INVESTIGATION OF THE PHONATORY MECHANISM*

Thomas Baer

Abstract. A rational approach toward the development of improved techniques for the prevention, detection, diagnosis, and correction of vocal pathologies rests on an improved understanding of voice mechanisms. To achieve these goals, we need to better understand the dimensions of phonatory performance and their dependence both on the state of laryngeal structures and on patterns of control. Because of the inaccessible location of the larynx, few direct measurements of this performance are possible. Quantitative mathematical modeling is a useful vehicle for studying laryngeal vocal function. Continuation and extension of excised-larynx and animal studies can provide detailed data in support of the development and testing of these models. Human experiments, in vivo, aimed at factoring out the phonatory consequences of variations in individual laryngeal control parameters are suggested as a means of further extending such studies.

INTRODUCTION

A rational approach toward the development of improved techniques for the prevention, detection, diagnosis, and correction of vocal pathologies rests on an improved understanding of voice mechanisms. For prevention, we hope to understand the pattern of control, and its correlates in vibratory performance, whose breakdown leads to physiological failures in the laryngeal structures. Our research in detection and diagnosis is directed toward isolating non-invasive multidimensional measures capable of differentiating performance of larynges with different pathologies from the performance of normal larynges and from each other. In the area of correction, we hope to improve the conceptual framework for voice training and therapy, and improve the ability of surgeons to predict the phonatory consequences of alternative procedures. To achieve these goals, we need to better understand the dimensions of phonatory performance and their dependence both on the state of laryngeal structures and on patterns of control.

The process of phonation can be separated into three components: a phonatory system, its inputs, and its outputs. The system consists of two subsystems: one aerodynamic (the glottis), and the other mechanical (the

*A version of this paper was presented at the Conference on Assessment of Vocal Pathology, Bethesda, Md., April 1979. (Proceedings to be published in ASHA Reports.)

Acknowledgement. This work was supported by NINCDS Grant NS13870 and BRS Grant RRO5596.

vocal folds). Inputs to this system are muscular adjustments, transglottal pressure, and some other less significant variables. Outputs may be considered to be the pattern of mechanical vibrations in the vocal folds, or, more significantly for voice production, the pattern of airflow into the vocal tract. This latter output then serves as input to another system—the vocal tract—whose output is the radiated voice signal.

The myoelastic-aerodynamic theory of phonation (van den Berg, 1958) accounts grossly for the nature of phonation in terms of a passive interaction between the two phonatory subsystems when an appropriate combination of inputs is applied. The acoustic theory of speech (Fant, 1960) accounts for the effects of the vocal tract in transforming the glottal source signal to a radiated acoustic output signal. Although both of these theories have been well known for two decades or more, there are significant details that remain poorly understood. Thus, we have only limited ability to estimate the glottal volume velocity waveform by canceling the effects of the vocal tract from the speech output signal, and we have only limited ability to separate the influences of inputs to the phonatory system from the influences of the system itself on details of its output. Because of the inaccessible location of the larynx, few direct measurements of this output are possible.

Investigations into the mechanisms of phonation and its control have relied heavily on research with models. Much basic knowledge can be derived from experiments with excised larynges (e.g., van den Berg & Tan, 1959) and with live animal preparations, which serve as simplified models of their intact counterparts but which can be more carefully observed and more systematically controlled. Fabricated mechanical models have also been used to test hypotheses about the mechanism. For example, Smith (1962) experimented with a "membrane-cushion" model, which seems to incorporate some elements of the more recent "cover-body" theory of Hirano (1974, 1975, 1977). Mostly, however, mathematical descriptions and computer simulations have been used to formalize and refine knowledge about the mechanisms. Thus, the development of these models is both a goal and a tool of phonatory research.

The history of these modeling efforts parallels the improvement of our understanding of the system. As our understanding has become more complete, the models have become more complex. Building on the aerodynamic studies of van den Berg, Zantema, and Doornenbal (1957), Flanagan and Landgraf (1968) modeled the vocal folds as a simple mass-spring system performing horizontal movements with one degree of freedom. It soon became apparent that an additional degree of freedom was required to account for vertical phase differences. Ishizaka and Matsudaira (1972) corrected some errors in van den Berg's aerodynamic analysis, and showed that a two-mass model of the vocal folds could more realistically account for the conditions under which phonation could be initiated. Ishizaka and Flanagan (1972) simulated the two-mass model, extending the results of Ishizaka and Matsudaira, but were limited by this model's inability to account realistically for the closed period of the glottal cycle. Titze (1973, 1974) increased the number of masses to 16, in order to allow a distribution of vibrations along the anterior–posterior direction. This model also allowed for some vertical movements. Finally, Titze and Talkin (1979) have been investigating more sophisticated models that explicitly model the layered structure of the vocal folds (Hirano, 1974) and their behavior as a vibrator, and that incorporate tissue viscosity and bulk incompressibility.
Though it is understood that models must be complex to account realistically for the phonatory mechanism, there is also a danger inherent in the growth of complexity. As the number of degrees of freedom and the number of independent parameters multiply, the possibilities for accurately modeling the detailed mechanism improve, but so do the possibilities for producing apparently realistic behavior due to mechanisms that may not represent those of the real larynx. For our purposes, models must be mechanistically correct as well as descriptive of the output. It is therefore essential to determine as many of their parameters as possible and the constraints among them by direct measurement, and to evaluate the performance of these models in the greatest possible detail. Furthermore, we ought to be able to make directly testable predictions on the basis of our modeling efforts.

Further progress in understanding the detailed mechanism of phonation and in developing an accurate model of it thus depends on detailing the mechanical characteristics of vocal folds and determining their variation as functions of laryngeal control. It also depends on improved methods for measuring more detailed performance characteristics of real larynges, for comparing model performance to the performance of real larynges, and for generating testable predictions from modeling studies. Hirano has discussed, both at the Conference on Assessment of Vocal Pathology and in other publications (Hirano, 1975, 1977), measurements of mechanical properties of the vocal folds and some patterns of their variation with the contractions of individual muscles. Other papers at the conference will discuss techniques for obtaining detailed measurements, and Titze's paper will discuss methods for comparing the performance of models with these measurements on in vivo larynges. In the remainder of this paper, the continuation and extension of excised larynx and animal studies is urged because of their ability to produce detailed data for the direct testing of models. Then, some experiments in vivo, aimed at factoring out the phonatory consequences of variations in individual control parameters, are suggested as a means of further extending these studies.

I. EXPERIMENTS WITH EXCISED LARYNGES AND ANIMALS

It is well known that excised larynges, both canine and human, can simulate many of the vibratory characteristics of normal human larynges when they are attached to a pseudosubglottal system that supplies suitably conditioned airflow and when the positions of the laryngeal cartilages are suitably controlled, using strings to simulate the functions of muscles. As a simplified model of their intact counterparts, excised larynges offer several advantages. Because they are more accessible, they can supply observations and measurements that cannot be made in vivo. For example, both Matsushita (1969) and Baer (1975) have developed techniques for observing vibration patterns both from the normal supraglottal aspect and from the subglottal aspect. Baer also developed a technique for marking the vocal folds with small particles and tracking their frontal-plane movement trajectories throughout a glottal cycle using a microscope and stroboscopic illumination. Measurements could be made from both the supraglottal and subglottal aspects, and with the aid of qualitative observations, vocal fold shapes in the frontal plane throughout a cycle could be reconstructed from the measurements. With excised larynges, measurements of subglottal pressure and glottal airflow can be simplified. Furthermore, almost any technique for measuring characteris-
Figure 1. Schematic diagram of apparatus for measuring vibration patterns of excised larynges.
tics of phonatory vibrations can be used more effectively on an isolated larynx. Additional advantages are that the configuration of an excised larynx can be held constant or systematically varied, that its structures can be experimentally modified to determine the effects on vibration, and that they are accessible for measurement of mechanical properties in their configuration for voice production. The major limitations of the excised preparation—namely, that its death changes some of its mechanical properties, including its ability to tense the vocalis muscle—can be overcome by using live animal preparations and stimulating the muscles electrically. However, these advantages have not been fully exploited.

Baer's work with excised larynges was directed toward elucidating the phonatory mechanism in excised canine larynges. Although there is not space here to describe these experiments in detail, some of the most significant results are summarized below.

The experimental apparatus is shown schematically in Figure 1. A larynx was mounted on a pseudo-trachea, which made a right-angle turn just below the larynx, allowing a window to obtain a subglottal view. A stroboscope synchronized to subglottal pressure variations was mounted in front of the preparation. The phase at which the stroboscope was triggered could be adjusted to any point within the glottal cycle. Airflow was delivered at regulated flow rate or pressure, and both average pressure and average flow rate were measured. The subglottal system was intended to simulate the acoustic properties of the real subglottal tract. The apparatus was mounted on the top of a rotary indexing table, whose tabletop could be rotated, so that observations could be made through the microscope at any angle. The tabletop could also be translated along its two horizontal axes. A measurement system was devised by which the locations of any points observed through the microscope could be determined in three dimensions.

With respect to gross aspects of the performance of excised larynges, observations already made by others were replicated. In addition, it was observed that, for a given laryngeal configuration, phonation could be maintained at values of subglottal pressure below those required for initiating phonation. As the tissues desiccated, the separation between conditions for onset and conditions for maintenance increased. Thus, mobility of the surface tissues appeared to be important for initiating phonatory vibration. Perhaps this observation has some implications for the assessment of pathologies.

Figure 2 shows data from a run in which the frontal-plane trajectories of three particles were measured at eighth-cycle increments while the larynx sustained steady-state vibration. One particle was on the lateral superior surface of the vocal folds, a second was near the medial superior surface of the folds, and a third was on the lower (subglottal) surface. These trajectories are typical. They were roughly elliptical, in the clockwise direction (for the coordinate system shown). The minor axis of the ellipses decreased as average distance from the midline increased. Subglottal particles moved primarily in a horizontal direction, while supraglottal particles well off the midline moved primarily in a vertical direction. Trajectories of particles near the midline often exhibited complex perturbations near the superior-medial parts of their trajectories. Trajectories of the two upper
Figure 2. Frontal-plane trajectories of three particles during a single glottal cycle. Measurements were made at eighth cycle increments, numbered 0 through 7. The inset to the right of the trajectories contains notes about the measurements, including the angle, \( \theta \), of the tabletop for which each measurement was made. The schematic sketch at the top of the inset indicates the particle locations with respect to the margin of the vocal fold.
particles crossed, so that the particles were nearly vertically aligned during one measurement and horizontally aligned during another. Thus, the vibrations were complex. Some aspects of the trajectories and of vibrations in general were consistent with the notion of a displacement wave, progressing up the medial surface at a velocity of about 1m/sec, and then progressing laterally on the superior surface at .3-.5m/sec. The supraglottal wave was easily observed, as with normal human larynges, and its velocity was measured directly. Glottal closure also exhibited wavelike properties. Tissues at the lower edge of closure were peeled apart, while tissues above the point of closure were still coming together. The depth of closure was often almost negligible immediately before the glottis opened. The middle particle in Figure 2 appeared to be on the superior part of the vocal folds for part of the cycle, and was below the point of closure for part of the closed phase. Thus, it is evident that the vibrations are complex and cannot be well modeled, in detail, as simple translations of a small number of lumped-parameter masses.

Although some aspects of the vibration patterns seemed best describable by surface waves along the cover of the vocal folds, vibrations of the edge also appeared to be describable as string vibrations (that is, whole-body translation and torsional flexure). There may have been components of both types of vibrations. This interpretation is interesting, because interactions between the two types of vibration as a function of variations in control parameters may help to explain fine control over voice quality variations.

Detailed shapes of the vocal folds during the eight phase increments in Figure 2 were estimated and are shown in Figure 3. A two-mass model approximation could be superimposed on these shapes if vertical movements of the masses were allowed. Given this approximation, the aerodynamic theory of Ishizaka and Matsudaira (1972) was capable of reconciling average subglottal pressure with average flow rate. It was also shown, as expected, that the aerodynamic model provided for the efficient transfer of energy from the aerodynamic system to the mechanical system (Stevens, 1977), given the nature of vertical phase differences. The mechanical parts of the two-mass model did not well account for these data, however. Thus, to the extent it could be tested, the aerodynamic aspect of the two-mass model seemed accurate, but the mechanical part of the model seemed inadequate.

A change in particle trajectories was observed as the tissues desiccated and vibrations eventually ceased. These and other measurements suggested that particle trajectories could be considered as oscillations around an unstable equilibrium position. This result implies that small-signal modeling techniques, such as those of Ishizaka and Matsudaira (1972), which account for voice onset by finding unstable solutions to linear equations, are justified.

Excised larynges were able to produce nearly normal vibrations even when the vocalis muscle on one or both sides was completely removed. However, these preparations did not seem capable of falsetto vibrations. Wave motions with velocity similar to that of the normal case were still seen to propagate upward on the medial wall. Particle trajectories were somewhat similar to the normal case, although they differed in some details. These observations should be especially useful for testing models that account for the layered structure of the vocal folds.
Figure 3. Sketches of vocal-fold shape during a vibratory cycle. These shapes were estimated on the basis of the data shown in Figure 2, which is superimposed in each panel. Bilaterally-symmetric shapes are shown for display purposes, although measurements were actually made on only one side. The corner in the upper right of each panel indicates 1 mm scales. Individual shapes at eighth cycle increments are shown at the lower part of the figure. The top panel shows all of them superimposed.
The experiments described above illustrate the potential value of developing a model specifically for excised larynges, as a step in developing a model for the in vivo case. An advantage to modeling the excised preparation explicitly is not only its versatility, as illustrated by the experiments with excised vocalis muscles, but also the fact that measurements of mechanical properties can be made on the same preparation on which the vibration patterns are measured.

Optical techniques for measuring frontal plane vibration patterns, such as those used by Baer, are limited because they are time consuming and because only vibrations of the vocal fold surfaces can be measured. Radiographic techniques may provide a solution to the problem of measuring vocal fold shapes throughout a cycle. There have been some radiographic studies of vocal fold vibrations in vivo. Sovak, Courtois, Haas, and Smith (1971) described a high-speed radiographic study capable of resolving the details of a glottal cycle. Hollien, Coleman, and Moore (1968) developed the technique of stroboscopic laminography, in which an x-ray source is pulsed stroboscopically during a laminographic procedure. For steady phonation, images of a frontal section could thus be obtained at successive phases within a cycle. The usefulness of these studies was limited by the poor quality of the images obtained. Furthermore, they may be no longer practical, in view of modern concerns about radiographic dosage, especially to the thyroid gland. However, such techniques could be applied safely and more effectively to the study of excised or animal larynges. A promising improvement on these techniques was recently described by Saito (1977) and Saito, Fukuda, Ono, and Isogai (1978). Small lead pellets were affixed to the vocal fold surfaces and also implanted within the vocal folds, so that both internal and external vibrations could be monitored. Stroboscopic radiography, synchronized to the voice, was then used to track the movements of these particles throughout cycles of vibration. Such measurements might be made even more effectively with a computer-controlled x-ray microbeam system (Fujimura, Kiritani, & Ishida, 1973; Kiritani, 1977), if its detector output were stroboscopically sampled or its source stroboscopically pulsed, because of the improved spatial resolution of this device. Conceivably, radiopaque medium could be introduced through the circulatory system, as a further improvement of this technique.

II. MEASUREMENTS IN VIVO: RESPONSES TO INDIVIDUAL CONTROL VARIABLES

There are many parameters controlling phonation in the normal human larynx. Control is exerted most directly through the effects of the intrinsic muscles on laryngeal configuration and through transglottal pressure. Forces exerted by the extrinsic laryngeal muscles and other extrinsic structures also have an effect. Acoustic load can modify the patterns of airflow through the glottis and probably the mechanical vibrations as well. There are probably other effects, such as control of vascular and mucous supply, which are less well understood. During voluntary control of phonation, variations in several of these parameters are intercorrelated (see, for example, Atkinson, 1978). Although such variables as the levels of electromyographic activity in individual muscles and subglottal pressure can be correlated with corresponding changes in fundamental frequency or other aspects of phonatory performance, correlation does not guarantee causality, because of the intercorrelations among control variables. Therefore, it has been difficult to isolate
the detailed phonatory response to any one of them. Nevertheless, these
detailed effects must be known in order to determine the relevance of data
from excised larynx and animal experiments, to adequately test detailed
phonatory models, and, in general, to fully understand phonatory function.

One method for isolating the effects of a given parameter is to
externally apply involuntary perturbations and observe the phonatory response
while other parameters remain constant. This technique has been most success­
fully used for examining the effects of changes in subglottal pressure on
fundamental frequency. Several experiments have been reported in which
subglottal pressure is increased by a sudden push on the chest or abdomen of a
phonating subject, and both subglottal pressure and fundamental frequency are
monitored during an interval for which no muscular response is assumed to
occur (for example, van den Berg, 1957; Isshiki, 1959; Ladefoged, 1963; Öhman
& Lindqvist, 1966; Fromkin & Ohala, 1968). This experiment was recently
replicated by Bauer (1979), who also monitored the electromyographic activity
of laryngeal muscles to ensure the absence of a response. Transglottal
pressure can also be varied supraglottally, through modulation of intraoral
pressure (Lieberman, Knudson, & Mead, 1969; Hixon, Klatt, & Mead; 1971;
Rothenberg & Mahshie, 1977). When pressure modulations are oscillatory, at
frequencies of about 5-10Hz, continuous muscular compensation does not seem to
occur, although EMG evidence to support this claim has not been published.

Although results of these induced-pressure-change experiments differ in
some details, their consensus indicates that fundamental frequency varies with
transglottal pressure at rates of about 3-5Hz cm H2O within the speech range,
with higher rates at higher fundamental frequencies or in falsetto register.
These results, as well as correlation between fundamental frequency and
subglottal pressure during voluntary control (Atkinson, 1978), suggest that
the phonatory response to pressure change is fast, perhaps within the interval
of one or two glottal periods.

The effects of involuntary perturbations in acoustic load on fundamental
frequency have also been investigated through systematic variation in the
length of a tube that artificially extends the vocal tract (Ishizaka,
fundamental frequency of as much as 20Hz were obtained by varying the length
of the tube. However, it was not determined in these experiments whether
there was any compensatory laryngeal response. It is easily shown that such
artificially increased acoustic loads can have an effect on phonation. If one
phonates an ascending scale into an artificially extended vocal tract (such as
a mailing tube), the voice will typically break or switch to falsetto when the
fundamental frequency nears the first resonance frequency of the tract. A
lower order manifestation of this phenomenon might account for the intrinsic
pitch of vowels (Peterson & Barney, 1952). In any case, such experiments
could be repeated more carefully to further constrain the performance of
phonatory models.

The logical counterpart to these studies for quantifying the effects of
individual muscles on phonatory performance would probably require electrical
stimulation of the muscles. There are no accounts of any such studies on
normal human subjects, and it is unclear whether stimulation experiments are
possible in practice. However, an alternative method, which isolates the
effects of single-motor-unit contractions, has recently been used by Baer (1978) for investigating the effects of individual muscles on fundamental frequency. Rather than analyzing gross aspects of fundamental frequency control, this method relates very small changes in fundamental frequency (namely, pitch perturbations) to very small changes in muscle tension, which can be related to single-motor-unit activity. Statistical independence between motor-unit inputs can then be exploited to uncorrelate the muscles, and examine their individual causal effects on fundamental frequency.

This method extends the use of an averaging technique that was first developed for studying properties of single motor units in skeletal muscles (Milner-Brown, Stein, & Yemm, 1973). Single-motor-unit action potentials (see Harris, 1981) must be identified in an electromyographic recording while the muscle sustains a contraction. A simplified muscle model, which is approximately valid at low to moderate levels of contraction, is assumed. This model is shown in Figure 4. Its inputs are the action potential trains from individual motoneurons. Each of these can be considered a random point process, and they are statistically independent across units. Each motor-unit action potential triggers a mechanical twitch—a positive pulse of tension whose detailed characteristics vary across motor units. At least some of these units fire at low enough rates so that adjacent twitches do not overlap. The output tension of the whole muscle is the summation of its constituent motor unit outputs. Although many of the motor unit outputs are trains of pulses, they sum to an approximately constant, though noisy, value because they are statistically independent. The relative amplitude of this noise depends on the number of motor units and their firing rates.

Given the model in Figure 4, the contribution of a single motor unit to the output tension (its contraction properties) can be estimated if its input action potentials can be identified and if these inputs are isolated by intervals great enough to ensure against overlap of adjacent contractions. Samples of the output tension waveform following the inputs are aligned and averaged. The output of the isolated motor units is always the same within these intervals, while the outputs of all other motor units are random and thus average to a constant value.

To apply this technique to investigation of fundamental frequency control, we note that motor-unit firings are statistically independent across muscles as well as within a muscle. We then hypothesize that muscle-tension variability contributes to the fundamental frequency perturbations that can be measured when a normal phonating subject attempts to sustain a steady tone. The resulting model for pitch perturbations is then indicated in Figure 5. Laryngeal muscles produce roughly constant output tensions that are noisy because of single-unit effects. The noise components across muscles are uncorrelated. The complex effect of muscle forces on the vocal folds, which we have lumped under the term "vocal fold tension," is also roughly constant, but noisy. Output fundamental frequency then depends on this tension and other independent inputs such as subglottal pressure and, perhaps, mucosity and other random effects. All the detailed inputs to this model are thus statistically independent. According to the model, then, fundamental frequency as a function of time can be treated as an output and be averaged just as muscle tension in earlier studies to estimate the effects of single-motor-unit contractions in that muscle. The effects of other muscles and other inputs average to a constant value.
Figure 4. Simplified model of a muscle during a sustained contraction.
MODEL FOR PITCH PERTURBATIONS

Figure 5. Model for pitch perturbations during production of a steady tone.
To obtain data for such a study, a subject is asked to sustain a steady tone for several breaths. Electromyographic (EMG) activity, obtained through hooked wire electrodes from a laryngeal muscle under study, and the voice signal obtained through a standard microphone are recorded and input to a digital computer. After instantaneous fundamental frequency as a function of time is derived, this waveform is offset by approximately its average value and amplified to exaggerate the perturbations. Isolated single-motor-unit firings are identified in the EMG waveform. Then, samples of the EMG waveform and the F_0 perturbation waveform are aligned around the single firings and averaged. The sample window extends from 100ms before to 300ms after these firings.

Figure 6 shows a 1.5s sample of data when the muscle under study was the cricothyroid, whose function as a vocal-fold tenser and hence as a pitch raiser is well known. Fundamental frequency was about 100Hz, which is in the lower part of the subject's range, in order to keep the number of recruited units and their firing rates low. As this figure shows, fundamental frequency was estimated to 1 Hz resolution. Although cycle-to-cycle variations rarely exceed 1Hz, perturbations over larger time intervals were about 4Hz wide. Two firings have been isolated in this record, and the corresponding sample intervals are indicated by horizontal lines.

Figure 7 shows the results of the averaging calculation for this experiment after 19 suitable firings were identified. The upper panel shows the averaged EMG signal, which exhibits a pulse only at the lineup point, as expected. The lower panel shows the average F_0 perturbation. This signal is approximately at baseline both to the left of the lineup point and to the far right of the window. However, there is a positive pulse beginning immediately after the lineup point. This pulse reaches its peak amplitude of 1Hz at a latency of about 70-80ms. The pulse appears to indicate that the single-motor-unit contraction caused, on the average, a 1Hz increase in fundamental frequency.

A similar calculation was performed for one of the strap muscles, an extrinsic laryngeal muscle whose possible function in lowering F_0 has been a source of some controversy. When fundamental frequency was in the middle of the subject's range, no systematic effect was found. Results when the fundamental frequency was low are shown in Figure 8. Although these data are somewhat noisier than those in Figure 7, they appear to exhibit a negative pulse in the interval immediately after the lineup point. Thus, the strap muscle is shown to have a causal effect in lowering fundamental frequency from an already low level.

The confirmation of a muscular contribution to F_0 perturbations is itself interesting, since perturbations have been used as an indicator of vocal pathology. These results show that care must be taken when interpreting patterns of perturbation. More relevant to this discussion, however, is the fact that we can show the response to a short duration pulse of tension in a single muscle, and that these data can thus be used to constrain the performance of laryngeal models. It was noted that the average pitch perturbation for the cricothyroid muscle begins immediately after the lineup point. This shows that the phonatory response must begin within one glottal cycle. The latency of the peak of the response, 70-80ms, includes contribu-
Figure 6. Short segment of data during production of a steady tone at about 100 Hz. Top: voice waveform; Middle: EMG activity of the cricothyroid muscle; Bottom: "instantaneous fundamental frequency" extracted from the voice waveform. Two sets of horizontal lines indicate intervals from 100 ms before to 300 ms after single-motor-unit firings in the cricothyroid muscle.
Figure 7. Ensemble-average waveforms of EMG activity from the cricothyroid muscle and corresponding instantaneous fundamental frequency. All waveforms have been aligned at the time of a single-motor-unit firing for purposes of averaging.
Figure 8. Ensemble-average waveforms of EMG activity from an unspecified strap muscle and corresponding instantaneous fundamental frequency. All waveforms have been aligned at the time of a single-motor-unit firing for purposes of averaging.
tions due to muscle contraction time, mechanical response latency in the larynx, and latency of phonatory response. Since both the latency and the amplitude of the mechanical motor-unit contractions can be estimated in animal experiments, these data might be further applied to the detailed testing of models of laryngeal performance, especially in comparison with data reported by Hirano (1975) relating changes in shape and mechanical properties of vocal folds to stimulation of various muscles. These data might also shed some further light on the pattern of motor control. For example, the relatively large amplitude of the $F_0$ perturbation pulse in Figure 7 relative to the overall perturbation in Figure 6 suggests that very few motor units were firing at rates low enough to show the effects of individual twitches. However, it is unclear how many other units may have been in tetanus. Perhaps the greatest value of the single-unit technique will be in elucidating the phonatory function of muscles such as the vocalis, whose gross patterns of activity are so intercorrelated with those of other muscles during ongoing regulation of phonation that their detailed effects have remained obscure.

In considering the function of individual control parameters in this section, we have only discussed measurements of their effects on fundamental frequency. The reason for this is that, with few exceptions, these are the only measurements that have been made. Fundamental frequency by itself, however, is evidently not a very complete descriptor of phonatory activity. As fundamental frequency is varied, attributes of the vocal source waveform that contribute to intensity and voice quality also vary. It is important to determine how these parameters covary when changes are produced by different control mechanisms, and, for purposes of assessing vocal pathology, how these relationships change in different pathological states.

Techniques to be discussed in today's session can be used to measure some of these different parameters of phonatory performance, such as amplitude of the glottal pulse and open quotient. When these parameters are measured cycle-to-cycle, the same techniques described in the section for studying fundamental frequency control can be utilized to assess the effects of different control parameters. These data, together with such anatomical and physical studies as those reported by Hirano (1975), are needed to improve our understanding of the phonatory mechanism and constrain the performance of mechanistic models. Thus, these studies should be pursued. Furthermore, if it were possible, it would be even more useful to study not only changes in vibratory performance characteristics as a function of these control parameters, but also intermediate variables such as the positions of the laryngeal structures and their mechanical properties. However, these experiments must await the development of techniques for measuring these parameters.

Finally, further insights are needed into the detailed conditions necessary for initiating and sustaining phonation, as well as for regulating ongoing phonation. An example of how such studies might be performed in vivo is by using involuntary perturbations of subglottal pressure. For example, a subject might be asked to assume a configuration appropriate for voicing but to maintain subglottal pressure at a level below the threshold for voice onset. Transglottal pressure might then be suddenly increased, say using a chest push procedure, to a level for which phonatory vibrations are initiated, while laryngeal configuration remains constant. Conditions for voice onset could then be determined, in terms of the level of subglottal pressure, as a
function of variations in the configuration. With negative transglottal pressure perturbations, conditions for voice offset could also be studied.

REFERENCES


Hirano, M. Morphological structure of the vocal cord as a vibrator and its variations. Folia Phoniatrica, 1974, 26, 89-94.


Ishizaka, K., & Matsudaira, M. Fluid mechanical considerations of vocal cord vibration. SCRL Monograph (Speech Communication Research Laboratory, Santa Barbara), 1972, No. 8.


Isshiki, N. Regulatory mechanism of the pitch and volume of voice. Oto-Rhino-Laryngology (Kyoto), 1959, 52, 1065-1094.

Kiritani, S. Articulatory studies by the x-ray microbeam system. In


van den Berg, Jw. Sub-glottal pressure and vibrations of the vocal folds. Folia Phoniatrica, 1957, 9, 65-71.

van den Berg, Jw. Myoelastic-aerodynamic theory of voice production. Journal of Speech and Hearing Research, 1958, 1, 227-244.

van den Berg, Jw., & Tan, T. S. Results of experiments with human larynges. Practica Oto-Rhino-Laryngology, 1959, 21, 425-450.