this review.

Until the time of this experiment, no essentially comprehensive investigations had been carried out directly on just how pitch (F0) changes were accomplished. Some authors appeared to feel that they were achieved by a lengthening of the VF, whereas Negus [17] and others argued that such changes were not possible. On the other hand, Moore [21-22] and Farnsworth [23] had reported observing some sort of lengthening-frequency shifts in their ultra-high-speed motion pictures [24]. Furthermore, both Irwin [25] and Brackett [26] also had sometimes found similar, if somewhat varying, increases.

Of course, individuals attempting to conduct research of this type faced a number of serious challenges. Included among them were problems such as actually visualizing the VF, then accurately photographing them and finally measuring them with reasonable precision. At issue, especially, was variation in “lens-to-field distance”—i.e., that distance between the folds and the photographic film. This variable will not only fluctuate because of differences in the anatomical size of an individual, but also, because (as had been observed) the larynx sometimes is elevated with variation in vocal pitch. Since the size of the photographic image is a function of these distances, it follows that measurements on laryngeal films could be subject to error. That is, the reported data could be error prone if the variation in lens-to-folds distance is not known, and suitable corrections made.

These problems were addressed [20] by using updated systems, new measurement techniques, and by training the subjects to better expose their VF as they tolerated the attendant discomfort. The equipment consisted of a mounted laryngeal mirror, a parabolic head mirror (directing light), a 500-watt light source with a condensing lens and prism, and an Eastman Cine-Kodak 16-mm motion picture camera with a four-inch telephoto lens. A motion picture camera was used to permit selection of the optimum exposures for measurement. Control of subjects’ F0 was obtained by means of a reference tone at the required frequency level provided by an ordinary chromatic pitch pipe.
Figure 3. A tracing of the vocal folds and larynx as seen in laryngoscopic photographs. E is the epiglottis; F, the vocal folds; and T, the tubercles formed by the corniculate and arytenoid cartilages. Line A is drawn tangent to the most anterior extent of the vocal folds; and line B, tangent to the tubercles, is considered to be the posterior boundary of the vibrating folds. All measurements (in mm) were made from A to B.

Measurements were made by projecting the selected frames onto a surface with a microfilm projector. Figure 3 provides a drawing of the image most often available. Measurements were made of the maximum anteroposterior extent of the VF. These values were later validated by re-measures based on the upgraded protocol developed for the second VF length study (see below). Correction procedures for this set of data were established by obtaining the distance of the VF to the laryngeal mirror from the x-rays available for each subject (from the first study) and then photographing a millimeter grid at those identified distances. A table of corrections from the projected images of these grids was developed and applied.

Of the original group, only 16 subjects (i.e., four from each group) were able to successfully complete this procedure and, unfortunately, even these data could not be obtained for 17% of the experimental conditions. The problem here resulted from the subject not being able to tilt the epiglottis far enough forward to expose the most anterior extent of the folds. Data from the acceptable views can be found in Table 2. First, it should be noted that the VF size varied for the groups. That is, average VF length descended from low male to high male to low female and finally to high female. These measurements are both in substantial agreement with ADMET and the findings for general laryngeal size. In addition, the following relationships were noted:

a) As F0 of phonation was raised, the VF were systematically lengthened.

b) In the abducted position (i.e., at rest or when the subject was breathing), the VF were longer than for any condition of modal register phonation.

c) Low pitched individuals exhibited generally longer VF than did individuals with higher pitch ranges; this relationship was found both between sexes and within a sex.
Table 2. Measurements of vocal fold length for 16 subjects during abduction and phonation in the modal register. All values are group means—corrected for variation in film-to-fold distance. They are reported in millimeters.

<table>
<thead>
<tr>
<th>Fundamenta Frequency</th>
<th>Group</th>
<th>N</th>
<th>Abducted</th>
<th>Low</th>
<th>Middle</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low Pitched Males</td>
<td>4</td>
<td>19.5</td>
<td>10.3</td>
<td>11.3</td>
<td>15.2</td>
</tr>
<tr>
<td></td>
<td>High Pitched Males</td>
<td>4</td>
<td>14.5</td>
<td>9.9</td>
<td>10.3</td>
<td>12.9</td>
</tr>
<tr>
<td></td>
<td>Low Pitched Females</td>
<td>4</td>
<td>11.3</td>
<td>7.6</td>
<td>8.2</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>High Pitched Females</td>
<td>4</td>
<td>10.0</td>
<td>6.5</td>
<td>6.9</td>
<td>8.5</td>
</tr>
</tbody>
</table>

Due to the limitations in this, the first experiment, an additional series of investigations, employing even more sophisticated techniques, were carried out. Specifically, the issues addressed were: 1) can a more accurate measurement procedure be developed, 2) can the finding of VF lengthening, in association with rise in fundamental frequency, be validated, 3) can the nature of these trends (e.g., slopes, intergroup relationships) be specified, and 4) can the data indicating that they are shortened for phonation (relative to their length during abduction) also be authenticated? This last finding (i.e., the VF being shorter for phonation than when at rest) was somewhat unexpected. Indeed, it contradicted many long-held opinions about phonation. One, in particular, was that for phonation, the VF are adducted and stretched [27-29]—a postulation which originated in the 19th century. In any event, it appeared necessary to obtain more comprehensive data.

The Second length investigation

Data from this experiment were acquired by means of color motion picture photography [30]. These investigators had developed more sophisticated techniques (including color) for viewing the vocal cords and measuring them. Subjects here were six men ranging in age from 27 to 53 years who were able to expose the full anteroposterior extent of their vocal folds while phonating fundamental frequencies encompassing their modal range. Three subjects were classified as bass or baritone, and three as tenors. All had histories essentially free of laryngeal pathology and/or formal singing training. The photographic exposures were made with an Arriflex 16 motion picture camera with a six-inch telephoto lens mounted on an extension bellows. The lighting system consisted of a 2000-watt incandescent lamp, a water cell cooler, and a focusing lens (see Figure 4).

During the experiment each subject was requested to phonate at the musical tones of C, E, and A within all possible octaves from the lowest to the highest tones sustainable in their modal register. Vocal pitch was controlled by having subjects match reference tones provided by an audio-generator coupled to a speaker system. No control of vocal intensity was attempted other than to request that the subjects produce all pitches at a “comfortable” loudness level.
Figure 4. The laryngeal photographic system used in this research. The cameraman is observing the subject’s vocal folds through the lens/bellows system of the Ariflex motion picture camera. The subject has placed the fixed laryngeal mirror in the back of her oral pharynx. That mirror is fixed so that (later) the grid, placed on a rack and pinion gear can be photographed when in position. The structures in the background house the light, water cell cooler and lens.

The challenge here was to obtain demonstrably accurate measurements of VF length. A new procedure was developed in order to do so. It was made possible by exploitation of certain features of the Arriflex 16 camera. That is, with a telephoto lens and extension bellows, this camera provided both a magnified image of the folds and one at a very shallow depth of field. Thus, it was possible to bring the photographic image into very accurate (and shallow) focus and, then later, bring a millimeter grid to that same level, by means of a rack and pinion gear, while the focus remained constant. It was possible to do so (individually) for each experimental condition. Photographs were then made and used to directly measure the length of the VF for that specific condition. Error due to variation in the distance from the VF to the film (caused by the relative elevation of the folds for different pitches and differences in subject size) were all corrected by a single process.

A number of measurements could be made (see again Figure 3). Included was the entire anteroposterior length of the VF from the anterior commissure to the posterior commissure. However, as with the first length investigation (see Table 2), the measurements employed in this and subsequent research on this subject were made from the most anterior point of the folds (A) to a line drawn tangent to the anterior borders of the tubercles formed by the protrusions of the corniculate and arytenoid cartilages (C). These values thereby parallel those from the first and subsequent investigations. Best yet, they delineated the length of the vibrating portion of the VF—and they were found to provide a consistency that was much superior than those for prior measurements. Accordingly, a single measurement set was again chosen for evaluation; please see Table 3. The findings:

a) Excepting for a minor enlargement difference, the data and trends exhibited a close relationship to those from the preceding study.

b) The length of VF increases in concert with vocal pitch for the modal register.
c) The VF in abduction are always longer than for any phonatory condition in the modal register.

d) VF length appears to coordinate rather well with pitch level.

e) These data clearly support ADMET.

**Table 3.** Means of vocal fold length (in mm) of two male groups of N=3 each. Selection was based on frequency range with one group consisting of three subjects with low pitched voices and the other three exhibiting high pitched ranges.

<table>
<thead>
<tr>
<th>Semitone Level</th>
<th>Hz</th>
<th>Baritones</th>
<th>Tenors</th>
</tr>
</thead>
<tbody>
<tr>
<td>E2</td>
<td>82</td>
<td>11.6</td>
<td>-</td>
</tr>
<tr>
<td>A2</td>
<td>110</td>
<td>12.8</td>
<td>9.8</td>
</tr>
<tr>
<td>C3</td>
<td>131</td>
<td>13.4</td>
<td>11.7</td>
</tr>
<tr>
<td>E3</td>
<td>165</td>
<td>15.7</td>
<td>12.8</td>
</tr>
<tr>
<td>A3</td>
<td>220</td>
<td>19.4</td>
<td>13.7</td>
</tr>
<tr>
<td>C4</td>
<td>262</td>
<td>-</td>
<td>14.0</td>
</tr>
<tr>
<td>Abduction</td>
<td>20.9</td>
<td>19.4</td>
<td></td>
</tr>
</tbody>
</table>

It also should be noted that several other investigators independently demonstrated that the method employed in this research—and the findings also—were valid. First, Wendler [31] directly studied—and verified—both the procedure and data. The second “validation” involved a lateral soft tissue x-ray approach [32] carried out at the present author’s laboratory but by a research team led by a guest investigator. Again, the data confirmed the lengthening of the VF with rise in phonated F0. Finally, yet further confirmation (although somewhat tangential) was provided by Sonninen [33-34] and Luchsinger and Pfister [35]; reliability also was established by the next study and by Hollien et al. [36].

**Third VF length investigation**

In order to further validate the obtained findings and provide parallel data for females, a replication of the above research was conducted on 10 adult women [37] ranging in age from 18-32 years. They were required to meet all of the selection criteria as did the subjects in the previous two investigations. Moreover, in order to provide parallel data for the other studies, only those four with the lowest pitched voices and the four with the highest were used in the present analyses—and, of course, only data obtained from the modal register was considered. Finally, protocol for acquiring the images and the measurement of the VF, followed exactly those for the immediately preceding experiment on men (see again Figures 3 and 4).

The results may be found summarized in Table 4. As can be seen, the data there follows the same patterns as did those from the first two studies. Basically, 1) systematic VF lengthening occurred with rise in F0, 2) the folds are shorter for all modal register phonation than they are for abduction, 3) the length values parallel those for women in the first study and 4) the data supported ADMET.
Table 4. Means of vocal fold length (in mm) for two groups of females (N=4 each). Selection was based on frequency range; hence, one group was of four subjects with low pitched ranges and the other with high. NM = No Measurement (i.e., insufficient data).

<table>
<thead>
<tr>
<th>Semitone Level</th>
<th>Hz</th>
<th>Alto</th>
<th>Soprano</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3</td>
<td>131</td>
<td>NM</td>
<td>-</td>
</tr>
<tr>
<td>E3</td>
<td>165</td>
<td>7.6</td>
<td>7.1</td>
</tr>
<tr>
<td>A3</td>
<td>220</td>
<td>8.9</td>
<td>7.6</td>
</tr>
<tr>
<td>C4</td>
<td>262</td>
<td>9.8</td>
<td>8.5</td>
</tr>
<tr>
<td>E4</td>
<td>330</td>
<td>10.4</td>
<td>9.4</td>
</tr>
<tr>
<td>A4</td>
<td>440</td>
<td>10.8</td>
<td>10.3</td>
</tr>
<tr>
<td>C5</td>
<td>523</td>
<td>11.5</td>
<td>11.0</td>
</tr>
<tr>
<td>E5</td>
<td>659</td>
<td>-</td>
<td>NM</td>
</tr>
<tr>
<td>Abduction</td>
<td>14.3</td>
<td>13.7</td>
<td></td>
</tr>
</tbody>
</table>

Mechanics of VF lengthening

The data on VF lengthening obtained from the studies cited above will not be combined and discussed until a brief review is provided about their relevant operational mechanisms. The primary issue to be addressed is on how VF lengthening is accomplished.

On the surface anyway, the basic process appears to be a simple one. That is, once adduction is accomplished—primarily by contracting the thyro-arytenoid (TA) muscles (especially the vocalis) in concert with interarytenoids and lateral cricothyroid—all frequency related lengthening results from systematic contraction of the crycothyroid (CT) muscle. This position has been held for some time and by most authors. Yet, it has been shown that the entire extent of the lengthening process cannot be achieved by CT activity alone. Specifically, data published by the present author and his associates [38-39] indicate that, while the CT configurations they tested provided for nearly all of the lengthening at the low and middle (phonated) frequencies, it could account for only about 70-80% of the actual lengthening which occurred at the upper modal frequencies. Accordingly, it is proposed that this additional lengthening is accomplished by contraction of the interarytenoid muscles and the posticus. While such contractions have been reported by Gay et al. [40], somewhat mixed data has been published by others [41-44]. On the other hand, Shipp [45] suggests that some evidence of this activity occurred at the upper frequencies and, also, that some of the relevant activity quite possibly could have been missed due to the difficulty in accurately placing the hook-wire electrodes downward into the posticus.

The proposed mechanism also is validated by von Leden and Moore [46-47] who demonstrated the anterioposterior movement of the arytenoids along the track provided by the cricoid. In any event, VF lengthening appears to be primarily accomplished by the TA muscle operating in opposition to the CT muscle supplemented by the posterior arytenoid muscles.
Interpretation

The task now is to merge all the length data (two sets for males and two for females) in order to create a model of VF lengthening as related to frequency (F0) change. The resultant curves are based on data from seven males exhibiting low pitched voices and seven with high pitched ones (two studies) plus eight low pitched and eight high pitched females. The sum total was of three studies populated by 14 men and by 16 women which provided 122 specific data points in all.

![Figure 5. Variation in vocal fold length with change in vocal pitch. Mean vocal fold lengths of the subjects are plotted in mm. Relative frequency level values—in the modal register only—are calculated as percentages of the subject’s total pitch range above his lowest sustainable tone.](image)

The combined data then result in the curves found in Figure 5. These curves—i.e., those for two each male/female groups at two levels—can be seen from LM (top) to HF (bottom). As can be noted, most data points are close to a straight line for that gender and pitch level, and the trends show a systematic increase of VF length with rise in pitch. Also, please note that the angle of the slopes increases with laryngeal/VF size. From these data, it is clear that change in fundamental frequency in the modal register is mediated by a lengthening of the VF. Since the total mass of the VF should not vary, this pattern (logically anyway) should result in their thinning. The next step, then, was to determine if this is what actually happens.

Research on VF thickness

The next objective, then, was to study the cross sectional mass—or thickness—of the VF as a function of ΔF0 in the modal register. Prior to 1954 when the first of these studies were initiated, only Griesman [48] had used the laminographic procedure (referred to as planigrams) to study laryngeal phenomena. He presented and discussed data of this type obtained from several singers while they were producing a variety of vocal tones. His work suggested that systematic trends might exist (especially with respect to fundamental frequency) in the coronal cross-sectional size of the VF. Later, others employed this
same technique. However, they did so to either study voice pathology [49-50] or unusual phonation—ordinarily that which was related to singing—rather than for basic research purposes [51-53].

The First VF thickness experiment

The objective of this research [54] was to determine if such VF dimensions as size, area or thickness would systematically vary with changes in phonatory behaviors—especially with respect to fundamental frequency. This investigation employed the same 24 subjects who populated the experiments of laryngeal size and the initial one on VF length. In the present instance, however, the procedure used was laminographic x-ray (previously referred to as “planigraphic” and now as “tomographic”). The unit employed was a Keleket Selecto-plane laminographic x-ray unit. Subjects were placed supine on the x-ray table with a head immobilizer and sponge rubber padding positioned to ensure proper placement (and comfort). A marker system permitted 1) the VF to be located accurately and 2) to have the x-rays made on a plane which passed approximately through the midpoint of their anteroposterior length. Exposures also were taken at either 1 cm (women) or 2 cm (men) both anterior and posterior to the midpoint plane.

Subjects were required to phonate at the same three F0s as they did in the first VF length study (i.e., at 10%, 25% and 50% of their total range). Again, no data on the falsetto register is included. However, in this case vocal intensity also was controlled. The procedure here required all subjects to produce three samples of each of their experimental frequencies at a “comfortable” level. These levels were measured and the mean calculated. A voice level monitor with two neon reference lights was then used to provide intensity cues to the subject during the experiment. That is, one of the neon lights was programmed to glow when the subject reached a vocal intensity within 2 dB of that desired. If he/she exceeded this level by 2 dB or more, both lights would glow. Once the subject was in position, with the proper x-ray settings and subject output established, laminagrams were made at each of the experimental F0 levels.

Measurements were attempted only on film made during phonation as those made of the nonphonation or rest conditions (i.e., with the VF abducted) did not show them. That is, they usually were not distinguishable from the lateral laryngeal walls. An example of the experimental material may be seen in Figure 6.

The measurements which were made will be described by reference to Figure 7. Of the several, two are most relevant to this review; they were: 1) the area of mesial VF projection into the airway and 2) mean VF thickness. Area is shown as the shaded portion on the figure and mean thickness was obtained by dividing area by the maximum distance from A-B to D-E, parallel with C). Since these measurements are common both to this study and all of those to follow, they will be described in additional detail. That is, in Figure 7, the VF are identified by the letter F; the ventricular folds by V, and the trachea by T. Lines A-B define the mesial borders of the laryngeal tract; line C the superior surface of the vocal folds, and lines D-E constitute the standard reference line included to close the area of the folds laterally; please note line D-E was shown not to vary significantly from condition to condition. The reported values were obtained by taking the individual measurement for each fold (i.e., for either area or thickness) and averaging them to provide a single value.

Most of the findings from this experiment were expected; all were statistically significant. Basically, it was found that individuals with low pitch levels exhibit larger, more massive VF than did individuals with higher pitch levels—so did male size exceed that for females. Second, as the fundamental frequency of an individual’s voice was raised, the VF were reduced in cross-sectional area and became thinner.
What was not expected was the high correlation between VF thickness and absolute F0 of phonation (see Figure 8). As can be seen, the thickness of the VF appears to be substantially similar at each F0 no matter if the subject was male or female or had a high pitched or low pitched voice. Thus, it appears that the per-unit-mass of the VF relates to the frequency produced no matter how massive (or not) these structures are naturally [55]. This finding was pursued further and for two reasons. One was to assess its validity and the other to determine if this change is in symbiotic relationship with the lengthening patterns cited above for parallel phonatory conditions. Indeed, if lengthening and thickness combine to provide a common structure for specific frequencies, they then provide 1) robust evidence of the accuracy of ADMET and 2) a basis for the calculation of VF mass. They also should lead to a possible model. In any event, the first step in this quest was to evaluate the validity of the hypothesis that VF thickness is directly related to the specific F0 phonated.
The Second VF thickness study

As stated, the goals for the second experiment [56] were primarily to test the hypothesis that the thickness of the VF correlates directly with the F0 produced by an individual and does so more-or-less irrespective of laryngeal size or gender.

Subjects for this investigation were three males and three females ranging in age from 19-33 years. The usual selection criteria were applied and each subject’s pitch range was then obtained by standard procedure [36]. In order to create a group as heterogeneous as possible, subjects were chosen so that low, medium and high pitched voices were represented within each gender.

All equipment used was virtually identical to that employed in the first experiment of this kind with one exception. In this instance, special Auer film packs, which featured multiple x-ray plates, were used. This procedure allowed five coronal planes (5 mm apart) to be made simultaneously along the anteroposterior dimension of the VF. Thus, in turn, it permitted measurements to be made along (nearly) the entire A-P length of the VF.

Subjects were required to phonate six pitches within their modal register; they were 123, 147, 165, 220, 262, and 294 Hz for the men; 220, 262, 294, 349, 392, and 440 Hz for the women. Please note that 1) the only three frequencies which overlapped the male-female groups were 220, 262, and 294 Hz and 2) any F0 that was phonated in falsetto was not included in this project.

Subjects were placed supine on the laminagraph table. Sponge rubber padding, a head immobilizer, and a chest strap were used to ensure maintenance of appropriate positioning—and do so with a minimum of subject discomfort. Each subject was placed so that the central ray of the equipment would pass through his or her larynx at the level of the VFs.

As with the earlier study, it was considered important to make measurements on a coronal plane which passed approximately through the anteroposterior midpoint of the VF—and do so for all subjects. In this case, however, it was also desirable to attempt to determine if the thickness of the VFs varied, or did not vary, along their entire A-P length. Both of these controls were accomplished by utilization of the cited Auer multi-film pack.

Once the subject was in position and all proper settings/outputs determined, a laminagraph was made. The procedure was continued until all experimental conditions were...
satisfied. The measurements were made as per the first thickness study, excepting they were made from photographs of the x-ray plates.

The obtained findings confirmed the high correlation between VF thickness and F0 (see Figure 9). Indeed, a correlation coefficient of -0.91 was obtained. This robust correlation can be considered to be even more remarkable when the population studied is considered. That is, it must be remembered that the six subjects formed a heterogeneous group both with respect to gender, physical size, laryngeal size, pitch range, and so on. Thus, such small variation in the frequency-thickness relationship was somewhat unexpected. Moreover, it should also be noted that VF thickness varied but little over much of its anteroposterior length.

To expand the available database, yet another experiment was carried out [57]. The purposes of this investigation were to: 1) further test the hypothesis cited above, 2) study the obtained curves in more detail, 3) investigate the intensity-VF thickness relationships and 4) study VF thickness at falsetto frequencies. Of course, only the F0 thickness data for the modal register (i.e., goals 1 and 2) will be included in this review.

In any event, data for 10 additional subjects (in the modal register) could be added to the 30 individuals populating the first two studies. It now appeared that enough data were available to combine the length and thickness measures in an attempt to generate the basis for a model of VF operation relative to frequency change in the modal register.

**Evaluation of potential error sources**

However, an effort in the direction cited above was judged to be somewhat premature, and for three reasons. First, concern had been expressed that the effects of gravity might distort the laryngeal structures when subjects were placed in the supine position (rather than vertically) for research of this nature. Second, it had not yet been fully demonstrated that the VF exhibited the same thickness throughout their entire A-P dimension for all experimental conditions. Third, concern also had been articulated relative to the possibility that variation of VF thickening at lower frequencies and thinning at higher ones might possibly be due (at least in part) to the complexity of their movement rather than just to their lengthening and thinning. In this regard, it must be remembered that each laminographic x-ray took a full second to complete. Hence, the resulting picture consists of a number of complete vibratory cycles ranging from a few hundred, to many hundred,
repetitions. It was argued that it could be possible that the lack of stiffness at the low frequencies would result in the folds moving in a complex fashion for those frequencies but in a much simpler pattern for the higher ones. If that were true, they would appear larger at the lower ones than for the higher.

The first concern was resolved by Curtis and Hollien [58] and Hollien [59]. Identical trends for the relationships between VF cross-sectional size and voice pitch were found for subjects studied in both the vertical and supine positions under what were otherwise identical conditions. Accordingly, it was judged that valid research of this type can be carried out with subjects supine.

The second issue, thickness vs. A-P length, was addressed by the use of the Auer cassette which features multiple x-ray plates (5 mm apart). First, a random sample of 25% of all experimental conditions was carried out as part of the second experiment. The data in that instance were validated by an independent and random analysis of 30% of the plate sets generated by a third study. While these x-rays covered the entire A-P length in only 80% of the cases, little-to-no variation was found among any of the measurements. Specifically, the mean of the plate-to-plate differences was only 2.6%. Hence, the second concern was answered. That is, it appears that vocal fold thickness varies but little along their entire anterior-posterior dimension.

**Stroboscopic laminography (STROL)**

The third concern was triggered by the fact that each laminographic image of the VF was created by x-raying a number of vibratory cycles. Thus, it was suggested that they might appear larger for conditions where their movement was complex (i.e., at low frequencies) rather than when it was relatively simple (as when high frequencies were produced). That is, if the outline on an x-ray photograph was made from the blurring of many vibratory cycles, measurements could have been made on an image which did not show them in any single position. Indeed, both the shape and the size of the obtained shadows could be affected by such dynamic factors as amplitude of motion, relative duration of one or more phases, and so on. In short, it was argued that conventional laminagrams could be contaminated by the inadequacies of that observational technique and the validity of vocal fold cross sectional thickness measurements questioned.

Stroboscopic laminograph (STROL) was developed as a response to this rather serious problem [60-61]. This system (Figures 10a and 10b) provides a series of laminographic x-ray photographs showing coronal views of the vocal folds (usually N=10) when they are at each of several phases distributed equally throughout the vibratory cycle. Although each picture is a composite of several short exposures obtained from a number of vibratory cycles, they occur only when the folds are in exactly the same position. Thus, in practice, the folds are seen in 10 different positions throughout a vibratory cycle and accurate measurements of area and thickness can be performed on them.

As would be expected, this technique also is capable of providing motion pictures of the vibrating VF in the coronal cross-section [62]. However, in this review, the material provided by STROL will be employed primarily to validate the accuracy of the measures of VF thickness obtained from ordinary laminography.

Since STROL was a new technique, the characteristics are described below (Figure 10). Basically, it uses an x-ray which is both powered and controlled by a Keleket Noveomatic 500 console. To do so, that unit features an internal pulse-forming network which converts an experimental subject’s sustained phonational output (i.e., his or her F0 when at a fixed frequency) into sharp pulses. These pulses, in turn, are amplified and used to drive the x-ray tube. The x-rays produced by these circuits pass through the subject and impinge on an image intensifier. In turn, a 35 mm Nikon camera is placed facing the intensifier screen; it takes a 0.5 s exposure during each swing of the yoke and advances
one frame at the end of each swing. To reiterate, the x-ray is flashed at the same rate as the subject’s phonated fundamental frequency during each laminographic swing with its phase advanced 36° at the end of each arc. Thus, with the 10 steps in each complete circuit, a full 360° phase shift sequence is accomplished to produce the 10 photographs found in Figure 11. Seen there are the 10 sequential positions of the vocal folds progressing from their closed position through a complete vibratory cycle and, then, back to closure.

**Figure 10.** Block diagram of STROL of the laminographic drive, x-ray and optical systems of STROL. The components include: x-ray tube, image intensifier tube, 35 mm camera, TV camera, yoke drive, and yoke stand.

**Figure 11.** An illustrative series of stroboscopic laminagrams showing the vibratory action of the VF in coronal cross-section. As seen, the VF are closed in the first photograph; they then open, progress through a full vibratory cycle, and ultimately move toward closure—which can be seen in the last (lower right) photograph.
Measurements of VF thickness for four subjects producing seven frequencies [63] can be found in Figure 12. Note that both their actual thickness and the slope of their thinning (as a function of increased F0) almost exactly parallel similar data for ordinary laminagrams. Hence, the third concern—that of potential errors due to varying and/or complex VF movement—has been met and the data specifying a reduction in VF per unit mass (i.e., thickness) as a function of change in F0 may be considered accurate.

Yet the finding of the extent of the per unit mass does not address the issue as to whether or not total VF mass varies in some manner while all the changes reviewed above are taking place. The question to be asked then is: does vocal fold mass vary as a function of F0 changes in the modal register?

**VF mass**

The measurements of the VF already carried out (and reviewed in the cited references) permit reasonably good estimates of total VF mass to be calculated. That is, if the cross-sectional area—or thickness—of the folds is known, and their length also determined, mass will be the sum of the first multiplied by the second (M = L x T). Moreover, if these two values can be calculated for specific frequencies throughout the entire extent of the modal register, actions that cause changes in mass—and when they occur—can be identified.

Specifically, data of this type can be developed simply by using the experiments described in this report. That is, they first may be calculated for the eight males and eight females who populated the first of the length and thickness experiments. Following this, an independent evaluation can be carried out on two groups of matched males and the two similar groups of females.

The first set of comparisons can be found in Table 5 (males) and Table 6 (females). Note that but little variation of total VF mass occurs over the measures at the three frequencies. The same is true for the two additional sets of matched subjects (see Table 7 for the males; Table 8 for the females). Not only are the total VF mass values found for the second sets of individuals throughout the 5-6 F0 conditions, but they also are very much like those found for the initial groups (note the closeness of the mean values). It also can be observed that the size measurements are different between genders and between the pitch ranges. Conclusions based on these data appear obvious. VF mass does not
vary (or, at least, varies but little) as a function of the fundamental frequency produced within the modal register. In addition, it can be noted that these are the first compilations of vocal fold mass measurements to be reported. Finally, it should be pointed out that the relatively small variation among the data points in the mass trends probably is mostly due to the difficulty in making precise measurements of data such as these. Indeed, the nature of the structures studied and the procedures employed in collecting the images, sometimes are not as clear as would be desirable. Hence, minor error in judgment of where a border exists can result in the small variations seen in the VF mass data.

Table 5. Total VF mass measurements for 8 subjects from two categories (low and high pitched males). Values were calculated by multiplying each subject’s VF length by VF area at the three experimental frequencies. All values are group means in millimeters.

<table>
<thead>
<tr>
<th>Frequency in Hz</th>
<th>Low Pitched Males</th>
<th>High Pitched Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>566</td>
<td>-</td>
</tr>
<tr>
<td>110</td>
<td>524</td>
<td>370</td>
</tr>
<tr>
<td>156</td>
<td>-</td>
<td>320</td>
</tr>
<tr>
<td>185</td>
<td>577</td>
<td>-</td>
</tr>
<tr>
<td>262</td>
<td>-</td>
<td>348</td>
</tr>
<tr>
<td>Mean</td>
<td>556</td>
<td>346</td>
</tr>
</tbody>
</table>

Table 6. Total VF mass measurements for 8 subjects from two categories (low and high females). Values were calculated by multiplying each subject’s VF length by area at the three experimental frequencies. All values are group means in millimeters.

<table>
<thead>
<tr>
<th>Frequency in Hz</th>
<th>Low Pitched Females</th>
<th>High Pitched Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>151</td>
<td>220</td>
<td>-</td>
</tr>
<tr>
<td>208</td>
<td>221</td>
<td>-</td>
</tr>
<tr>
<td>226</td>
<td>-</td>
<td>156</td>
</tr>
<tr>
<td>311</td>
<td>-</td>
<td>138</td>
</tr>
<tr>
<td>349</td>
<td>233</td>
<td>-</td>
</tr>
<tr>
<td>523</td>
<td>-</td>
<td>146</td>
</tr>
<tr>
<td>Mean</td>
<td>225</td>
<td>147</td>
</tr>
</tbody>
</table>

Subglottic air pressure (Ps)

The final step in this research program was to determine if the relationships established by the many experiments already carried out were consistent with the predicted aerodynamic reality. After all, it now appears clear that the VF are first contracted for phonation by moving them to the midline by means of the thyroid-arytenoid (TA) muscle system operating in opposition to the other (intrinsic) laryngeal muscles. As attempts are made to increase absolute frequency, the TA contracts yet further with the cricothyroid and posticus (and related muscles) also further contracting in opposition. The end result is a continual thinning of the VF. In turn, they stiffen (as they thin) and air flow impedance is increased. Theoretically, air flow itself might not increase markedly but, it would be expected that subglottic pressure would do so. That is, it would if the conditions specified by ADMET are to be met.
Table 7. VF mass as a function of frequency for six male subjects. Values are vocal fold length multiplied by VF area. The data for the first value (i.e., length) was obtained from three low pitched males and three with high pitched ranges (see Table 3). The area measurements were made on a second set of low and high pitched males. The second group of subjects (i.e., re: area) were actually comparable to the first but they were populated by different individuals. Values are in mm.

<table>
<thead>
<tr>
<th>Frequency in Hz</th>
<th>Low Pitched Males</th>
<th>High Pitched Males</th>
</tr>
</thead>
<tbody>
<tr>
<td>82</td>
<td>575</td>
<td>-</td>
</tr>
<tr>
<td>110</td>
<td>584</td>
<td>380</td>
</tr>
<tr>
<td>131</td>
<td>525</td>
<td>336</td>
</tr>
<tr>
<td>165</td>
<td>551</td>
<td>360</td>
</tr>
<tr>
<td>220</td>
<td>574</td>
<td>316</td>
</tr>
<tr>
<td>262</td>
<td>-</td>
<td>365</td>
</tr>
<tr>
<td>Mean</td>
<td>562</td>
<td>351</td>
</tr>
</tbody>
</table>

Table 8. VF mass as a function of frequency for eight females. Values are VF length multiplied by vocal fold area. Again, the data for VF length was obtained from two sets of four women each (low pitched and high) and the area values from a second set of closely comparable women also with low and high pitched voices. Values are in mm.

<table>
<thead>
<tr>
<th>Frequency in Hz</th>
<th>Low Pitched Females</th>
<th>High Pitched Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>156</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>165</td>
<td>225</td>
<td>141</td>
</tr>
<tr>
<td>220</td>
<td>222</td>
<td>128</td>
</tr>
<tr>
<td>262</td>
<td>219</td>
<td>136</td>
</tr>
<tr>
<td>330</td>
<td>210</td>
<td>117</td>
</tr>
<tr>
<td>440</td>
<td>233</td>
<td>145</td>
</tr>
<tr>
<td>523</td>
<td>207</td>
<td>128</td>
</tr>
<tr>
<td>554</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>219</td>
<td>133</td>
</tr>
</tbody>
</table>

It was not necessary for the present author to embark upon an airflow/pressure series of experiments because research in this area already had been carried out and reported. First, the relevant studies provide little, if any, data that specified an airflow-frequency relationship for the modal register—that is, except (perhaps) for a small increase at the very top of the frequency continuum [42,65]. On the other hand—and as postulated—these authors, plus Ladefoged [66], provide data that indicates that, except for very low modal register frequencies, subglottic pressure (P_s) increases systematically with rise in F0. This relationship pretty much completes those being established and provides the relevant data about operations necessary for vibration in the modal register. They also provide the final link for development of a model. Incidentally, these findings are further confirmed by data on trends in vocal fold tilting as a function of increase in fundamental frequency [67]. That is, if subglottic pressure is increased at the higher frequencies in the modal range (and the VF are thinner), compensatory activity would be expected with the folds pushed upward as they opened in the vibrating cycle. This tilting
would be expected to be muted somewhat with respect to the thicker folds at lower F0. In any event, those were the patterns reported.

A Model of VF action in F0 control

The model which results from the cited corpus of experiments turns out to be a relatively simple one. It can be articulated as follows. Voice frequency change in the modal register is accomplished by a systematic thinning of the VF with the thickness related directly to the actual F0 phonated. This relationship appears to hold no matter what the size or gender of the individual phonating—or the basic size of their VF. This operation is mediated by a continual, but orderly, fold lengthening.

The mechanism that supports the cited processes also is a fairly straightforward one. That is, the vibratory portion of the VF is adducted and stretched to systematically reduce their cross sectional area. They do so without violating the extent of their mass—which remains constant. These relationships are controlled primarily by the action of the cricothyroid muscle (CT) complex operating in opposition to the thyro-arytenoid (TA) group. In turn, this action is supplemented by certain intrinsic laryngeal muscles (especially the posticus) plus several of the extrinsic laryngeal muscles, and, of course, the subglottal pressure (Ps) provided by the respiratory system.

Relevance to the three prior principal theories of voice production

The data reviewed in this chapter—and the conclusions drawn can be employed to refine certain aspects of the Aerodynamic-Myoelastic Theory of phonation in humans. In general, however, it serves to strongly support this theory. Finally, it also can be employed as a starting point for study of other classes of VF activity.

References


62. Hollien, H., 1964. Stroboscopic laminography, sound motion picture, 16 mm, color and black and white, 10 minutes.