

Reflex activation of laryngeal muscles by sudden induced subglottal pressure changes

Thomas Baer

Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06510
(Received 7 August 1978; revised 25 January 1979)

In measuring the effect of subglottal pressure changes on fundamental frequency (F_0) of phonation, the effects of changing laryngeal muscle activity must be eliminated. Several investigators have used a strategy in which pulsatile increases of subglottal pressure are induced by pushing on the chest or abdomen of a phonating subject. Fundamental frequency is then correlated with subglottal pressure changes during an interval before laryngeal response is assumed to occur. The present study was undertaken to repeat such an experiment while carefully monitoring electromyographic (EMG) activity of some laryngeal muscles, to discover empirically the latency of the laryngeal response. The results showed a consistent response to each push, with a latency of about 30 ms. Despite this response, analyses of fundamental frequency versus subglottal pressure changes during the interval of constant EMG activity were in general agreement with previously published values. With respect to the nature of the electromyographic response itself, its timing was found to be within the range of latencies appropriate for peripheral feedback, and was also similar to that for an acoustically — or tactually — elicited startle reflex.

PACS numbers: 43.70.Bk

INTRODUCTION

The effect of subglottal pressure changes on fundamental frequency (F_0) of phonation has been a subject of theoretical and research interest for at least the past twenty years. However, measurement of this effect is difficult because, during normal phonation, changes in subglottal pressure and in the activity of laryngeal muscles are usually correlated (for example, Atkinson, 1978). To measure the effects of subglottal (or transglottal) pressure changes on fundamental frequency, the effects of changing laryngeal muscle activity must be eliminated. A commonly adopted strategy intended to eliminate laryngeal muscle effects on fundamental frequency is to induce changes in subglottal pressure experimentally, and to measure fundamental frequency before laryngeal reaction is presumed to occur. Several studies have been reported in which pulsatile increases of subglottal pressure are produced by pushing suddenly at random intervals on the chest or abdomen of a phonating subject (for example, van den Berg, 1957; Isshiki, 1959; Ladefoged, 1963; Ohman and Lindqvist, 1966; Fromkin and Ohala, 1968). Van de Berg (1957), for instance, argued that "no known human reflexes can occur faster than 100 ms," so that if fundamental frequency is correlated with subglottal pressure during this interval, laryngeal configuration should be considered constant.

Results from these experiments and others with a somewhat different paradigm (Lieberman, Knudsen, and Mead, 1969; Hixon, Klatt, and Mead, 1971) have shown general agreement. The sensitivity of fundamental frequency to subglottal pressure is usually found to be in the range of 3–7 Hz/cm H_2O , with the higher values occurring at the higher fundamental frequencies or in falsetto.

Constancy of laryngeal configuration can only be ensured if electromyographic (EMG) signals from the laryngeal muscles are recorded simultaneously. Although some of the cited investigators apparently moni-

tored EMG signals and eliminated from processing those tokens for which the EMG showed sizable variations, these results were not reported in detail. Furthermore, the experiments reported depend on the unproven assumption that reflex latencies must exceed 100 ms. The assumption is certainly true for voluntary reaction times. For example, Draper, Ladefoged, and Whitteridge (1960) report a minimum reaction time of 140 ms in a respiratory control task, and Netsell and Daniels (1974) report a similar latency for EMG signals associated with voluntary lip movements. Voluntary adjustments for initiating phonation also appear to be limited by a similar latency (Izdebski and Shipp, 1976), and indeed, voluntary reaction times for F_0 change have recently been measured (Miller, Hanson, and Shulak, 1978) and found to be similar. However, a peripheral reflex might well be much faster. The mechanical eyeblink response to an acoustic startle stimulus begins at about 40 ms (Landis and Hunt, 1939). Reflex activation of respiratory muscles with EMG latencies of 33–80 ms to sudden changes in pressure has been reported by Sears and Newsom Davis (1968). The laryngeal protective-closure reflex is similarly biologically basic, and the laryngeal adductor muscles are among the fastest in the body, with mechanical response times as short as 15–20 ms (Sawashima, 1974; Atkinson, 1978).

Because, as the above discussion indicates, reflex adjustments could have interfered with previous measurements, it seemed useful to repeat a chest-pushing experiment while carefully monitoring EMG signals from laryngeal muscles, as well as the subglottal pressure and voice waveforms. A series of experiments with one subject was performed. The results showed a rapid and consistent EMG response to each chest push. Although this response could have affected results of previous investigations, analyses of fundamental frequency versus subglottal pressure during the period before this response could occur were in general agreement with published values. The nature of the EMG response itself is of interest, and is considered further.

I. METHODS

In the initial experiment, the subject sat upright in a dental chair and produced steady phonation at, successively, three different fundamental frequencies (nominally 94, 110, and 220 Hz) in chest voice and one (240 Hz) in falsetto while the experimenter pushed sharply and at random intervals on his chest. Some double pushes were used, but these were eliminated during later stages of processing. Pressure was measured through a catheter inserted into the lower subglottal space of the larynx, while the voice signal was recorded through a standard microphone. The EMG signals were recorded with hooked-wire electrodes. Suitable signals were obtained from two laryngeal adductor muscles, the vocalis (VOC) and the interarytenoid (INT). Unfortunately, signals from the cricothyroid muscle are not available. Some variations of this procedure were introduced in two later runs, as discussed below.

The data were processed using the facilities at Haskins Laboratories. Measurements of subglottal pressure and rectified electromyographic activity were obtained by integrating the signals in contiguous, non-overlapping 5-ms intervals. The voice waveform synchronous with the pressure and EMG signals was sampled and digitized at a 10-kHz rate. From this waveform, an estimate of the fundamental frequency over each 5-ms interval was obtained by an autocorrelation method (Lukatella, 1973). Measures of the amplitude of the audio waveform were also made over 5-ms intervals. Thus, the sampling rate for all these signals was 200/s, and no further smoothing was applied.

Ensemble averages and standard deviations were calculated for each of the four F_0 conditions. For each condition, the data were carefully aligned for averaging at the onset of the rise in subglottal pressure. Since the lineup point was the beginning of the push, as measured from the onset of pressure rise, the ensemble averages reflect average responses to a push.

II. RESULTS

Some results from the first experiment are illustrated in Fig. 1. This figure shows results from the lowest- F_0 condition (94 Hz). Each column contains waveforms for one variable. The lineup point is shown by the vertical line through each waveform. The top row shows, for each column, the average waveform based on 18 tokens. The four rows underneath contain waveforms from the first four of these 18 tokens. The columns contain, from left to right, waveforms for subglottal pressure, acoustic variables (voice amplitude envelope and fundamental frequency), and electromyographic signals from the vocalis and interarytenoid muscles.

As this figure shows, the subglottal pressure builds up rapidly at the onset of a push, and then falls more slowly. As could be expected, the amplitude envelope reflects, at least grossly, the variations in subglottal pressure, as do variations in fundamental frequency. The EMG signals for individual tokens are noisy, but

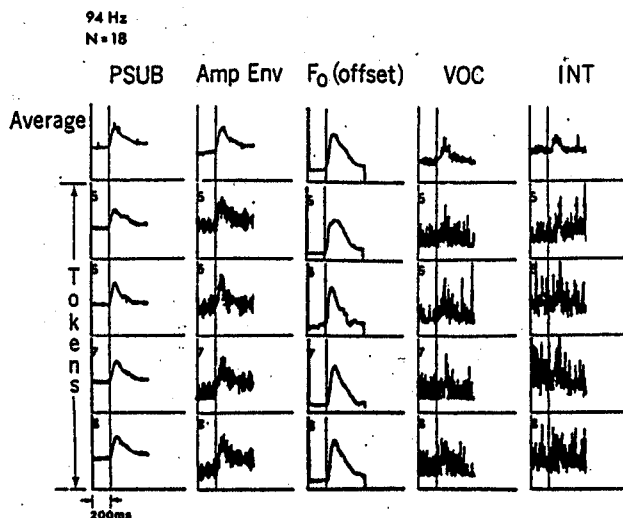


FIG. 1. Results from the lowest-fundamental-frequency condition (see text). Ordinate scale values, not shown here, are the same as in the next figure.

clearly contain peaks of activity in an interval after the lineup point. These peaks appear consistently. Although it would be difficult to measure the timing of the peaks from individual pushes, their timing appears grossly constant, so that averaging seems appropriate. Timing of the average waveform is more easily measured than that for individual tokens. This timing will be discussed below.

Figure 2 shows the average waveforms from each of the four conditions in vertical alignment to clearly display the timing relationships. In all four conditions, subglottal pressure, amplitude envelope, and fundamental frequency begin to rise immediately at the lineup point. In all three chest voice conditions, vocalis

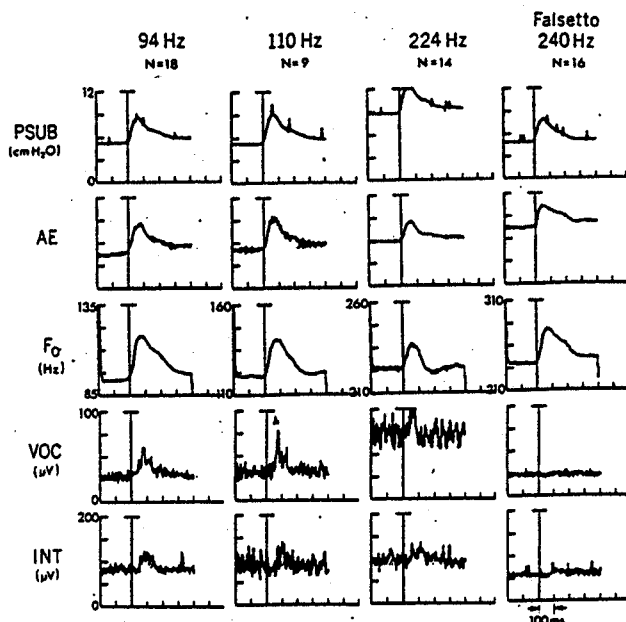


FIG. 2. Average waveforms from all four conditions. Note that the scale factor for the F_0 plot in the falsetto condition is different from that for the other three conditions.

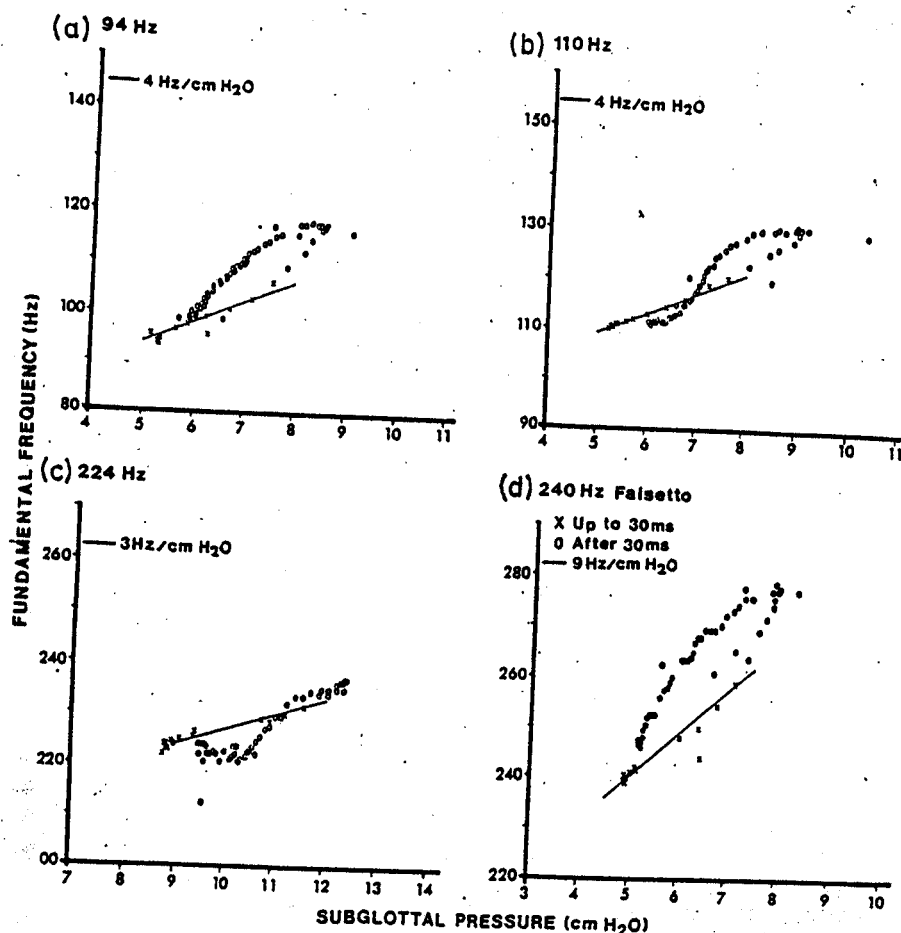


FIG. 3. Scatter plots of fundamental frequency versus subglottal pressure for the conditions: (a) 94 Hz, (b) 110 Hz, (c) 224 Hz, and (d) 240 Hz, (falsetto).

and interarytenoid activity rise sharply at a latency of about 30–40 ms after the lineup point. In the lowest- F_0 condition, there is also a gradual rise of activity immediately after the lineup point. This apparent activity may reflect a "microphonic effect."¹ For the falsetto condition, the activity of the laryngeal adductors was relatively low, and there appeared to be little or no response to the chest push.

Figure 3 shows scatter plots of the average fundamental frequency versus the average subglottal pressure waveforms for the four conditions shown in Fig. 2. Points corresponding to times up to 30 ms after the lineup point are plotted with different symbols from those which occur later. Points in the interval up to 30 ms are treated differently because they are uncontaminated by muscle responses. Additionally, their trajectories are reasonably well fit by straight lines. For the three chest-voice conditions [Figs. 3(a), (b), and (c)], the slopes of these lines are 4, 4, and 3 Hz/cm H_2O , respectively. For the falsetto condition, shown in Fig. 3(d), the slope was greater (9 Hz/cm H_2O). This aspect of the results agrees qualitatively with the consensus of earlier studies. Thus, although the results confirm the initial hypothesis, that muscle activity could have interfered with previous investigations, the actual results of those investigations are substantiated.

Beyond 30 ms, the trajectories depart from linearity. Perhaps the onset of muscular activity accounts for

some of this apparent nonlinearity. However, this departure from linearity cannot be fully explained on the basis of EMG activity whose latency is 30 ms. At least 15 additional ms would be required to account for the mechanical response time of the muscles. Therefore, nonlinearity seems to be inherent in the relationship between fundamental frequency and subglottal pressure.

Response latencies of 30–40 ms are too short to be centrally mediated, and therefore must be due to peripheral reflexes. However, before it can be concluded that the observed EMG responses were actually due to a peripheral reflex, it must be established that the subject had no cue for the impending chest push prior to the onset of the pressure rise. In the foregoing results, all latencies were referred to the beginning of the pressure rise. The possibility was considered that the beginning of the push may have significantly preceded the pressure rise, and therefore tactile cues may have been available to the subject.

To establish this latency, another smaller experiment was performed. The chest-pushing procedure was repeated, and a contact switch was fitted on the subject's chest. The experimenter pushed directly on the switch. A signal that suddenly increased with switch closure was recorded in parallel with the voice waveform and with the EMG activity from the vocalis muscle. Subglottal pressure was not recorded, since the foregoing results show that increase of subglottal

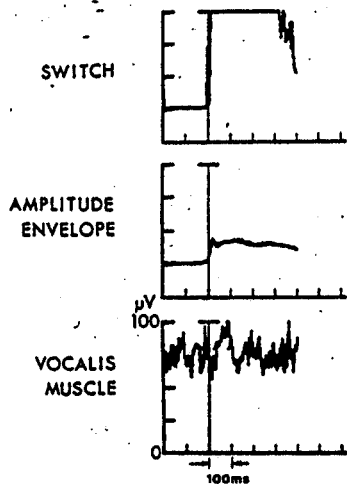


FIG. 4. Results from experiment with contact switch. The upper panel shows the average switch contact signal. These average signals are based on 23 pushes.

pressure is immediately reflected in the voice amplitude envelope.

Results from this experiment are shown in Fig. 4. The upper plot shows the switch contact signal. Results from 23 pushes have been averaged, and all signals have been aligned with respect to switch closure. The middle plot shows the average voice amplitude envelope. This signal begins to increase exactly at the lineup point. Hence, the delay between the onset of the physical push and the beginning of the pressure rise must be within the 5-ms sampling interval of the processing system. This figure shows that the response latency may be measured from either the onset of pressure rise or from the tactile stimulus of the push. The bottom plot in Fig. 4 shows the average simultaneous activity of the vocalis muscle. Although the electrodes inserted into this muscle produced extremely noisy waveforms, it can be seen that there is a peak in EMG activity beginning at a latency of about 30 ms from the lineup point. The possibility that other cues, such as visual or auditory correlates of the experimenter's preparation to push, contributed to the subject's time to respond is unlikely (though, unfortunately, there were no controls for this).

Having established that there is a rapid and consistent EMG response of the laryngeal adductors to chest pushes, we must now ask what is the nature or function of this response. Two alternative hypotheses are suggested: (1) that the response represents the correction of a closed-loop control system to a perturbing signal (e.g., the increase of pressure, which might tend to abduct the folds), or (2) that the response represents the laryngeal protective-closure component of the startle reflex. We have not collected any data that can distinguish between these two possibilities. However, we performed a simple experiment to show that the second alternative is consistent with the data at hand.

The subject again produced steady sustained phonation. Startle responses were elicited both by clapping

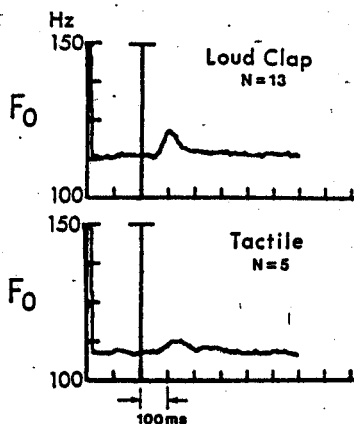


FIG. 5. Average perturbations in fundamental frequency caused by startle reflexes elicited by auditory and tactile stimuli.

loudly near the subject's ear and by the use of a sudden, painful tactile stimulus on the forearm. Fundamental frequency was tracked with the aid of an electrical glottagraph. Average F_0 curves based on 13 claps and on five tactile stimuli are shown in Fig. 5. The vertical line marks the onset of the startle-eliciting stimulus. It can be seen that there is a perturbation of fundamental frequency associated with the startle response. The latency of onset of this perturbation is about 50 ms. This combined latency could be accounted for by a 30–35 ms EMG latency and a 15–20 ms mechanical response latency for the muscle.

III. DISCUSSION

The existence of a rapid and consistent EMG response to chest pushes appears to contradict the assumption by past investigators that laryngeal configuration can be considered constant for at least 100 ms after the push. Nevertheless, the relationship between fundamental frequency and subglottal pressure derived by correlating those variables within a 30-ms interval after the push agrees with the results of past investigations.

The trajectories in Fig. 3 exhibit hysteresis in addition to the nonlinearity already noted. One component of the hysteresis is undoubtedly due to muscular forces, both from the adductor muscles whose EMG activity was recorded and from other muscles. The complex nature of some of these trajectories, in which the descending part crosses the ascending part, suggests a pattern of control in which an early F_0 -raising response by one set of muscles is corrected for by an F_0 -lowering response from another set, which overshoots. However, EMG data must be obtained from a larger set of muscles before any such hypothesis can be tested. In addition, there may also be an inherent component for both the nonlinearity and the hysteresis, as suggested by the results for the falsetto condition, where no muscle response is apparent, and from the results of experiments with excised larynxes (van den Berg and Tan, 1959; Baer, 1975).

From a consideration of the latency alone, there are

two plausible explanations for the EMG response. The results of Sears and Newsom Davis (1968) provide an example of another respiratory muscle with a response at comparable latency to pressure changes introduced at the mouth. These responses were interpreted in terms of peripheral feedback control. However, we have shown that the latency is also consistent with that for a startle response. Measurements from a larger set of muscles, and probably a more elaborate experimental paradigm, are necessary to elucidate the nature of the observed responses. Continuation of these studies is worthwhile for what it may reveal about reflex mechanisms available for the ongoing control of laryngeal function during speech.

ACKNOWLEDGMENT

The author is grateful to Dr. Hajime Hirose and Dr. Seiji Niimi, who performed all the physiological procedures reported here. This research was supported by NIH grants DE01774 and NS13870.

The "microphonic effect" is sometimes observed in vocalis-muscle insertions when the electrode tips are close to the mucosal surface. The amplitudes of vocal-fold vibration are large at low phonatory frequencies, and the resulting movements probably modulate the interelectrode distance, producing a modulation of the recorded signal at the phonatory frequency. With increases of subglottal pressure, the amplitudes of vibration, and hence the microphonic effect, increase. The microphonic signal can sometimes be eliminated by pulling back the electrodes, presumably so that the amplitudes of mechanical vibration are reduced.

- Atkinson, J. E. (1978). "Correlation analysis of the physiological features controlling fundamental voice frequency," *J. Acoust. Soc. Am.* 63, 211-222.
- Baer, T. (1975). "Investigation of phonation using excised larynxes," Ph.D. thesis, Massachusetts Institute of Technology (unpublished).

- Draper, M. H., Ladefoged, P., and Whitteridge, D. (1960). "Expiratory pressures and airflow during speech," *Br. Med. J.* 1, 1837-1843.
- Fromkin, V., and Ohala, J. (1968). "Laryngeal control and a model of speech production," *Working Papers in Phonetics (UCLA)* 10, 98-110.
- Hixon, T. J., Klatt, D. H., and Mead, J. (1971). "Influence of forced transglottal pressure on fundamental frequency," *J. Acoust. Soc. Am.* 49, 105(A).
- Isshiki, N. (1959). "Regulatory mechanism of the pitch and volume of voice," *Oto-Rhino-Laryngology (Kyoto)* 52, 1065-1094.
- Izedebski, K., and Shipp, T. (1976). "Voluntary reaction times for phonatory initiation," *J. Acoust. Soc. Am.* 60, S66(A).
- Ladefoged, P. (1963). "Some physiological parameters in speech," *Lang. and Speech* 8, 109-119.
- Landis, C., and Hunt, W. A. (1939). *The Startle Pattern* (Farrar and Rinehart, New York).
- Lieberman, P., Knudson, R., and Mead, J. (1969). "Determination of the rate of change of fundamental frequency with respect to subglottal air pressure during sustained phonation," *J. Acoust. Soc. Am.* 45, 1537-1543.
- Lukatella, G. (1973). "Pitch determination by adaptive autocorrelation method," *Haskins Labs. Status Rep. on Speech Res.* SR-33, 185-194.
- Miller, A., Hanson, R. J., and Shulak, C. (1978). "Sensorimotor integration just prior to F_0 change," *J. Acoust. Soc. Am.* 64, S90(A).
- Netsell, R., and Daniels, Billie (1974). "Neural and mechanical response time for speech production," *J. Speech Hear. Res.* 17, 608-618.
- Ohman, S., and Lindqvist, J. (1966). "Analysis-by-synthesis of prosodic pitch contours," in *Quarterly Progress and Status Report, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm STL-QPSR* 1/1966, 1-6.
- Sawashima, M. (1974). "Laryngeal research in experimental phonetics," in *Current Trends in Linguistics, Vol. 12: Linguistics and Adjacent Arts and Sciences*, edited by T. A. Sebeok *et al.* (Mouton, The Hague), pp. 2303-2348.
- Sears, T. A., and Newsom Davis, J. (1968). "The control of respiratory muscles during voluntary breathing," *Ann. NY Acad. Sci.* 155 (Art. 1), 183-190.
- van den Berg, Jw. (1957). "Sub-glottal pressure and vibrations of the vocal folds," *Folia Phoniatr.* 9, 65-71.
- van den Berg, Jw., and Tan, T. S. (1959). Results of experiments with excised larynxes, *Pract. Oto-Rhino-Laryng. (Basel)* 21, 425-450.