Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice

Eva B. Holmberg
Boston University, Department of Communication Disorders, 48 Cummington Street, Boston, Massachusetts 02215; Massachusetts Institute of Technology, Research Laboratory of Electronics, Building 36, Room 521, Cambridge, Massachusetts 02139; and University of Stockholm, Institute of Linguistics, Department of Phonetics, S-10691 Stockholm, Sweden

Robert E. Hillman
Boston University, Department of Communication Disorders, 48 Cummington Street, Boston, Massachusetts 02215 and Massachusetts Institute of Technology, Research Laboratory of Electronics, Building 36, Room 511, Cambridge, Massachusetts 02139

Joseph S. Perkell
Massachusetts Institute of Technology, Research Laboratory of Electronics, Building 36, Room 543, Cambridge, Massachusetts 02139

(Received 17 November 1987; accepted for publication 3 May 1988)

Measurements on the inverse filtered airflow waveform (the "glottal waveform") and of estimated average transglottal pressure and glottal airflow were made from noninvasive recordings of productions of syllable sequences in soft, normal, and loud voice for 25 male and 20 female speakers. Statistical analyses showed that with change from normal to loud voice, both males and females produced loud voice with increased pressure, accompanied by increased ac flow and increased maximum airflow declination rate. With change from normal voice, soft voice was produced with decreased pressure, ac flow and maximum airflow declination rate, and increased dc and average flow. Within the loudness conditions, there was no significant male–female difference in air pressure. Several glottal waveform parameters separated males and females in normal and loud voice. The data indicate higher ac flow and higher maximum airflow declination rate for males. In soft voice, the male and female glottal waveforms were more alike, and there was no significant difference in maximum airflow declination rate. The dc flow did not differ significantly between males and females. Possible relevance to biomechanical differences and differences in voice source characteristics between males and females and across loudness conditions is discussed.

PACS numbers: 43.70.Aj, 43.70.Jt

INTRODUCTION

The main purpose of this study is to establish quantitative group data for a number of measures of vocal function for normal male and female speakers. Measures of subglottal air pressure, glottal airflow, and sound intensity were derived from intraoral air pressure, oral airflow, and sound pressure signals. The signals were recorded noninvasively during the pronuncation of specially constructed syllable strings at three levels of vocal effort (soft, normal, and loud voice). The results are to be used in two ways: first, to understand more about normal vocal function, and, second, as norms for the evaluation of pathological voices. This report is concerned with the first of these applications.[1] Group means, variances, and correlations of measures are used to infer mechanisms that underlie (1) differences in sound pressure level (SPL) for normal versus loud voice and normal versus soft voice, (2) differences between male and female voices, and (3) relationships among measured parameters.

I. BACKGROUND

Vocal fold vibration results from an alternating balance between subglottal air pressure that drives the vocal folds apart and muscular, elastic, and Bernoulli restoring forces that draw them together. The movements of the vocal folds take the form of a surface wave that propagates in an upward direction (cf. Stevens, 1977; Baer, 1981). To a certain extent, these movements can be characterized by the behavior of a two-mass model (Ishizaka and Matsudaira, 1968; Ishizaka and Flanagan, 1972). This type of model and others (cf. Titze, 1981) are useful in describing the effects of forces acting within and upon the focal folds (Stevens, 1977). The levels of these forces depend on interactions among the amount of transglottal pressure drop, separation of the vocal folds, and tension on and stiffness in the folds. Vibration may occur over a range of values of these parameters, but the detailed pattern of vibration, airflow waveform, and acoustic characteristics of the resulting sound source will depend on particular conditions, or sets of values, of the parameters (cf. Stevens, 1977). The complicated interactions among aerodynamic, biomechanical, physiological, and acoustic factors that characterize the glottal vibration pattern are far from being understood in detail (cf. Fant, 1982). These interactions are currently the subject of a great deal of research on details of the vibratory mechanism (Baer *et al.*, 1987; Titze and Scherer, 1985; Bless and Abbs, 1983; Stevens and Hirano, 1981).

 0001-4966/88/080511-19$00.80

Different techniques are used in studying glottal function, resulting in different kinds of representations of the vocal fold movements (Hillman and Weinberg, 1981). With some techniques the vibratory pattern of the vocal folds is represented by a waveform as a function of time. Studies have been made of how such a waveform varies with intensity (cf. Monsen and Engebretson, 1977; Baer *et al.*, 1983) and subglottal pressure (Rothenberg, 1973) and of waveform differences between male and female speakers (cf. Monsen and Engebretson, 1977; Karlsson, 1985; Cheng and Guerin, 1987). Some general waveform characteristics have been found for intensity variation: In comparison with normal intensity, a waveform in high intensity has a briefer closing portion and a more abrupt closure (Monsen and Engebretson, 1977), indicating that the folds close simultaneously over a large area along the entire length of the membranous portion (Baer *et al.*, 1983). In low intensity, the waveform has been found to be more symmetrical (Monsen and Engebretson, 1977), with the opening portion only slightly longer than the closing portion. The relative length of the closed portion for high and low intensity in comparison with normal intensity has been found to vary across speakers (Baer *et al.*, 1983). There are few studies of male–female waveform differences. Male waveforms have been found more asymmetrical, with the closing portion shorter than the opening portion, while female waveforms have shown a more equal duration of opening and closing portions (Monsen and Engebretson, 1977; Cheng and Guerin, 1987).

Since transglottal air pressure provides the energy for vocal fold vibration, the relationship between transglottal air pressure and vocal intensity is relevant to the study of vocal function. In addition, measurements of average airflow have been obtained, with special interest for studies of pathological voice function. Both air pressure and airflow have been found to increase with increased intensity (Isshiki, 1964; Schutte, 1981; Tanaka and Gould, 1983).

Large interspeaker variation has been found in measurements of glottal function (cf. Monsen and Engebretson, 1977; Schutte, 1981). Therefore, in order to establish reliable norms for vocal function, it is necessary to study large groups of speakers. However, as a result of the inaccessible location of the larynx and vocal folds, most techniques for studying vocal function are more or less invasive; therefore, most studies have reported data on a relatively small number of subjects. The intent of the present study was to use some of the ideas from previous work and, with a noninvasive technique, study a relatively large number of speakers. The study uses a technique that allows us to compare glottal airflow waveform measures with an indirect measure of transglottal air pressure, thus providing new insight into functional differences between male and female voices and across intensity conditions.

II. METHODS

In this study, we use speech material and instrumentation that allow us to derive indirect estimates of transglottal air pressure and glottal airflow from recordings of intraoral air pressure and oral airflow (Rothenberg, 1973; Smitheran and Hixon, 1981; Lofqvist *et al.*, 1982). SPL is derived from the recorded sound pressure signal. A number of measures are calculated from the derived signals and those measures are analyzed statistically. In the following subsections we describe the experimental design. Section II A describes the subjects, speech material, the transduced signals, the calibrations, and the recording procedure. Section II B gives a detailed description of the computer processing and analyses of the recorded signals and the measures derived from those signals. Section II C describes the statistical analyses that are used to examine male–female differences; differences between soft, normal, and loud voice; and relationships among the parameters.

A. Recordings

1. Subjects

Forty-five adult American speakers (25 males and 20 females) with no history of speech, voice, or hearing problems served as subjects. The males, who ranged in age from 17–30 years, had an average age of 22 years and 5 months. The average age for the females was 24 years and 3 months, with a range of 18–36 years. The subjects were nonsmokers with no professional speaking or singing training. Speakers with nasal[2] or creaky voice were excluded. The number of speakers who had to be excluded for these reasons was minimal.

2. Speech material

The speech material consisted of strings of five repetitions of the syllable /pæ/. The subjects were asked to produce the syllable strings in five speech conditions: three levels of vocal effort (henceforth called "loudness conditions"), namely, normal, soft, and loud voice, and two pitch conditions, low and high pitch.[3] Since we wanted the productions to be as similar as possible to the subjects' natural speech, the subjects were not asked to attain prescribed levels of loudness (and pitch), but instead were free to use a "comfortable" level for each condition. The only instructions given to the subjects vis-á-vis the loudness condition were to not whisper in soft voice or shout in loud voice. The syllable strings were repeated five times per condition. Intraoral air pressure (for the occlusion of the stop consonant /p/), oral volume velocity (for the vowel /æ/), and the acoustic signal were recorded as described below. The detailed rationale for the use of this speech material is given below in the discussion of the measurements.

3. Transduced signals

a. Intraoral air pressure. Intraoral air pressure was transduced using a thin, short catheter, with one end passed between the subject's lips into the oral cavity and the other end connected to a differential pressure transducer (Glottal Enterprises) with a flat frequency response up to about 30 Hz. The catheter passes through a fitting in the face mask of the flow transducer (see below) that holds it in the correct position with respect to the subject's lips.

b. Oral volume velocity. Oral volume velocity was transduced with a high time-resolution pneumotachograph attached to a circumferentially vented face mask (Rothenberg, 1977).[4] The venting screen provides a linear flow resistance such that the volume velocity of the air passing through the screen is linearly proportional to the pressure drop across the screen. This transducer system has a frequency response that is flat (± 2 dB) from zero to over 1 kHz (Rothenberg, 1977). The accuracy of all flow measurements depends on a tight seal between the mask and the subject's face. As described below, great care was taken during the recordings to assure a tight seal.

At the time of the recording the flow signal was low-pass filtered at 900 Hz by a linear phase (eight-pole Bessel) filter in order to eliminate the effects of vocal tract resonances above the first formant.[5]

c. Sound pressure. Sound pressure was transduced with a small electret microphone (Sony model ECM 50) that was affixed to the flow pneumotachograph so as to maintain a constant reproducible distance of 15 cm from the subjects' lips. The output from the microphone was amplified by a Shure model M67 microphone mixer/preamplifier. To cover the entire range of SPL that could be produced from soft to loud voice, the four channels of the microphone mixer were used, each at a different gain setting, permitting the optimal use of the dynamic range of the tape recorder. The recorder VU meter was used to choose the appropriate channel, which was then selected with a rotary switch.

The signals were recorded on FM tape using a Honeywell 5600E FM tape recorder at a speed of $7\frac{1}{2}$ in./s ($\pm 0.3\%$ of full scale, dc—5 kHz).

4. Calibration

Calibrations of the air pressure, airflow, and sound pressure signals were recorded at the beginning of each recording session (in the same sound isolated booth used for the recordings of the subjects; see below).

a. Air pressure: A series of six pressure levels, from 0–30 cm H_2O, was produced with a pressure source (syringe) and monitored with a manometer.

b. Airflow: dc flow: A series of six dc flow levels, from 0–0.395 l/s was generated using a flow source from an air tank and a rotometer to monitor flow level. A custom device was constructed to couple the flow source to the face mask of the transducer system. Prior to each calibration level, a zero flow level was recorded with the calibration device removed from the mask in order to minimize errors due to potential drift in the flow transducer system.

c. SPL: A tone generator connected to a loudspeaker was used to generate the sound calibration tones (sinusoidal, 300 Hz). An SPL meter (on a linear scale, 400-ms averaging time) was placed alongside the microphone to monitor the sound pressure levels. For each of the microphone channels, a zero and four sound levels were recorded (with a range from 70–100 dB SPL for all channels combined) while the loudspeaker was held tightly to the face mask.

The calibration signals were subsequently digitized and processed in the same way as signals from subjects (see below). Data were extracted from the calibration signals and linear regression was used to derive scale factors relating extracted values (in arbitrary, software-based units) to actual values, i.e., cm H_2O/unit, (l/s)/unit, and dB (SPL)/unit.

5. Recording procedure

The experiment was conducted inside a sound-isolated booth. One experimenter inside the booth prompted the subject and held the face mask against the subject's face firmly, taking care to assure a good fit and a tight seal so as to prevent a dc flow leak.[6] Outside of the booth, a second experimenter used a multichannel oscilloscope to monitor the signals as they were recorded to ensure that all the experimental criteria were met. Before the start of the recording, the subject practiced the speech task, coached by the experimenters.[7] The time needed for practice was generally short. At recording, the speech conditions were produced in the following order: normal loudness, softer than normal, and louder than normal (and low- and high-pitched voice), with a short break for instructions and practice between conditions.

Before each syllable string, the face mask was removed from the subject's face for a zero reference in subsequent analysis (in order to minimize errors due to possible drift in the flow transducer system).

B. Derived signals and measures

The air pressure, airflow, and sound pressure signals were low-pass filtered, digitized at 256 samples per second (sps), and processed to produce "low-bandwidth" signals.[8] From these studies, "low-bandwidth data" were extracted interactively for the middle three tokens of the five syllable strings from each loudness condition (making 15 tokens per loudness condition). For each measure, the resulting 15 values were averaged to produce a single averaged value for each loudness condition. A second copy of the airflow signal was (again) low-pass filtered (at 900 Hz; see below, Sec. II B *2*), sampled at 8192 sps, and processed to produce signals from which "high-bandwidth" or "glottal waveform" data were extracted. All the signals were digitized simultaneously and then demultiplexed. Time alignment was preserved with the use of appropriate delays in digital signal processing (to compensate for delays introduced by hardware and software filtering). In the following paragraphs, detailed descriptions of the processing and data extraction for each signal and descriptions of derived measures are given with reference to Fig. 1, which illustrates the scheme for A/D conversion.

1. Transglottal air pressure

The oral air pressure signal (fifth channel in Fig. 1) was low-pass filtered at 80 Hz (10th-order Butterworth filter for antialiasing purposes) and digitized at 256 sps. Transglottal air pressure was estimated from the intraoral pressure signal, as described below.

Transglottal air pressure is one of the primary variables controlling the vibratory behavior of the larynx (Stevens, 1977). During vowel production, transglottal pressure is equal to the pressure measured just below the larynx, i.e., subglottal presssure, because there is no pressure drop in the unobstructed vocal tract for normal (open) vowel articula-

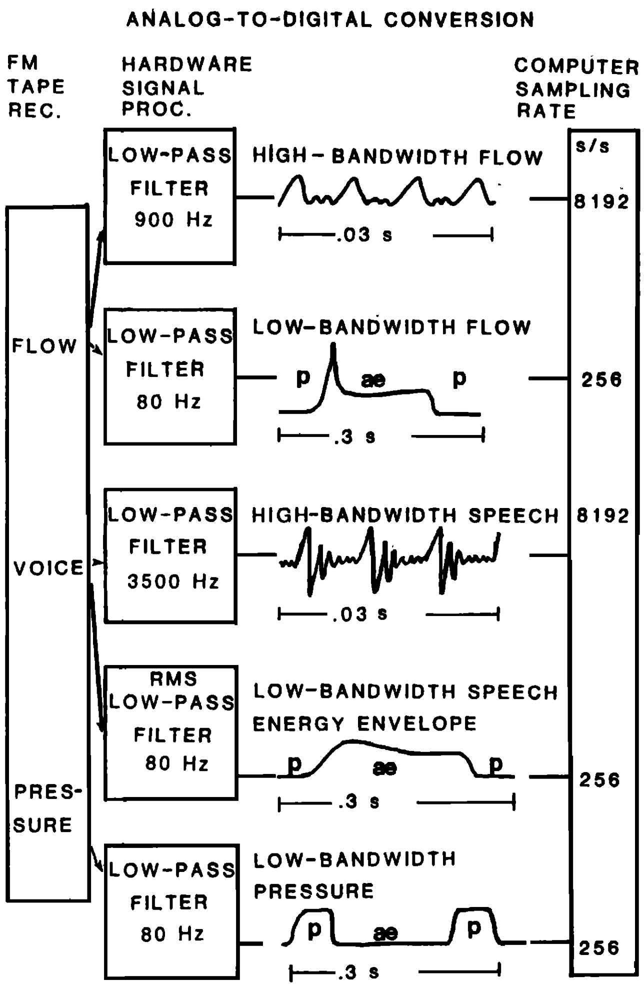

FIG. 1. Scheme for analog-to-digital conversion. The traces are from top to bottom: the high-bandwidth flow signal, the low-bandwidth flow signal, the high-bandwidth speech signal (for listening), the low-bandwidth sound power signal, and the low-bandwidth pressure signal versus time.

tions. Subglottal pressure was not measured directly in this study (e.g., by use of a tracheal puncture; see Isshiki, 1964). Instead, the midvowel subglottal pressure was estimated by a method that makes use of oral pressure measurements in surrounding labial consonants (Rothenberg, 1973).

This technique is illustrated in Fig. 2. Figure 2 shows an example of a low-bandwidth signal stream for one sequence of five syllables elicited from a subject. The traces from top to bottom correspond to intensity, oral volume velocity, and intraoral air pressure versus time. During the /p/ production the vocal folds are spread and the lips occlude the oral cavity. In this case, oral pressure builds up rapidly to a value equal to subglottal pressure—easily reaching subglottal pressure before the plosive is released, at least under normal conditions where the /p/ is indeed voiceless. This is clearly the case for the data of Fig. 2; the pressure trace is nearly rectangular during the presumed closure interval for each /p/ that is surrounded by vowels. Tokens that deviated from this expected pattern were discarded.

Assuming that the subglottal pressure is essentially constant during the production of a /pæ/ syllable string,[9] subglottal pressure during the vowel can be found by drawing a line through the peak oral pressure values observed in each

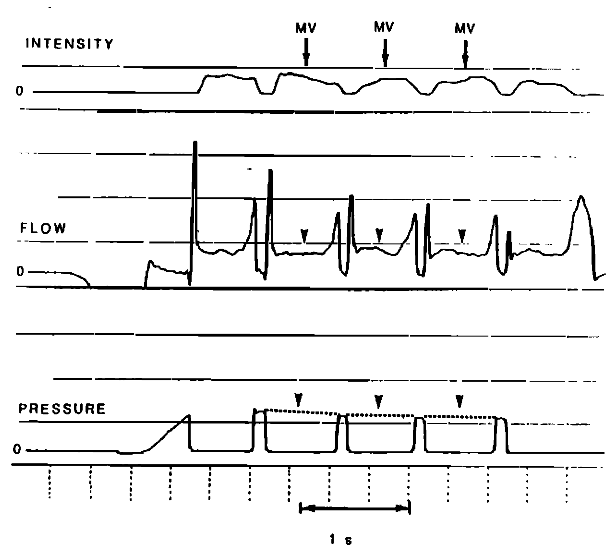

FIG. 2. Example of low-bandwidth signal stream for one sequence of five syllables and low-bandwidth data extraction. The traces from top to bottom correspond to intensity, average oral airflow, and intraoral air pressure versus time.

adjacent consonant (see dashed line in Fig. 2) The value of the interpolated line was measured at the vowel midpoint[10] for each of the central three /pæ/ tokens in the five repetitions of the syllable sequence, obtained for each of the loudness conditions. For each loudness condition, a single average (over 15 tokens) value of pressure in cm H_2O was calculated from the interpolated values using the scale factor obtained from the calibration procedure.

2. Glottal airflow

As mentioned above, the oral airflow signal was digitized at two different rates, yielding low- and high-bandwidth copies of the flow signal. The second "channel" in Fig. 1 shows one copy of the flow signal. It was low-pass filtered at 80 Hz (10th-order Butterworth filter) in order to remove most of the oscillations due to the vocal fold vibrations and it was digitized at a rate of 256 sps. The digitized signal was further low-pass filtered in software (with a 35-pole FIR filter, -3 dB at 10 Hz) and used for measurements of "average flow" through the glottis. Data from this signal were extracted at the vowel midpoint (see Fig. 2), averaged (over 15 tokens), and the value in l/s was calculated using the scale factor obtained from the calibration procedure.

The other copy of the oral airflow signal (top channel in Fig. 1) was low-pass filtered at 900 Hz (a second time, to further assure removal of the effects of formants above $F1$) and digitized at 8192 sps.[11] This signal was inverse filtered with an antiresonator to cancel the effects of the first formant. The 1 token out of the 15 for each loudness condition that was closest to the mean in terms of SPL was located and inverse filtered. At the midvowel location of this token, LPC (linear prediction coding) and DFT (discrete Fourier transform) spectra (62.5-ms Hamming window) were calculated and displayed simultaneously for the interactive extraction of the value of the first formant resonance of the vocal tract transfer function, to be used as the center frequency of the antiresonator. The bandwidth was fixed at 70 Hz (Fant,

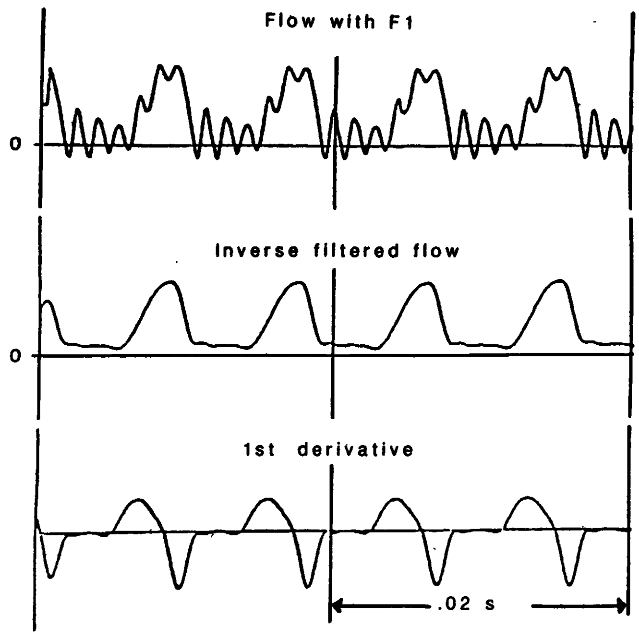

FIG. 3. Example of high-bandwidth flow signals versus time. The top trace shows the flow signal before inverse filtering of $F1$, the middle trace shows the flow signal after inverse filtering, and the bottom trace shows the first derivative of the inverse filtered flow signal. In the two top traces, the horizontal line indicates zero flow level.

1972) because of the known difficulties in estimating formant bandwidths using LPC techniques.[12] The resulting "glottal airflow signal" was then differentiated for measurement of the flow signal's maximum airflow declination rate, the flow measure most closely related to maximum velocity of vocal fold closing movement. Figure 3 shows (from top to bottom) an example of the high-bandwidth airflow signal (after low-pass filtering at 900 Hz) with the effect of $F1$ superimposed on the glottal airflow waveform, the glottal waveform as derived by inverse filtering, and the first derivative of the glottal waveform. In the first two traces, the horizontal straight line represents zero flow as extracted from the zero flow level just before the syllable string.

3. Sound pressure level

Two copies of the acoustic signal were digitized. One copy (Fig. 1, channel 3) was low-pass filtered at 3.5 kHz with an elliptical filter (110-dB/oct roll-off) and digitized at 8192 sps. This signal was only used for listening to the speech during interactive data extraction.

The other copy of the acoustic signal was processed by a hardware root-mean-square (rms) circuit (Analog Devices AD536 true rms to dc converter with an exponential time constant set to 12 ms), low-pass filtered at 80 Hz (10th-order Butterworth filter) in order to remove most of the pitch-synchronous ripple in the signal, and digitized at 256 sps. In software, the signal was squared and further low-pass filtered (with an FIR filter, -3 dB at 5 Hz) to give a signal that corresponds to the mean square of the sound pressure, a quantity that is proportional to intensity (sound power per unit area at a fixed distance from the lips). The midvowel value of the intensity signal was extracted interactively, averaged over 15 tokens and SPL (and other derived measures; see below) was calculated using the scale factor obtained from the calibration procedure.

SPL value depends on complex ways on the transglottal pressure and the way that the vocal folds vibrate. The usual way to measure the SPL of speech is to place a microphone at a standard distance, such as 1 m, from the speaker's lips in a nonreverberant field. In this experiment, the design required the use of the vented face mask with the microphone fixed to the mask handle. Therefore, the effects of the presence of the vented mask and a microphone position rather close to the near field were assessed, by comparing data from one trained speaker with mask in place versus without mask, for two different microphone distances of 15 cm and 1 m from the lips in a sound-treated room. Results of this test are presented in Fig. 4, which shows sound spectra of the low-pass-filtered, digitized acoustic signals for the different recording situations. The effect of the mask was seen mostly in an attenuation of frequencies in the $F3$ and, to a smaller extent, $F1$ regions. In comparing the acoustic spectrum from the

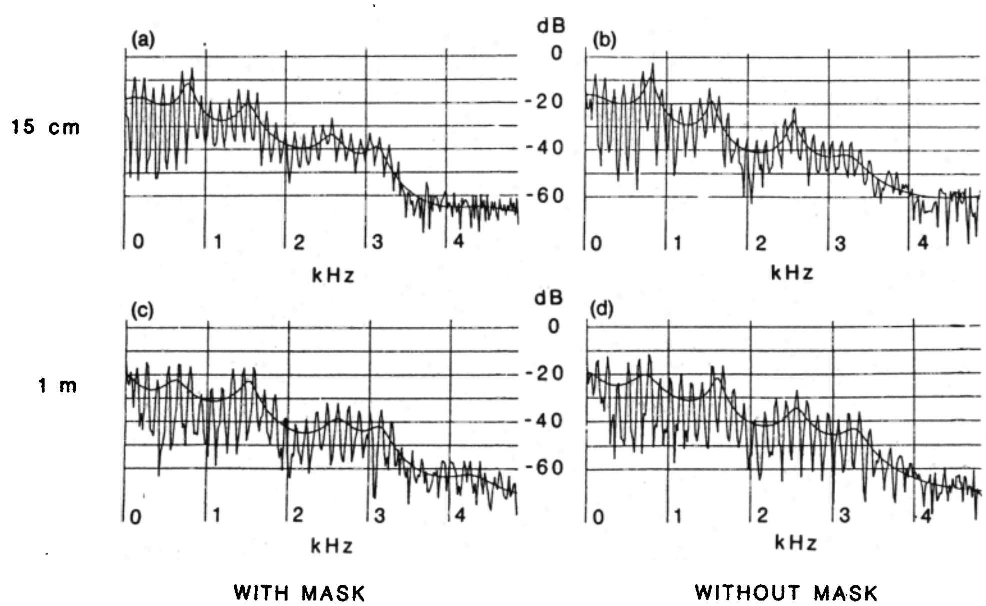

FIG. 4. Sound spectra from recordings of one trained male speaker phonating a sustained vowel /æ/ with different recording procedures. From the top left to the bottom right the spectra shown are from phonation, with the following recording conditions: (a) with the face mask in place, 15-cm lip–microphone distance; (b) without mask, 15-cm distance; (c) with mask, 1-m distance; and (d) without mask, 1-m distance.

standard recording situation (1-m lip–microphone distance, without face mask) with the spectrum from our recording situation (15-cm lip–microphone distance, with face mask), the main difference was a higher amplitude of harmonics in the $F1$ frequency region for our recording procedure.

Although our SPL measure is not standard, it is the same across all subjects and it should be possible to make valid comparisons across subjects and loudness conditions within this study. For comparisons with SPL values of other studies, some care may be appropriate.

4. Other measures dervived from the low-bandwidth air pressure, airflow, and sound pressure signals

The values of average pressure, flow, and sound pressure signals were used to define two general measures of glottal function. A measure of "vocal efficiency" was defined to be the ratio of sound power to the product of air pressure times average flow (Schutte, 1981). A measure of "glottal resistance" was defined to be the ratio of air pressure to average flow (Smitheran and Hixon, 1981).[13]

5. Fundamental frequency

The fundamental frequency was measured for each token, using a PM voice pitch analyzer (Voice Identification, Inc.) and averaged over the 15 tokens per loudness condition.

6. Characterization of the glottal volume velocity waveform

Measurements were extracted interactively from the glottal airflow waveform and its first derivative. The data points for the time and amplitude-based measurements were chosen using consistent, subjective criteria.[14] The measures were chosen with reference to previous research by others (cf. Hirano, 1981) in studies of both normal and pathological voice function. Means and standard deviations for the measurements, described below, were calculated over the four middle glottal cycles for the single most typical token (in terms of SPL) for each loudness condition. Figure 5 shows a schematic diagram of one glottal cycle, marked with the points in time at which data were extracted. In Fig. 5 we also show the time- and amplitude-based flow waveform measures indicated below.

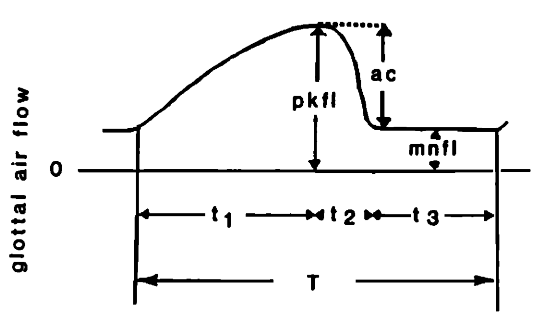

FIG. 5. High-bandwidth data extraction. Schematic diagram of one glottal cycle. The high-bandwidth parameters are the fundamental period (T), open quotient $[(t_1 + t_2)/T]$, speed quotient (t_1/t_2), closing quotient (t_2/T), peak flow (l/s), minimum flow (l/s), ac flow (l/s), and ac/dc (rms of ac flow around its mean/mean of ac). [In addition, maximum airflow declination rate (l/s^2) was extracted as the maximum amplitude of the negative peak from the first derivative.] The measurements were extracted interactively.

The time-based parameters are: the *fundamental period* (T), the *open quotient*, which is the ratio of open time to the period $[(t_1 + t_2)/T]$, the *speed quotient*, which is the ratio of opening to closing time (t_1/t_2), and the *closing quotient*, which is the ratio of closing time to the period (t_2/T).

The flow amplitude-based parameters are: the *peak flow*, which is the amount of flow from zero flow baseline to maximum flow; the *minimum flow*, which is the amount of flow from zero flow to the minimum during the closed phase (or "dc offset"); the *ac flow*, which was calculated as peak flow minus minimum flow; and the *ac–dc ratio*, the rms of the ac portion around its mean divided by the mean of the ac portion, calculated over a 60-ms rectangular time window centered at the middle of the four cycles.

In addition, the *maximum airflow declination rate* was extracted as the maximum amplitude of the negative peak from the first derivative.[15]

For the time-based glottal waveform parameters and maximum airflow declination rate it is important to acknowledge that there is a nonlinear relationship between area waveforms and flow waveforms (Rothenberg, 1973) and that several factors such as acoustical properties of the vocal and subglottal tracts (Rothenberg, 1981), inertia of the moving air column in the vocal tract (Rothenberg, 1973), and horizontal phase differences in the vocal fold movement (cf. Sundberg and Gauffin, 1979; Baer *et al.*, 1983) can affect the airflow waveform. Open quotient and speed quotient have mainly been used to characterize glottal width or cross-sectional waveforms, as derived from high-speed motion pictures (Timcke *et al.*, 1958). While the nonlinear relationship between area and flow waveforms influence our being able to make statements about absolute values of vocal fold closing velocity (represented by the maximum airflow declination rate), speed quotient, and closing quotient, we use these parameters in a comparative manner to make within- and across-subject comparisons between male and female voices and among the different loudness conditions.

C. Statistical analysis

There were three primary goals for the statistical analysis, as follows.

(i) The first goal was to determine the extent to which males and females were significantly different in terms of the glottal waveform and low-bandwidth measures. Simple univariate analysis of covariance procedures was performed within each loudness condition to test for significant male–female differences. In each analysis the covariate was SPL. As already mentioned, subjects were not required to attain prescribed levels of SPL within loudness conditions, but were free to choose levels that felt "comfortable" or "natural." However, it was known that SPL was highly correlated with a number of measures (see the discussion of pairwise relationships in Sec. III B 3 and Appendix B, and thus analysis of covariance with SPL as the covariate was used to reduce variation in the data that was due to intersubject differences in SPL (Winer, 1971; Tabachnick and Fidell, 1983). A primary motivation of this analysis was to correct for the fact that males were on average 3 dB louder than females in all loudness conditions prior to testing for significant male–female differences. This statistically based *post hoc* control for SPL variation should improve the chances of obtaining measures that are more representative of "normal" vocal function rather than having subjects attain a prescribed SPL level. To offset the problem with artificially increased differences that can be associated with such repeated univariate analysis, the tests for significant male–female differences were performed at the conservative level of $p = 0.025$. The mean values after covariation are referred to in the following as the "adjusted" mean values. The unadjusted mean values for males and females in soft, normal, and loud voice are provided in Appendix A.

(ii) The second goal was to determine which glottal waveform and low-bandwidth parameters were significantly altered when loudness was manipulated (normal to loud voice and normal to soft voice). Multivariate (Hotelling's T^2) and univariate t tests for paired comparisons were performed. A conservative level of significance ($p < 0.025$) was used for the univariate t tests.

(iii) The third goal was to determine the extent to which (1) cross-subject variation in SPL was systematically (linearly) related to variation in the glottal waveform and low-bandwidth parameters and (2) parameters were systematically related to one another, both within each loudness condition and for change between conditions. Pearson product moment correlations were calculated to examine simple (pairwise) linear relationships between SPL and the $\log_{10}$[16] of each of the other parameters and for pairs of the other parameters. This was done for males and females separately *within* each of the loudness conditions and for *change* (difference values) between normal and loud and normal and soft voice. Two variables were considered to be highly correlated if r is greater than or equal to 0.70. This means that we account for at least 49% of the variance (i.e., $r^2 = 0.49$).

All statistical analyses were carried out using the BMDP Statistical Software (1983).

III. RESULTS

A. General observations on the glottal airflow waveforms

Figures 6 and 7 show glottal airflow versus time waveforms in soft, normal, and loud voice for six male and six female speakers, marked A–F. The waveforms were chosen to show typical male and female waveforms, as well as to illustrate qualitatively the differences that accounted for a large amount of intersubject variation in quantitative measurements within and across loudness conditions. As seen in Figs. 6 and 7 the inverse filtering result was not always completely successful and residuals of the first formant are superimposed on the waveforms. When this is the case, the formant residuals are seen primarily in loud voice during the "closed" portion of the waveform and should not affect the above-described measurements to any great extent.

Loud voice was typically produced with higher fundamental frequency than normal loudness, while the fundamental frequency in soft voice was either higher or lower. This result was also found by Monsen and Engebretson (1977).

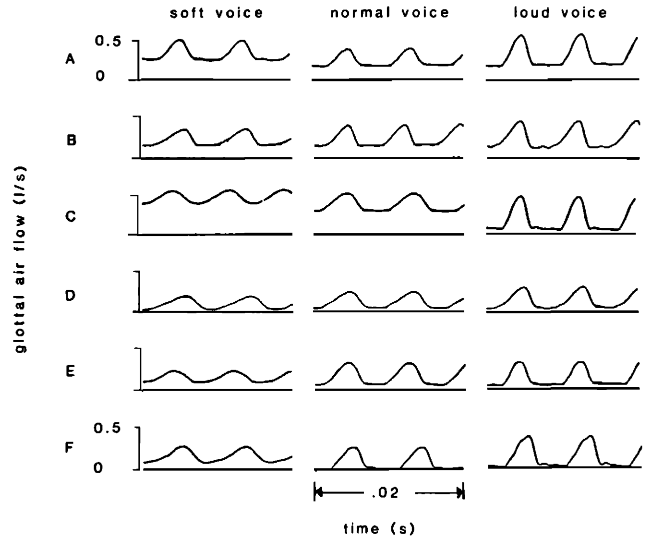

FIG. 6. Glottal airflow versus time waveforms in soft, normal, and loud voice for six male speakers.

In most cases for both male and female speakers, the flow waveform never decreases completely to zero; there is a "dc flow offset" during the "closed" portion of the cycle (cf. Rothenberg, 1973, 1985; Stevens, 1977; Karlsson, 1985). The amount of the dc flow offset varies. For some speakers, the glottal waveforms are well above the flow baseline in all loudness conditions (cf. Fig. 6, speakers A and B; Fig. 7, speakers A–D). For other speakers, the waveforms show little or no dc flow (cf. Fig. 6, speaker C in loud voice, all conditions for speaker D, normal and loud for speakers E and F; Fig. 7, all conditions for speaker E and normal and loud voice for speaker F).[17] In the following, we will call the flat portion of the glottal waveform "the closed portion," even in cases of a dc flow.

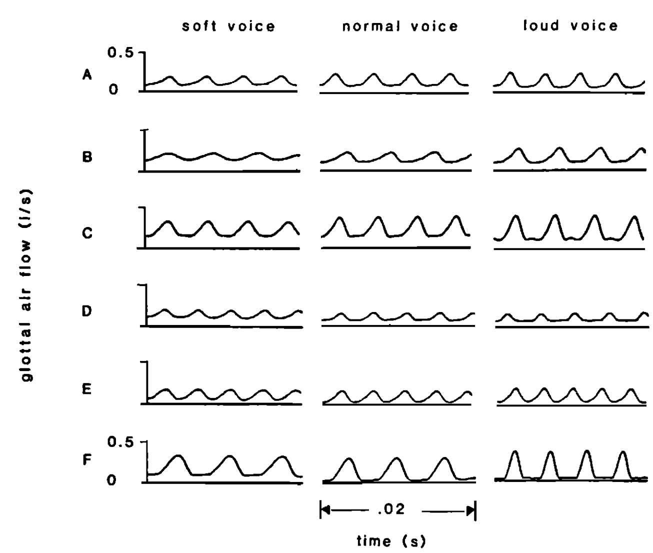

FIG. 7. Glottal airflow versus time waveforms in soft, normal, and loud voice for six female speakers.

TABLE I. Adjusted mean values and for male and female speakers in normal, loud, and soft voice. P values are listed for significant ($p < 0.025$) male–female differences. Per = period (s), mfdr = maximum airflow declination rate (l/s/s), acfl = ac flow (l/s), pkfl = peak flow (l/s), mnfl = minimum flow (l/s), clqu = closing quotient (closing time/T), opqu = open quotient (open time/T), spqu = speed quotient (opening time/closing time), ac/dc = ratio of ac to dc flow (rms of ac flow around its mean/mean ac flow), pres = pressure (cm H_2O, avfl = average flow (l/s), glre = glottal resistance (pressure/average flow), and voef = vocal efficiency (sound power/pressure × average flow).

	Normal			Loud			Soft		
	Male	Female	P	Male	Female	P	Male	Female	P
per	0.0087	0.0048	<0.001	0.0081	0.0044	0.008	0.0087	0.0049	<0.001
mfdr	239.08	117.04	<0.001	418.91	261.32	<0.001	134.15	129.53	···
acfl	0.234	0.141	<0.001	0.346	0.185	<0.001	0.193	0.103	<0.001
pkfl	0.354	0.226	<0.001	0.461	0.271	<0.001	0.367	0.232	<0.001
mnfl	0.101	0.076	···	0.103	0.075	···	0.152	0.109	···
clqu	0.218	0.279	<0.001	0.191	0.279	<0.001	0.301	0.328	···
opqu	0.609	0.735	<0.001	0.569	0.689	<0.001	0.748	0.785	···
spqu	1.794	1.666	···	1.978	1.450	<0.001	1.481	1.382	···
ac/dc	0.519	0.398	0.009	0.671	0.506	0.001	0.318	0.298	···
pres	5.91	6.09	···	8.39	8.46	···	4.79	4.79	···
avfl	0.185	0.139	0.005	0.197	0.140	0.001	0.235	0.162	0.004
glre	32.53	44.33	0.006	43.21	61.08	0.003	20.72	29.69	0.006
voef	8.87[a]	11.48[a]	(0.033)	27.59[a]	38.10[a]	0.005	3.02[a]	4.33[a]	0.019

[a] Figures for vocal efficiency have been multiplied by 10^5.

1. Male versus female glottal waveforms

In all loudness conditions, obvious differences between male and female glottal waveforms are the shorter female period and lower peak and ac flow. In comparison with males, in the normal and loud voice conditions the closed portions of some female waveforms are somewhat less well defined (cf. Fig. 7, speakers B and E) due to a more gradual opening and/or closing. In soft voice, this male–female difference is less obvious.

2. Waveforms in soft versus normal voice

Typically, waveforms in soft voice show a dc flow offset even in cases of no dc flow in normal voice (cf. Stevens 1981; Figs. 6 and 7, speaker F). The opening and closing are more gradual than in normal loudness, making the waveforms more rounded (cf. Baer *et al.*, 1983) and the closed portion relatively shorter (cf. Fig. 6, speakers C–F; Fig. 7, less clearly, speakers A, B, and D). These features cause an increase of the open quotient and suggest a less abrupt closure (and opening) of the vocal folds. In comparison with normal loudness, the maximum flow amplitude in soft voice was sometimes lower (cf. Fig. 6, speakers B, D, and E; Fig. 7, speakers A and C. For other speakers, the entire waveform is elevated from the flow baseline, making minimum, peak, and average flow higher than in normal loudness (cf. Fig. 6, speakers A and C; Fig. 7, less clearly, speakers D and F). However, the *modulated*, *ac part* of the flow waveform is usually lower in soft voice than in normal loudness (cf. Fig. 6, speakers B, D, and E; Fig. 7, less clearly, speakers A and C).

3. Waveforms in loud versus normal voice

For both males and females in loud voice, the waveforms often show a sharp angle between end of closing and beginning of the closed portions, suggesting an abrupt closure of the vocal folds (cf. Fig. 6, speakers A–C and F; Fig. 7, speakers A, C, and F). The closed portion is often well defined and sometimes relatively longer, which would result in a smaller open quotient (cf. Fig. 6, speakers A, C., and E; Fig. 7, speakers C and F). Differences in asymmetry between normal and loud voice are not always obvious. However, while most male waveforms are skewed to the right in loud voice, some female waveforms are nearly symmetrical (cf. Fig. 7, speakers A and E) and others are even skewed to the left (Fig. 7, speaker D). Both peak and ac flow are often higher than in normal loudness (cf. Fig. 6, speakers A, B, D, and F; Fig. 7, speakers B, C, E, and F). For many voices, there is no obvious difference in dc flow offset between normal and loud voice. However, for voices with relatively high dc offset in normal loudness, dc flow is often decreased in loud voice, which may result in lower peak flow than in normal loudness. Even in these cases, the *modulated*, *ac portion* of flow is higher than in normal loudness (cf. Fig. 6, speaker C).

B. Quantitative results

1. Within loudness conditions: Differences between males and females

As described above (Sec. II C, male–female differences for the glottal waveform and low-bandwidth parameters were tested using univariate analysis of covariance with SPL as the covariate. Table I shows mean values "adjusted" for intersubject differences in SPL within each loudness condition. P values indicate significant male–female differences.

a. Normal loudness. [In Appendix A, Tables AI and AII give summary statistics with unadjusted values in normal loudness for males and females, respectively. The average SPL value was 79.5 (3.3) dB for males and 76.4 (4.0) dB for females. The average $F0$ was 116 (12) Hz for males and 205 (24) Hz for females.[18,19]]

In order to graphically represent the difference between male and female glottal airflow waveforms, stylized waveforms were drawn using adjusted mean flow amplitude and

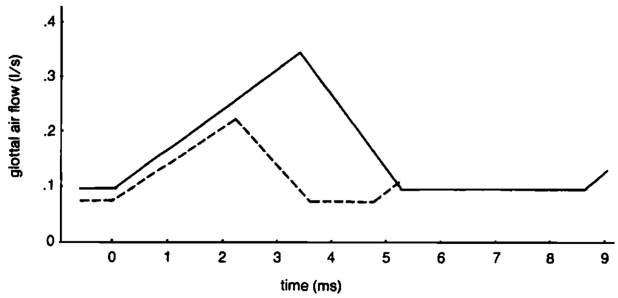

FIG. 8. Stylized glottal waveforms, drawn from mean values adjusted for SPL variation across subjects within the male and female groups in normal loudness: —, males; ---, females.

time-based values. Figure 8 shows such waveforms for male and female voices in normal loudness. Straight lines were drawn between adjusted mean flow values at the times of beginning of opening, peak opening, beginning of closed portion, and beginning of next opening. The resulting stylized waveforms illustrate most of the significant differences between male and female waveforms (although the use of straight lines for the closing portions makes the significant difference in airflow declination rate less apparent than it is in actuality).

As illustrated in Fig. 8 and seen in Table I, in the normal loudness condition there are significant male–female differences for several of the glottal waveform parameters. In comparison with females, males have significantly longer period ($p<0.001$), higher maximum airflow declination rate ($p<0.001$), higher peak flow ($p<0.001$), higher ac flow ($p<0.001$), smaller open quotient ($p<0.001$); smaller closing quotient ($p<0.001$), and higher ac/dc ($p=0.009$). No significant differences were found between males and females for minimum flow or speed quotient.

For the low-bandwidth parameters, males had significantly higher average airflow ($p=0.005$) and lower glottal resistance ($p=0.006$) than females. There were no significant male–female differences in air pressure or vocal efficiency [although the male–female difference for vocal efficiency was close to significant ($p=0.033$)].

b. Loud voice. [In Appendix A, Tables AIII and AIV give summary statistics with unadjusted values in loud voice for males and females, respectively. The average SPL value for males in loud voice was 86.0 (4.3) dB, an 8.9% increase from normal voice. For females the average SPL value was 83.3 (3.2) dB, a 9.2% increase from normal voice.]

Figure 9 shows stylized glottal waveforms for male and females voices in the loud voice condition. The waveforms were drawn in the same way as in Fig. 8.

In the loud voice condition, there are significant male–female differences for all of the glottal waveform parameters except for minimum flow. These differences, shown in Table I, are illustrated in Fig. 9. In comparison with females, males in loud voice have significantly longer period ($p<0.001$), high maximum airflow declination rate ($p<0.001$), higher peak flow ($p<0.001$), higher ac flow ($p<0.001$), smaller open quotient ($p<0.001$), smaller closing quotient ($p<0.001$), larger speed quotient ($p<0.001$), and higher ac/dc ($p=0.001$).

For the low-bandwidth parameters, males have significantly higher average airflow ($p=0.001$), lower glottal resistance ($p=0.003$), and lower vocal efficiency ($p=0.005$) than females, while there is no significant male–female difference in air pressure.

c. Soft voice. [In Appendix A, Tables AV and AVI give summary statistics with unadjusted values in soft voice for males and females, respectively. The average SPL value is 75.0 (2.5) dB for males, a 5.1% decrease of SPL from normal voice. For females, the average SPL value is 71.5 (4.9) dB in soft voice, a 5.3% decrease from normal voice.]

Figure 10 shows stylized male and female glottal waveforms for soft voice. The waveforms were drawn in the same way as in Figs. 8 and 9.

As illustrated in Fig. 10 and shown in Table I, in the soft voice condition there are significant male–female differences for only three of the glottal waveform parameters. In comparison with females, males have significantly longer period ($p<0.001$), higher peak flow ($p<0.001$), and higher ac flow ($p<0.001$). There are no significant male–female differences for the parameters of maximum airflow declination rate, closing quotient, open quotient, speed quotient, minimum flow, or ac/dc.

For the low-bandwidth parameters, males have significantly higher average airflow ($p=0.004$), lower glotttal resistance ($p=0.006$), lower vocal efficiency ($p=0.019$), and lower $F0$ ($p<0.001$) than females. There is no significant male–female difference in air pressure.

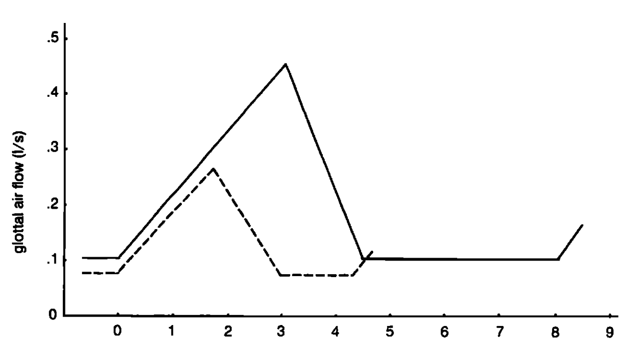

FIG. 9. Stylized glottal waveforms, drawn from mean values adjusted for SPL variation within the male and female groups in loud voice: —, males; ---, females.

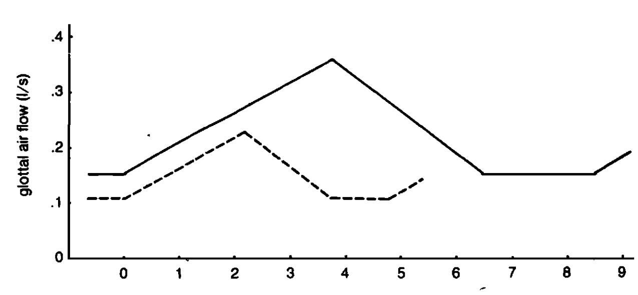

FIG. 10. Stylized glottal waveforms, drawn from mean values adjusted for SPL variation within the male and female groups in soft voice: —, males; ---, females.

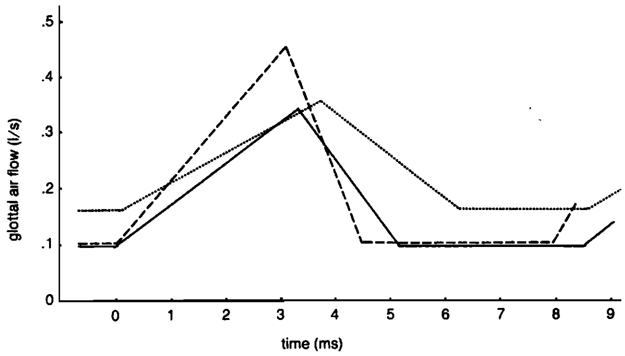

FIG. 11. Stylized glottal waveforms, drawn from group mean values for male voice in normal, loud, and soft voice: —, normal; – – –, loud; ---, soft.

2. Change between loudness conditions

As described above (Sec. II C) multivariate and univariate t tests were used to test for significant differences between loudness conditions.

a. Differences between normal and loud voice. For both males and females, there are significant differences between normal and loud voice for the *set* of glottal waveform parameters (males: $p < 0.0001$; females: $p = 0.0031$) and the *set* of low-bandwidth parameters (males: $p < 0.00001$; females: $p = 0.0014$).

For male voices, Fig. 11 shows stylized glottal waveforms in normal, loud, and soft voice. The waveforms were drawn in the same way as in Figs. 8–10, except that *unadjusted* group means were used.[20] (See Tables AI–AVI.)

As illustrated in Fig. 11, in loud voice for males, significant differences from normal loudness were found for the following individual glottal waveform parameters: period is shorter ($p < 0.0001$), maximum airflow declination rate is higher ($p < 0.0001$), ac flow is higher ($p < 0.0001$), peak flow is higher ($p < 0.0001$), closing quotient is smaller ($p = 0.0001$), open quotient is smaller ($p = 0.0030$), speed quotient is larger ($p = 0.0029$), and ac/dc is higher ($p = 0.0008$). Significant differences were found for the following low-bandwidth parameters: pressure is higher ($p < 0.0001$), glottal resistance is higher ($p = 0.0016$), vocal efficiency is higher ($p < 0.0001$), and $F0$ is higher ($p < 0.0001$). There is no significant difference between male normal and loud voice for the parameters of minimum flow and average flow.

For female voices, Fig. 12 shows stylized glottal waveforms in normal, loud, and soft voice. The waveforms were drawn in the same way as in Fig. 11.

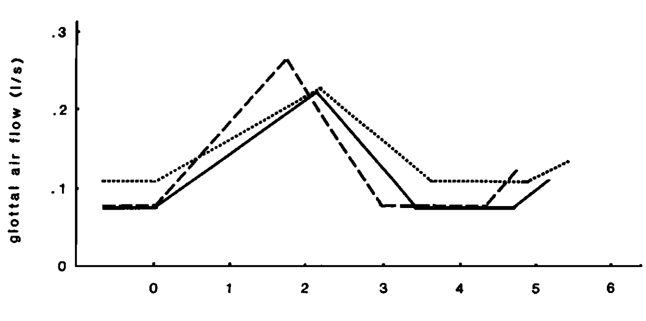

FIG. 12. Stylized glottal waveforms, drawn from group mean values for female voice in normal, loud, and soft voice: —, normal; – – –, loud; ---, soft.

As illustrated in Fig. 12, for females in loud voice as compared with normal loudness, significant differences were found for the following individual glottal waveform parameters: period is shorter ($p = 0.0001$), maximum airflow declination rate is higher ($p < 0.0001$), ac flow is higher ($p < 0.0001$), peak flow is higher ($p = 0.0001$), speed quotient is smaller ($p = 0.0208$), and ac/dc is higher ($p = 0.0004$). Significant differences were found for the following low-bandwidth parameters: pressure is higher ($p < 0.0001$), glottal resistance is higher ($p = 0.0001$), vocal efficiency is higher ($p = 0.0003$), and $F0$ is higher ($p = 0.0001$). There is no significant difference between female normal and loud voice for the parameters of minimum flow, closing quotient, open quotient, or average airflow.

b. Differences between normal and soft voice. For both males and females, there are significant differences between normal and soft voice for the *set* of glottal waveform parameters (males: $p < 0.00001$; females: $p = 0.0020$) and the *set* of low-bandwidth parameters (males: $p < 0.0001$; females: $p < 0.0001$).

For male voices, significant differences between normal and soft voice were found for the following individual glottal waveform parameters: maximum airflow declination rate is lower ($p < 0.0001$), ac flow is lower ($p = 0.0004$), minimum flow is higher ($p = 0.0002$), closing quotient is larger ($p < 0.0001$), open quotient is larger ($p < 0.0001$), speed quotient is smaller ($p < 0.0001$), and ac/dc is lower ($p < 0.0001$). Significant differences were found for the following low-bandwidth parameters: pressure is lower ($p < 0.0001$), average flow is higher ($p = 0.0002$), glottal resistance is lower ($p < 0.0001$), and vocal efficiency is lower ($p < 0.0001$). No significant difference between male normal and soft voice was found for period, peak flow, and $F0$.

For females (Fig. 12), significant differences were found for the following individual glottal waveform parameters: maximum airflow declination rate is lower ($p < 0.0001$), ac flow is lower ($p < 0.0001$), minimum flow is higher ($p = 0.0002$), closing quotient is larger ($p = 0.0007$), open quotient is larger ($p = 0.0045$), speed quotient is smaller ($p = 0.0044$), and ac/dc is lower ($p < 0.0001$). Significant differences were found for the following low-bandwidth parameters: pressure is lower ($p < 0.0001$), average flow is higher ($p = 0.0014$), glottal resistance is lower ($p < 0.0001$), and vocal efficiency is lower ($p = 0.0001$) in soft voice than in normal loudness. As found for males, there is no significant difference between female normal and soft voice for period or peak flow.

3. Pairwise relationships between parameters within and across loudness conditions

Within loudness conditions, there are generally high positive correlations between SPL and maximum airflow declination rate ($r > 0.79$) and between SPL and vocal efficiency ($r > 0.87$). Within all loudness conditions, other consistent high positive correlations were found between maximum airflow declination rate and ac flow ($r > 0.82$), average flow and minimum flow ($r > 0.72$), and vocal efficiency and ac/dc ($r > 0.72$) (with the exception of males in normal voice, where the correlation was low). A high negative correlation was found between glottal resistance and average flow ($r > -0.71$). These results and other high correlations within and across conditions are shown in Appendix B.

IV. DISCUSSION

A. Intrasubject and intersubject variation

As mentioned in Sec. I, there was large intersubject variation within each loudness condition. This finding agrees with findings by Schutte (1981), who found the largest variation for flow, followed by vocal efficiency and subglottal air pressure. Schutte also cites large intersubject variation that was found in several previous studies.

B. Increased SPL from normal to loud voice and decreased SPL from normal to soft voice

1. Pressure

Subglottal pressure can be considered to reflect the aerodynamic driving force for phonation and is the primary factor in raising voice intensity (Fant, 1982). It has been found to range from 5–10 cm H_2O for normal speakers during habitual (comfortable loudness and pitch) phonation (Hirano, 1981). In our study, pressure ranges from 4.4–9.6 cm H_2O across subjects at normal loudness (and pitch). Subglottal pressure has been found to be positively related to sound intensity in several studies (cf. Isshiki, 1964; Schutte, 1981; Tanaka and Gould, 1983). Our results show a significant increase of pressure from soft to normal and normal to loud voice for both males and females.

Others have shown that an increase of subglottal pressure by a factor of 2 is accompanied by an increase in SPL of about 9 dB (Isshiki, 1964; Tanaka and Gould, 1983), in agreement with results from modeling of the voice source (Fant, 1982). In our study, SPL increases about 13 dB for a doubling of pressure. Possible reasons for this discrepancy between previous data and ours are different speech tasks and pressure range. In the studies by Isshiki and Tanaka and Gould, the data are for sustained vowel phonation and the highest pressure values were about 20 cm H_2O, while our mean pressures for syllable repetitions are 9 (males) and 8 (females) cm H_2O, with the highest individual pressure being 15 cm H_2O. Within our pressure range the previous data and ours are roughly in agreement. It is reasonable to believe that results from increased intensity during sustained vowel phonation and productions of syllables in different loudness conditions are not quite compatible and that conditions of soft, normal, and loud voice represent three different "vocal modes" with separate and complex vocal fold adjustments, rather than points along a continuous change from low to high intensity (cf. Monsen and Engebretson, 1977). This idea is also supported by the fact that there is no consistent, high correlation across subjects between pressure and SPL. The loudest voice did not result simply from the highest pressure, but instead most likely from more complex interactions between pressure and vocal fold adjustments (see, also, Sec. III B *3*).

There is no significant male–female difference in pressure in any of the loudness conditions. Thus the observed male–female differences in the glottal waveforms are not primarily due to differences in respiratory forces.

2. Maximum airflow declination rate

The closing phase of the glottal cycle provides the main excitation of the vocal tract (Fant, 1979). The measure of maximum airflow declination rate is the closest we can come to measuring the abruptness of airflow change at the instant of closure, but because of low-pass filtering of the flow waveform at 900 Hz, the relationship of this measure to the magnitude of vocal-tract excitation is not exact.

The results of generally very high positive correlations between maximum airflow declination rate and SPL within loudness conditions agree with the predictions from acoustic theory that a voice of high intensity is primarily a result of high maximum airflow declination rate and a voice of low intensity is a result of low maximum airflow declination rate (cf. Fant, 1979). The results of very high positive correlations between maximum airflow declination rate and ac flow in all loudness conditions agree with the theory that with high flow through the glottis the vocal folds close at a higher maximum rate (van den Berg, 1957). Across loudness conditions, the increased pressure and ac flow for increased loudness is accompanied by an increased airflow declination rate. These results agree with other findings of increase of pressure (Aronson, 1980) and maximum airflow declination rate (Fant, 1979) for increase of intensity.

According to acoustic theory, there should be a 6-dB SPL increase per doubling of the maximum airflow declination rate (Fant, 1979). Our results for males are roughly in agreement with these predictions, while, for females, the SPL increase is at a rate of about 9 dB. We did not expect a male–female difference since this relationship should be independent of factors such as muscle forces and mass differences. One possible reason for the difference could be that the female overall higher $F0$ can increase SPL by about 3 dB. In addition, the $F0$ difference between normal and loud voice is almost twice as large for females as for males, while the $F0$ difference between normal and soft voice is about the same for males and females. There could also be male–female differences in the waveform at the instant of closure that are not measured by our technique. Further work is needed to explore this finding.

In normal and loud voice, males have significantly higher maximum airflow declination rate than females, suggesting male acoustic source spectra with more energy in the high-frequency regions than the female spectra. In soft voice, the maximum airflow declination rate is significantly lower than in normal voice and the angle between end of closing and beginning of closed portions is not sharp, probably due to an incomplete or gradual closure of the membranous part of the vocal folds. The less abrupt closure presumably results in an acoustic source spectrum that has less high-frequency energy in relation to low-frequency energy (Fant, 1979; Stevens, 1981). There is no significant male–female difference in maximum airflow declination rate in soft voice and in combination with the male–female similarities in the time-based quotients, this result suggests that the male and female acoustic source spectra have more similar slopes in soft voice than in normal and loud voice.

3. Amplitude-based flow measures

a. Average flow. Mean values for average flow during vowels for normal voices have previously been reported to be approximately 0.07–0.20 l/s, with large individual ranges (Hirano, 1981). Our group mean values for average flow in normal loudness are within this range. Average flow has been found to be both hgher and lower for males as compared with females in previous studies (Hirano, 1981). In our study, in all loudness conditions, average flow (as well as peak flow and ac flow) is significantly higher for males than for females. This finding is most likely due to the larger male glottis.

Average flow has been found to increase slightly with increased intensity for normal voices (Isshiki, 1964; Schutte, 1981; Tanaka and Gould, 1983). In a high-speed motion picture study, Timcke *et al.* (1958) found a greater peak glottal width for loud voice and a smaller peak width for soft voice. This observation, in combination with the increased pressure for loud voice and decreased pressure for soft voice, might suggest that average flow also should increase in loud voice and decrease in soft voice. Our results show a small, but not significant increase of average flow between the normal and loud voice conditions, but a significant increase of average (and minimum) flow between normal and soft voice. These results suggest that the high subglottal pressure in loud voice was accompanied by increased vocal fold tension and tighter approximation of both the posterior and membranous portions of the vocal folds. On the other hand, the low subglottal pressure in soft voice was accompanied by more relaxed vocal folds, letting more airflow through the entire length of the glottis (cf. also, Secs. IV B 3 b and IV B 3 c), so average (and minimum) flow is increased. However, the average flow is also dependent on the open quotient and future work will explore such interactions among parameters (using multiple regression techniques). High correlations between average and minimum flow and low correlations between average flow and ac flow suggests more of a systematic relationship between minimum flow and average flow than between ac flow and average flow.

b. Minimum flow. As observed in Sec. III, there was almost always a dc flow component of the flow during the "closed" phase of the glottal cycle. Only a few voices showed complete closure of the glottis and, in those cases, it was seen primarily in the loud voice condition. A small dc flow offset of 0.02–0.03 l/s can be due to vertical movements of the vocal folds (Stevens, personal communication). However, in cases of higher dc flow there is likely to be an air leakage through the glottis. Previous results from high-speed motion pictures (Koike and Hirano, 1973) suggest a posterior opening of the glottis, a "chink" between the arytenoid cartilages during normal phonation. In cases of a dc leakage in combination with a well-defined, flat portion during the closed phase (as was often the case in normal and loud voice), the most plausible explanation is that there is closure of the membranous part of the vocal folds and air is being shunted through the posterior "chink" (cf. Rothenberg, 1973, 1985; Stevens, 1977; Hirano *et al.*, 1987). The cartilaginous portion of the glottis has been measured to occupy 35%–40% of the entire glottis (Hirano *et al.*, 1987). Since a perceptually "normal" nonbreathy voice can apparently be achieved with a "chink," there seems to be no reason to adduct this part of the folds completely. On the other hand, in cases of high minimum flow accompanied by a rounded "closed" portion (as is often the case in soft voice), the closure is likely to be over a small portion of the membranous part of the vocal folds, resulting in increased flow through both a "chink" and the membranous portion. The increased flow in soft voice is likely to contribute to the soft voice quality, making it different from normal and loud voice (cf. Sec. IV B 1, "vocal modes").

Across the loudness conditions, minimum flow increased significantly between normal and soft voice for both males and females, while there was no significant difference in minimum flow between normal and loud voice. Within the loudness conditions, there is never any significant male–female difference in the parameter minimum flow. However, in normal and loud voice, it is possible that the presence of a dc flow component could have different acoustic effects for male and female voices. The higher male maximum airflow declination rate that suggests a higher closing velocity and more abrupt closure, in combination with the smaller open quotient, should result in a source spectrum with more energy in higher frequencies than the female voice (cf. Fant, 1979; Isshiki, 1981). The turbulent noise source from the dc flow component through the posterior glottis should have a broadband spectrum (Hillman *et al.*, 1983). In male voices, the turbulence noise would be masked by the higher-amplitude harmonic spectrum of the glottal waveform; however, it is less masked by the more rapidly declining harmonic spectrum of the female glottal waveform. Therefore, the dc flow component of the female voice (cf. Rothenberg, 1986) might be more apparent perceptually as breathiness (cf. Sundberg and Gauffin, 1979; Bickley, 1982) and contribute to the quality of female voice (cf. Klatt, 1986). In soft voice, the lack of significant male–female differences in the airflow declination rate or open quotient should result in more similar spectral slopes and the effect of the noise component should be more similar. Thus the soft male and female voices may possibly sound more alike in terms of breathiness. In order to further explore these male–female differences, the glottal waveform parameters should be analyzed in relation to quantified sound spectral data and perceptual judgments.[21]

c. Peak flow and ac flow. Since there usually seems to be an opening between the arytenoid cartilages during phonation and a resulting dc flow offset, the ac part of the flow waveform provides more accurate information about the magnitude of the vocal fold oscillation than the measurement of peak flow (which includes the dc flow offset). In an inverse filtering study, Sundberg and Gauffin (1979) found the magnitude of the ac flow to be positively related to SPL for individual speakers; our group mean results of significantly increased ac flow with increased loudness condition agree with these previous results.

4. Time-based glottal waveform measures

The speed quotient has been found to increase and the open quotient to decrease as intensity increases (Timcke *et al.*, 1958; Aronson, 1980). Our data for male differences between loudness conditions agree with these previous findings. On the other hand, the data for females do not show these differences between normal and loud voice. A more symmetric waveform has been suggested as being typical for female voice (Fant, 1980; Cheng and Guerin, 1987). In our study in the normal and soft voice conditions, there are no significant male–female differences in symmetry, as indicated by the parameter speed quotient. Only in the loud voice condition are male waveforms more asymmetric (indicated by a larger speed quotient) than the female waveforms.

As mentioned above, the closed portion of the glottal waveform was sometimes poorly defined in the soft voice conditions for both males and females and sometimes also in the normal voice condition for females. As the waveform became more rounded, the decisions for the times of beginning and end of closed portion became more subjective. Therefore, the results for the male soft voice and all voice conditions of the females for the time-based quotients become somewhat uncertain with respect to the underlying physiological mechanisms and relationships to our other measures; they should be interpreted with caution.

5. Glottal resistance

"Glottal resistance" (Smitheran and Hixon, 1981) should indirectly reflect the vocal fold adduction forces. Glottal resistance values between 30–40 cm H_2O/l/s have been reported for syllable productions in comfortable loudness and pitch (Smitheran and Hixon, 1981). In our data, average glottal resistance values for normal loudness were around 40 cm H_2O/l/s, with a large range of 12–93 cm H_2O/l/s for males and females combined.

Several factors may limit the usefulness of the parameter glottal resistance.

(i) The finding of an almost universal dc flow offset and the possibility of a posterior glottal chink suggest that a particular value of average flow can result from glottal openings with different degrees of approximation of the membranous portion of the vocal folds. Therefore, while glottal resistance, measured as the ratio of pressure to *average* flow, may provide a crude index of overall vocal fold adduction, a better index of the underlying adduction forces of the membranous vocal folds might be the ratio of pressure to ac flow.

(ii) Within the loudness conditions, the female glottal resistance values are significantly higher than male values (cf. Shaughnessy *et al.*, 1981). However, since there is no male–female difference in pressure, the higher female glottal resistance is due to their significantly lower average flow that might be expected simply on the basis of the female smaller larynx size. In addition, more frequent high correlations between glottal resistance and flow parameters than between glottal resistance and pressure suggest that glottal resistance primarily reflected the amount of flow through the glottis. Therefore, glottal resistance may not be very useful in comparison between male and female voice function.

(iii) Glottal resistance values have been found to be sensitive to speech rate (Holmberg *et al.*, 1984), making glottal resistance less useful in cross-study comparisons.

(iv) Finally, since ratios of both extremely high and low values of pressure and flow can result in the same resistance value, the parameter of glottal resistance should always be considered in combination with the underlying parameters of pressure and flow.

6. Vocal efficiency and ac/dc

The efficiency of voice production has to be considered in two ways: (1) efficiency in an *aerodynamic/acoustic* sense and (2) efficiency in a *physiological* sense. In other words, an aerodynamically efficient, loud voice may not be optimal for the health of the vocal folds. This is often the case in so-called pressed voices in which high closing velocities may be associated with trauma to the folds. In our study there are two parameters that can be considered measures of the aerodynamic/acoustic efficiency of the voice, namely "vocal efficiency" and the ac–dc ratio.

a. Vocal efficiency. Vocal efficiency (Schutte, 1981; Tanaka and Gould, 1983) may be considered to be the ratio of the output sound power to the underlying aerodynamic power. Schutte (1981) found efficiency values varying from 0.12×10^{-5}–400×10^{-5} for a sound intensity range of 47 dB. In our data, vocal efficiency ranges from 0.35×10^{-5}–145.2×10^{-5}, with mean values of 15.7×10^{-5} (12.7×10^{-5}) for males and 11.2×10^{-5} (9.1×10^{-5}) for females in normal loudness.

b. ac/dc. The ac–dc ratio (Isshiki, 1981) is an inverse measure of "breathiness" that can be a cause of an acoustically inefficient voice. Isshiki (1981) reports ac/dc values around 0.5 as a rough, critical index value for separating the "normal" voice from the "breathy" voice. In our study, the ac/dc ranges are large and mean value for males in normal loudness is slightly above this critical index, while the female mean value is below.

In our study, high correlations between vocal efficiency and ac/dc support the idea that the two parameters reflect similar aspects of voice function. As indicated above, it is important to acknowledge for both parameters that a higher value does not imply better glottal function from a physiological point of view (Isshiki, 1981; Hillman *et al.*, submitted). Both vocal efficiency (Schutte, 1981) and ac–dc ratio (Isshiki, 1981) have been found to increase with increased intensity and our results across the loudness conditions are in agreement with these previous results. Schutte (1981) found higher vocal efficiency values for females than for males. In that study, intensity was controlled at the time of the recording and all subjects phonated at a target value. In our study, the unadjusted vocal efficiency values are higher for males than for females simply because male voices were louder. However, after *post hoc* control for intersubject SPL variation, the relationship between males and females agreed with Schutte's results.

V. SUMMARY AND CONCLUSIONS

Measurements on the inverse filtered airflow waveform and of estimated average transglottal air pressure and average glottal airflow were made for production of sequences of the syllable /pæ/ in soft, normal, and loud voice by 25 male and 20 female normal, adult speakers. The measured parameters were transglottal air pressure, average glottal airflow, "glottal resistance," "vocal efficiency," fundamental frequency and period, maximum airflow declination rate, open quotient, speed quotient, closing quotient, peak flow, ac flow, minimum flow, and ac/dc. Summary statistics were calculated for each parameter in the three loudness conditions for males and females separately and statistical analyses were performed to test for significant differences across the loudness conditions, male–female differences, and systematic linear relationships between pairs of parameters.

Large *intersubject variation* was found for all parameters in each of the loudness conditions. This finding confirms the necessity for studying large subject groups in order to find representative patterns of normal voice function. However, some data suggested that the two large groups of males and females may consist of smaller subgroups in terms of interactions among the parameters. The strategy for increasing and decreasing intensity from normal to loud and normal to soft voice may vary across speakers depending on how normal voice is produced. Therefore, future analyses should look for subject subgroups with similar vocal behavior. Some of the variation was due in part to cross-subject SPL variation within the loudness conditions. Therefore, when it is desirable not to constrain SPL at the time of the recording, it is necessary to control for intersubject SPL variation in statistical analyses of differences between subject groups.

An almost universal *dc flow offset* was found in the male and female glottal airflow waveforms. The combination of a dc flow and a flat "closed" portion during the cycle suggests that while the membranous portion of the folds was closed, air was being shunted through a posterior opening between the arytenoid cartilages. It is suggested that when there is a rounded "closed" portion, the glottal opening extends into the membranous part of the folds and air flows through both the interarytenoid opening and part of the membranous portion. The frequent finding of a dc flow offset suggests that the parameters average flow and peak flow are less useful than a combination of the parameters ac flow and minimum flow. The possibility of a posterior glottal opening limits the usefulness of the parameter glottal resistance when calculated as the ratio of transglottal pressure to average flow; a ratio of transglottal pressure to ac flow is suggested as a better (indirect) indicator of the component of vocal fold adduction that has the greatest influence on the source spectrum.

The data in this study generally confirm results from other studies of vocal function. However, some data suggested that results may be speech task dependent. It seems plausible that syllable productions in soft, normal, and loud voice may represent phonation in three different "vocal modes," with vocal fold adjustments that are different from those accompanying increasing intensity during sustained vowel productions. It is also likely that dc flow values are somewhat dependent on phonetic context, which in this case consisted of voiceless aspirated stops.

In *change of loudness conditions* there were no large differences between males and females in terms of number of parameters altered between normal and loud and normal and soft voice. Across the three loudness conditions, the increase in SPL was somewhat greater for a given pressure than that found by others (Isshiki, 1964; Tanaka and Gould, 1983). However, within our pressure range, our data and the previous data are in agreement and we conclude that differences in results are speech task dependent.

Across the loudness conditions, the increase of maximum airflow declination rate for males agreed with the prediction of a 6-dB SPL increase per doubling of the vocal fold closing velocity (Fant, 1979), while the female maximum airflow declination rate increased at a rate of 9 dB. The female overall higher $F0$, and the fact that the $F0$ difference between normal and loud voice was almost twice as large for females as for males was suggested as one possible reason for this male–female difference in the relationship between SPL and maximum airflow declination rate.

Within the loudness conditions, in terms of number of parameters, the greatest male–female difference was found for loud voice, followed by normal and soft voice. No significant male–female difference was found for transglottal pressure, indicating that the observed differences in the glottal waveform were not primarily due to differences in the respiratory forces. In addition, there was no significant male–female difference in the parameter minimum flow in any of the loudness conditions. It was suggested that the turbulent noise source from the dc flow component may have different effects on the male and female source spectra, being less masked by the more rapidly dropping harmonic spectrum of the female glottal waveform than that of the male glottal waveform. Thus the aspirated flow component may be perceptually more apparent in female normal and loud voice and contribute more to the voice quality for females than for males.

For *pairwise relationships between parameters*, within the loudness conditions there were generally high correlations between SPL and (log of) maximum airflow declination rate, while correlations between difference values for these parameters in change across loudness condition were low. It was suggested that within each loudness condition, a voice with high airflow declination rate has high intensity, but in change across conditions, several other factors may contribute to resulting loudness. High correlations were found between ac flow and maximum airflow declination rate. Assuming a close relationship between the flow declination rate and the vocal fold closing velocity, this result agrees with the theory of a positive relationship between high flow rate through the glottis and high vocal fold closing velocity (van den Berg, 1957). The results of high correlations between average flow and minimum flow, but not between average flow and ac flow, indicated that average flow values are more systematically related to the dc flow component than to the ac flow.

In this initial attempt to explore relationships between the measured parameters and SPL variation, as well as among pairs of all parameters, we used linear regression analyses. However, it is possible that there are relationships that are not well characterized by a linear model, but are still systematic. In addition, because of interactions among parameters, correlations between sets of parameters and SPL could reveal meaningful relationships. Therefore, future work should include nonlinear, as well as multiple regression analyses in order to look for more complex relationships between the parameters and SPL variation, both within and across loudness conditions.

Finally, the results of this study should be useful as norms for the evaluation of pathological voices. However, in order to gain deeper understanding of the findings observed in this study, future work should also include modeling of vocal fold vibration patterns and acoustic studies, to explore interactions among measured parameters, and male–female differences in relation to voice source spectral differences.

ACKNOWLEDGMENTS

This research was supported by the National Institutes of Health under Grant No. RO1 NS21183. We wish to thank Tom Baer, Dennis Klatt, and Kenneth Stevens for their many helpful comments.

APPENDIX A

TABLE AI. Summary statistics. Mean, standard deviation, and range for glottal waveform and low-bandwidth parameters in normal loudness for 25 male speakers.

	X	s.d.	Range
Values averaged over 15 syllables			
SPL (dB)	79.5	3.3	73.4–85.9
$F0$ (Hz)	116	12	93–135
pressure (cm H_2O)	6.3	1.4	4.2–9.6
average flow (l/s)	0.19	0.07	0.1–0.3
glottal resistance (cm H_2O/l/s)	37.7	16.7	12.3–77.2
vocal efficiency (I/cm $H_2O\times$l/s)	15.7[a]	12.7[a]	2.3–62.0[a]
Values averaged over four cycles of the single inverse filtered token			
SPL (dB)	79.5	3.4	73.3–87.0
period (s)	0.009	0.001	0.007–0.011
maximum airflow declination rate (l/s^2)	279.6	90.4	139.5–470.1
peak flow (l/s)	0.38	0.09	0.21–0.54
ac flow (l/s)	0.26	0.07	0.16–0.42
minimum flow (l/s)	0.12	0.07	0.02–0.29
open quotient	0.60	0.07	0.46–0.77
speed quotient	1.82	0.28	1.32–2.58
closing quotient	0.22	0.04	0.14–0.27
ac/dc	0.60	0.23	0.23–1.12
pressure (cm H_2O)	6.4	1.4	4.3–9.7
average flow (l/s)	0.20	0.7	0.09–0.36
glottal resistance (cm H_2O/l/s)	37.3	17.5	13.8–85.3
vocal efficiency (I/cm H_2O/l/s)	15.7[a]	12.7[a]	2.3–62.0[a]

[a] Figures for vocal efficiency have been multiplied by 10^5.

TABLE AII. Summary statistics. Mean, standard deviation, and range for glottal waveform and low-bandwidth parameters in normal loudness for 20 female speakers.

	X	s.d.	Range
Values averaged over 15 syllables			
SPL (dB)	76.4	4.0	66.9–81.3
$F0$ (Hz)	205	24	162–237
pressure (cm H_2O)	5.8	0.9	4.4–7.6
average flow (l/s)	0.14	0.03	0.09–0.21
glottal resistance (cm H_2O/l/s)	42.2	8.1	27.9–92.5
vocal efficiency (I/cm $H_2O\times$l/s)	11.2[a]	9.1[a]	1.4–37.6[a]
Values averaged over four cycles of the single inverse filtered token			
SPL (dB)	76.4	4.0	66.9–81.4
period (s)	0.005	0.001	0.004–0.006
maximum airflow declination rate (l/s^2)	164.0	57.5	91.0–279.0
peak flow (l/s)	0.22	0.06	0.12–0.37
ac flow (l/s)	0.14	0.05	0.09–0.25
minimum flow (l/s)	0.09	0.04	0.01–0.16
open quotient	0.76	0.10	0.56–0.95
speed quotient	1.65	0.30	1.19–2.33
closing quotient	0.29	0.04	0.21–0.37
ac/dc	0.38	0.13	0.25–0.78
pressure (cm H_2O)	5.8	1.1	3.6–8.1
average flow (l/s)	0.14	0.03	0.09–0.22
glottal resistance (cm H_2O/l/s)	42.7	8.7	29.6–61.9
vocal efficiency (I/cm H_2O/l/s)	10.6[a]	8.5[a]	1.6–35.8[a]

[a] Figures for vocal efficiency have been multiplied by 10^5.

TABLE AIII. Summary statistics. Mean, standard deviation, and range for glottal waveform and low-bandwidth parameters in loud voice for 25 male speakers.

	X	s.d.	Range
Values averaged over 15 syllables			
SPL (dB)	86.0	4.3	79.4–93.0
$F0$ (Hz)	126	14	98–152
pressure (cm H_2O)	9.0	2.4	5.9–15.4
average flow (l/s)	0.20	0.06	0.12–0.34
glottal resistance (cm H_2O/l/s)	48.6	19.4	18.1–105.8
vocal efficiency (I/cm $H_2O\times$l/s)	49.4[a]	49.3[a]	8.9–145.2[a]
Values averaged over four cycles of the inverse filtered token			
SPL (dB)	86.0	4.3	80.1–93.0
period (s)	0.008	0.001	0.007–0.011
maximum airflow declination rate (l/s^2)	481.1	162.6	247.5–894.1
peak flow (l/s)	0.49	0.10	0.28–0.66
ac flow (l/s)	0.38	0.10	0.21–0.59
minimum flow (l/s)	0.11	0.05	0.03–0.26
open quotient	0.57	0.06	0.47–0.75
speed quotient	2.05	0.39	1.48–3.06
closing quotient	0.19	0.03	0.13–0.24
ac/dc	0.74	0.20	0.38–1.09
pressure (cm H_2O)	9.2	2.4	6.2–16.4
average flow (l/s)	0.21	0.06	0.13–0.36
glottal resistance (cm H_2O/l/s)	47.7	19.2	17.4–101.7
vocal efficiency (I/cm H_2O/l/s)	45.8[a]	39.3[a]	6.7–147.3[a]

[a] Figures for vocal efficiency have been multiplied by 10^5.

TABLE AIV. Summary statistics. Mean, standard deviation, and range for glottal waveform and low-bandwidth parameters in loud voice for 20 female speakers.

	X	s.d.	Range
Values averaged over 15 syllables			
SPL (dB)	83.3	3.2	77.1–90.0
$F0$ (Hz)	223	20	185–252
pressure (cm H_2O)	8.2	1.8	6.5–13.1
average flow (l/s)	0.15	0.04	0.08–0.22
glottal resistance (cm H_2O/l/s)	60.9	20.6	35.9–110.9
vocal efficiency (I/cm $H_2O \times$ l/s)	37.7[a]	30.8[a]	9.2–108.4[a]
Values averaged over four cycles of the single inverse filtered token			
SPL (dB)	83.4	3.2	77.4–90.2
period (s)	0.005	0.001	0.004–0.006
maximum airflow declination rate (l/s^2)	248.9	84.0	111.7–466.3
peak flow (l/s)	0.27	0.06	0.16–0.39
ac flow (l/s)	0.18	0.06	0.11–0.32
minimum flow (l/s)	0.09	0.04	0.03–0.19
open quotient	0.71	0.14	0.52–0.93
speed quotient	1.46	0.37	0.78–2.13
closing quotient	0.29	0.07	0.22–0.48
ac/dc	0.50	0.19	0.29–0.98
pressure (cm H_2O)	8.4	1.8	6.4–13.7
average flow (l/s)	0.15	0.04	0.071–0.23
glottal resistance (cm H_2O/l/s)	61.1	19.6	36.6–114.1
vocal efficiency (I/cm H_2O/l/s)	35.8[a]	20.8[a]	7.5–113.5[a]

[a] Figures for vocal efficiency have been multiplied by 10^5.

TABLE AV. Summary statistics. Mean, standard deviation, and range for glottal waveform and low-bandwidth parameters in soft voice for 25 male speakers.

	X	s.d.	Range
Values averaged over 15 syllables			
SPL (dB)	75.0	2.5	70.7–79.9
$F0$ (Hz)	113	12	90–132
pressure (cm H_2O)	5.1	1.2	3.3–7.6
average flow (l/s)	0.25	0.10	0.09–0.42
glottal resistance (cm H_2O/l/s)	23.6	10.4	8.8–54.1
vocal efficiency (I/cm $H_2O \times$ l/s)	5.1[a]	3.6[a]	1.3–13.4[a]
Values averaged over four cycles of the single inverse filtered token			
SPL (dB)	75.0	2.5	70.7–79.9
period (s)	0.009	0.001	0.007–0.011
maximum airflow declination rate (l/s^2)	171.1	71.3	67.6–328.8
peak flow (l/s)	0.40	0.12	0.18–0.69
ac flow (l/s)	0.22	0.07	0.12–0.37
minimum flow (l/s)	0.18	0.10	0.02–0.38
open quotient	0.73	0.10	0.50–0.91
speed quotient	1.56	0.30	1.10–2.46
closing quotient	0.29	0.05	0.18–0.43
ac/dc	0.39	0.17	0.13–0.81
pressure (cm H_2O)	5.0	1.2	3.2–7.7
average flow (l/s)	0.26	0.10	0.08–0.50
glottal resistance (cm H_2O/l/s)	22.3	9.5	8.6–45.6
vocal efficiency (I/cm H_2O/l/s	5.0[a]	3.9[a]	1.1–15.6[a]

[a] Figures for vocal efficiency have been multiplied by 10^5.

TABLE AIV. Summary statistics. Mean, standard deviation, and range for glottal waveform and low-bandwidth parameters in loud voice for 20 female speakers.

	X	s.d.	Range
Values averaged over 15 syllables			
SPL (dB)	83.3	3.2	77.1–90.0
$F0$ (Hz)	223	20	185–252
pressure (cm H_2O)	8.2	1.8	6.5–13.1
average flow (l/s)	0.15	0.04	0.08–0.22
glottal resistance (cm H_2O/l/s)	60.9	20.6	35.9–110.9
vocal efficiency (I/cm $H_2O \times$ l/s)	37.7[a]	30.8[a]	9.2–108.4[a]
Values averaged over four cycles of the single inverse filtered token			
SPL (dB)	83.4	3.2	77.4–90.2
period (s)	0.005	0.001	0.004–0.006
maximum airflow declination rate (l/s^2)	248.9	84.0	111.7–466.3
peak flow (l/s)	0.27	0.06	0.16–0.39
ac flow (l/s)	0.18	0.06	0.11–0.32
minimum flow (l/s)	0.09	0.04	0.05–0.19
open quotient	0.71	0.14	0.52–0.93
speed quotient	1.46	0.37	0.78–2.13
closing quotient	0.29	0.07	0.22–0.48
ac/dc	0.50	0.19	0.29–0.98
pressure (cm H_2O)	8.4	1.8	6.4–13.7
average flow (l/s)	0.15	0.04	0.071–0.23
glottal resistance (cm H_2O/l/s)	61.1	19.6	36.6–114.1
vocal efficiency (I/cm H_2O/l/s)	35.8[a]	20.8[a]	7.5–113.5[a]

[a] Figures for vocal efficiency have been multiplied by 10^5.

APPENDIX B

TABLE BI. Significant (r is greater than or equal to 0.70) pairwise correlations for males in normal voice.

SPL–vocal efficiency	$r=0.87$
SPL–maximum airflow declination rate	$r=0.87$
average flow–minimum flow	$r=0.92$
ac/dc–minimum flow	$r=-0.88$
maximum airflow declination rate–ac flow	$r=0.87$
glottal resistance–ac/dc	$r=0.85$
average flow–peak flow	$r=0.84$
open quotient–closing quotient	$r=0.83$
glottal resistance–average flow	$r=-0.80$
glottal resistance–minimum flow	$r=-0.78$
ac/dc–average flow	$r=-0.78$
speed quotient–closing quotient	$r=-0.75$

TABLE BII. Significant (r is greater than or equal to 0.70) pairwise correlations for females in normal voice.

SPL–vocal efficiency	$r=0.93$
SPL–maximum airflow declination rate	$r=0.79$
maximum airflow declination rate–ac flow	$r=0.94$
maximum airflow declination rate–peak flow	$r=0.87$
ac flow–peak flow	$r=0.83$
average flow–peak flow	$r=0.77$
ac/dc–vocal efficiency	$r=0.75$
glottal resistance–average flow	$r=-0.73$
glottal resistance–vocal efficiency	$r=0.73$
ac/dc–ac flow	$r=0.73$
average flow–minimum flow	$r=0.72$
pressure–peak flow	$r=0.70$

TABLE BIII. Significant (r is greater than or equal to 0.70) pairwise correlations for males in loud voice.

SPL–vocal efficiency	$r = 0.93$
SPL–maximum airflow declination rate	$r = 0.89$
SPL–ac flow	$r = 0.81$
maximum airflow declination rate–ac flow	$r = 0.95$
ac flow–peak flow	$r = 0.86$
average flow–minimum flow	$r = 0.85$
ac/dc–minimum flow	$r = -0.85$
maximum airflow declination rate–peak flow	$r = 0.79$
glottal resistance–pressure	$r = 0.75$
ac/dc–vocal efficiency	$r = 0.72$
glottal resistance–average flow	$r = -0.71$
glottal resistance–minimum flow	$r = -0.70$

TABLE BIV. Significant (r is greater than or equal to 0.70) pairwise correlations for females in loud voice.

SPL–vocal efficiency	$r = 0.87$
SPL–pressure	$r = 0.72$
glottal resistance–ac/dc	$r = 0.88$
maximum airflow declination rate–ac flow	$r = 0.84$
ac/dc–vocal efficiency	$r = 0.84$
average flow–minimum flow	$r = 0.83$
glottal resistance–vocal efficiency	$r = 0.82$
glottal resistance–average flow	$r = -0.78$
maximum airflow declination rate–peak flow	$r = 0.77$
ac flow–peak flow	$r = 0.77$
maximum airflow declination rate–ac/dc	$r = 0.76$
open quotient–closing quotient	$r = 0.72$

TABLE BV. Significant (r is greater than or equal to 0.70) pairwise correlation of delta values for change from normal to loud voice for males, where delta = value in loud voice minus value in normal voice.

SPL–vocal efficiency	$r = 0.87$
maximum airflow declination rate–ac flow	$r = 0.91$
ac/dc–minimum flow	$r = -0.85$
speed quotient–closing quotient	$r = -0.76$
peak flow–minimum flow	$r = 0.71$

TABLE BVI. Significant (r is greater than or equal to 0.70) pairwise correlation of delta values for change from normal to loud voice for females, where delta = value in loud voice minus value in normal voice.

SPL–vocal efficiency	$r = 0.97$
SPL–pressure	$r = 0.78$
SPL–maximum airflow declination rate	$r = 0.78$
SPL–$F0$	$r = 0.77$
SPL–ac flow	$r = 0.76$
open quotient–closing quotient	$r = 0.82$
glottal resistance–pressure	$r = 0.80$
glottal resistance–vocal efficiency	$r = 0.77$
peak flow–minimum flow	$r = 0.76$
maximum airflow declination rate–ac/dc	$r = 0.71$

TABLE BVII. Significant (r is greater than or equal to 0.70) pairwise correlations for males in soft voice.

SPL–maximum airflow declination rate	$r = 0.90$
SPL–ac flow	$r = 0.76$
average flow–minimum flow	$r = 0.97$
average flow–peak flow	$r = 0.91$
ac/dc–vocal efficiency	$r = 0.87$
maximum airflow declination rate–ac flow	$r = 0.84$
minimum flow–peak flow	$r = 0.82$
ac/dc–minimum flow	$r = -0.79$
glottal resistance–average flow	$r = -0.77$
open quotient–closing quotient	$r = 0.76$
glottal resistance–vocal efficiency	$r = 0.75$
glottal resistance–minimum flow	$r = -0.72$
ac/dc–average flow	$r = -0.72$
maximum airflow declination rate–closing quotient	$r = -0.71$

TABLE BVIII. Significant (r is greater than or equal to 0.70) pairwise correlations for females in soft voice.

SPL–vocal efficiency	$r = 0.95$
SPL–ac/dc	$r = 0.93$
SPL–maximum airflow declination rate	$r = 0.90$
SPL–open quotient	$r = -0.78$
SPL–closing quotient	$r = -0.74$
ac/dc–vocal efficiency	$r = 0.91$
average flow–minimum flow	$r = 0.89$
peak flow–average flow	$r = 0.86$
glottal resistance–average flow	$r = -0.85$
closing quotient–speed quotient	$r = -0.84$
maximum airflow declination rate–ac/dc	$r = 0.83$
open quotient–closing quotient	$r = 0.83$
maximum airflow declination rate–ac flow	$r = 0.82$
ac/dc–open quotient	$r = -0.76$
ac/dc–closing quotient	$r = -0.74$
open quotient–vocal efficiency	$r = -0.74$
glottal resistance–minimum flow	$r = -0.74$
maximum airflow declination rate–open quotient	$r = -0.73$
ac flow–peak flow	$r = 0.70$

TABLE BIX. Significant (r is greater than or equal to 0.70) pairwise correlation of delta values for change from normal to soft voice for males, where delta = value in normal voice minus value in soft voice.

peak flow–minimum flow	$r = 0.89$
open quotient–closing quotient	$r = 0.80$
peak flow–ac flow	$r = 0.76$

TABLE BX. Significant (r is greater than or equal to 0.70) pairwise correlation of delta values for change from normal to soft voice for females, where delta = value in normal voice minus value in soft voice.

SPL–vocal efficiency	$r = 0.98$
glottal resistance–average flow	$r = -0.89$
peak flow–minimum flow	$r = 0.86$
closing quotient–speed quotient	$r = -0.80$
open quotient–closing quotient	$r = 0.74$

[1]Initial use of these data as norms in a study of voice disorders is presented in: Hillman *et al.* (submitted to J. Speech Hear. Res.).

[2]Vowel nasalization introduces an extra pole-zero pair close to F_1 (Fujimura, 1960) that would interfere with the inverse filtering used to derive glottal airflow waveforms. Therefore, an informal test for nasal coupling was preformed by having the subjects utter a sustained vowel /æ/ with and without the nares occluded, while we listened for differences between the two conditions. Further controls for nasalization, such as measuring nasal airflow or vibration, have been considered, but they introduce methodological complications when recording large groups of subjects.

[3]Both loudness and pitch conditions were recorded and analyzed, but in the following, only loudness conditions are reported on and discussed.

[4]Strictly speaking, such a transducer records a combination of oral and nasal airflow. Since subjects with overt signs of nasalization were excluded, we assume that the transduced flow signal is of oral flow.

[5]Low-pass filtering has the undesirable side effect of "rounding the corners" of any waveform discontinuities that might be present, such as at the instant of vocal fold closure. Fortunately, low-pass filtering at 900 Hz does not affect the general shape of the underlying glottal flow waveform, so this limitation has not had a serious effect on our ability to quantify most waveform characteristics.

[6]In spite of having two mask sizes available, there were a few subjects for whom we could not assure a tight seal between the face and the flexible mask flange. Those subjects were excluded from the study. For future work, a new, more flexible foam gasket from Glottal Enterprises will make it easier to fit the mask well to a wider range of subjects.

[7]Results of a pilot methodological study had shown that control over subject speech rate and mode is essential to obtaining interpretable pressure and flow data. The rate has to be slow enough to obtain steady states in the physiological data for both consonants and vowels, but the mode of production of the syllable string must be continuous without breaks between the syllables (Holmberg *et al.*, 1984).

[8]Digital signal processing and data extraction were implemented with the MITSYN signal processing and data extraction languages (Henke and Perkell, 1982) on a PDP-11/60 computer.

[9]While subglottal pressure has been found to be nearly constant or slowly changing during the production of a string of words (Netsell, 1973), there are sometimes local increases during syllable peaks (Stetson, 1951), especially in stressed syllables, due to bursts of activity in various muscles such as the internal intercostals (Draper *et al.*, 1959). We have no way of knowing whether some of our subjects adopted this behavior during performances of the task. However, the subjects were instructed to produce the syllable strings in a monotonous fashion and we believe the potential error in our subglottal pressure estimates at vowel midpoint due to this factor is small.

[10]Vowel midpoint was defined as the temporal midpoint between the pressure peaks. Other measurement points were considered, but rejected as a result of the difficulty of being consistent across all subjects and speech conditions.

[11]Although a signal which is low-pass filtered at 900 Hz need not be sampled at a rate much higher than 1800 Hz (depending on the filter characteristics) to preserve all the contained information, a higher rate was used for the following reasons: The higher sampling rate was necessary for adequate visualization during interactive data extraction on the particular display that was used and an accompanying electroglottographic signal (not analyzed for this report) was low-pass filtered at 3.5 kHz.

[12]Other techniques, such as analysis by synthesis, might be employed in the future to obtain a best guess of formant bandwidth because the bandwidth can vary considerably in a situation like this depending on glottal impedance and the degree of acoustic coupling to the trachea. Bandwidth values ranging from 60–400 Hz have been observed in comparable phonetic environments in a study of a number of male and female speakers using analysis by synthesis (Klatt, 1987). Misestimation of formant bandwidth will lead to a waveform ripple at the formant frequency that is most noticeable during the closed phase of the inverse filtered airflow signal, but is present to some degree throughout the period.

[13]Because of the complex, time-varying nature of transglottal air pressure and flow, these parameters, which incorporate average pressure and flow, are not equivalent to strictly defined measures of efficiency and resistance. Nevertheless, they were considered to provide crude approximations to the efficiency with which input power is converted to sound and the resistance to flow through the glottis.

[14]In the beginning of the data collection, criteria for determining events in the time domain were determined and in order to avoid variability due to differences in judgments, one investigator extracted the data for all the subjects. Times of closing and opening were determined by visually estimating the locations of the intersections of lines tangent to the closing, closed, and opening phases of the waveform. This strategy assumes that there is always a closed portion, even for waveforms with a minimum that is rounded. Future work will explore alternative time-based measurements (cf. Colton *et al.*, 1985).

[15]A measure of maximum rate of airflow increase (the maximum amplitude of the positive peak of the flow derivative) was not used since the portion of increasing flow of the glottal cycle sometimes contained $F1$ residuals, making measurements during this phase unreliable.

[16]Transformation to the logarithmic domain was used because (1) SPL is $\log_{10}$ transformation of intensity, the underlying physical measure to which the other parameters might be linearly related; and (2) other investigators (cf. Isshiki, 1964; Schutte, 1981) have found linear relationships among SPL and $\log_{10}$ of measured parameters.

[17]The fact that in some cases the flow declines to a zero value provides some experimental support of Rothenberg's (1977) claim that the transducer system has a flat frequency response in our operating range. In other words, the system appears to have equivalent ac and dc gains under experimental conditions and the findings of dc offsets appear to be valid.

[18]Numbers in parentheses are standard deviations.

[19]SPL and $F0$ values from the low-bandwidth data, averaged over 15 tokens, may differ somewhat from values in the high-bandwidth data, averaged over four cycles of one token.

[20]Unadjusted means are used for comparisons between loudness conditions.

[21]It is likely that the vowel in our speech sample /pæp/ which is in between two voiceless aspirated stop consonants is breathier than it would be in a voiced context. The degree of context-dependent vowel aspiration could have varied across speakers, contributing to interspeaker variation in minimum flow. Future work should explore this issue by examining minimum flow results from additional utterance materials such as strings of the syllable /bæ/ or yæ/, in which there should be minimal context-induced aspiration.

Aronson, A. E. (**1980**). *Clinical Voice Disorders. An Interdisciplinary Approach* (Thieme–Stratton, Inc., New York), pp. 34–35.

Baer, T. (**1981**). "Investigation of the Phonatory Mechanism," in *Proceedings of the Conference on the Assessment of Vocal Pathology, ASHA Reports 11*, edited by C. L. Ludlow and M. O.-C. Hart (American Speech and Hearing Association, Rockville, MD), pp. 38–47.

Baer, T., Sasaki, C., and Harris, K. (Eds.) (**1987**). *Laryngeal Function in Phonation and Respiration, Vocal Fold Physiology Series*, volume based on presentations delivered at the Fourth International Vocal Fold Physiology Conference, 1986, New Haven, CT (College-Hill, Little Brown, Boston), 574 pp.

Baer, T., Titze, I. R., and Yoshioka, H. (**1983**). "Multiple Simultaneous Measures of Vocal Fold Activity," in Bless and Abbs (1983), pp. 229–237.

van den Berg, Jw., Zantema, J. T., and Doornenbal, P., Jr. (**1957**). "On the Air Resistance and the Bernoulli Effect of the Human Larynx," J. Acoust. Soc. Am. **29**, 626–631.

Bickley, C. A. (**1982**). "Acoustic Analysis and Perception of Breathy Vowels," Speech Communication Group, Research Laboratory of Electronics, MIT, Cambridge, Work. Pap. **2**, 71–82.

Bless, D. M., and Abbs, J. H. (Eds.) (**1983**). *Vocal Fold Physiology: Contemporary Research and Clinical Issues, Vocal Fold Physiology Series*, Volume based on presentations delivered at the Second International Vocal Fold Physiology Conference, 1981, Madison, WI (College-Hill, San Diego, CA), 466 pp.

Dixon, W. J. (Ed.) (**1983**). *BMDP Statistical Software* (Univ. of Calif., Los Angeles, CA), 733 pp.

Cheng, Y. M., and Guerin, B. (**1987**). "Control Parameters in Male and Female Glottal Sources," in Baer *et al.* (1987).

Colton, R. H., Brewer, D. W., and Rothenberg, M. (**1985**). "Vibratory Characteristics of Patients with Voice Disorders," oral presentation delivered at the Symposium on Voice Acoustics and Dysphonia, Gotland, Sweden, August 1985.

Draper, M. H., Ladefoged, P., and Whitteridge, D. (**1959**). "Respiratory Muscles in Speech," J. Speech Hear. Res. **2**, 16–27.

Fant, G. (**1972**). "Vocal Tract Wall Effects, Losses and Resonance Bandwidths," Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden, QPSR/2-3, 28–52.

Fant, G. (**1979**). "Glottal Source and Excitation Analysis," Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden, QPSR/1, 85–107.

Fant, G. (**1980**). "Voice Source Dynamics," paper presented at the 10th International Congress on Acoustics, Sydney, Australia.

Fant, G. (**1982**). "Preliminaries to Analysis of the Human Voice Source," Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden, QPSR/4, 1–27.

Fujimura, O. (**1960**). "Spectra of Nasalized Vowels," Speech Communication Group, Research Laboratory of Electronics, MIT, Cambridge, MA, QPR 58, 214–218.

Henke, W., and Perkell, J. S. (**1982**). "Application of Signal Processing Utilities to Speech Physiology Data," J. Acoust. Soc. Am. **71**, 185–201.

Hillman, R. E., and Weinberg, B. (**1981**). "Estimation of Volume Velocity Waveform Properties. A Review and Study of Some Methodological Assumptions," in *Speech and Language: Advances in Basic Research and Practice,* edited by N. Lass (Academic, New York), pp. 411–473.

Hillman, R. E., Osterle, E., and Feth, L. L. (**1983**). "Characteristics of the Turbulent Noise Source," J. Acoust. Soc. Am. **74**, 690–694.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., and Vaughan, C. (**1988**). "Objective Assessment of Vocal Hyperfunction: An Experimental Framework and Preliminary Results," submitted to J. Speech Hear. Res.

Hirano, M. (**1981**). *Clinical Examination of Voice: Disorders of Human Communication 5* (Springer, New York), 100 pp.

Hirano, M., Yoshida, T., Kurita, S., Kiyokawa, K., Sato, K., and Tateishi, O. (**1987**). "Anatomy and Behavior of the Vocal Process," in Baer *et al.* (1983), pp. 1–13.

Holmberg, E. B., Perkell, J. S., and Hillman, R. E. (**1984**). "Methods for Using a Noninvasive Technique for Estimating Glottal Functions from Oral Measurements," J. Acoust. Soc. Am. Suppl. 1 **75**, S7; also published in Speech Communication Group, Research Laboratory of Electronics, MIT, Cambridge, MA, Working Papers **5**, 47–58.

Ishizaka, K., and Matsudaira, M. (**1968**). "Analysis of the Vibration of the Vocal Cords," J. Acoust. Soc. Jpn. **24**, 311–312.

Ishizaka, K., and Flanagan, J. L. (**1972**). "Synthesis of Voiced Sound from a Two-Mass Model of the Vocal Cords," Bell Syst. Tech. J. **51**, 1233–1268.

Isshiki, N. (**1964**). "Regulating Mechanisms of Vocal Intensity Variation," J. Speech Hear. Res. **7**, 17–29.

Isshiki, N. (**1981**). "Vocal Efficiency Index," in Stevens and Hirano (1981), pp. 193–208.

Karlsson, I. (**1985**). "Glottal Waveforms for Normal Female Speakers," Speech Transmission Laboratory, Royal Institute of Technology, Stockholm, Sweden, QPSR/1, pp. 31–36.

Klatt, D. H. (**1987**). Personal communication.

Klatt, D. H. (**1986**). "Detailed Spectral Analysis of a Female Voice," J. Acoust. Soc. Am. Suppl. 1 **80**, S79.

Koike, Y., and Hirano, M. (**1973**). "Glottal-Area Time Function and Subglottal Pressure Variation," J. Acoust. Soc. Am. **54**, 1618–1627.

Lofqvist, A., Carlborg, B., and Kitzing, P. (**1982**). "Initial Validation of an Indirect Measure of Subglottal Pressure during Vowels," J. Acoust. Soc. Am. **72**, 633–634.

Monsen, R. B., and Engebretson, A. M. (**1977**). "Study of Variations in the Male and Female Glottal Wave," J. Acoust. Soc. Am. **62**, 981–993.

Netsell, R. (**1973**). "*Speech Physiology*," *Normal Aspects of Speech, Hearing and Language*, edited by F. D. Minifie, T. J. Hixon, and F. Williams (Prentice-Hall, Englewood Cliffs, NJ), pp. 211–234.

Rothenberg, M. R. (**1973**). "A New Inverse-Filtering Technique for Deriving the Glottal Air Flow Waveform during Voicing," J. Acoust. Soc. Am. **53**, 1632–1645.

Rothenberg, M. R. (**1977**). "Measurements of Airflow in Speech," J. Speech Hear. Res. **20**, 155–176.

Rothenberg, M. R. (**1981**). "Acoustic Interaction between the Glottal Source and the Vocal Tract," in Stevens and Hirano (1981), pp. 305–328.

Rothenberg, M. R. (**1985**). "Source-Tract Acoustic Interaction in Breathy Voice," in Titze and Scherer (1985), pp. 465–481.

Rothenberg, M. R. (**1986**). "Comment in Chairman's Summary of Session 3," J. Phon. **14**, 443–444.

Schutte, H. K. (**1981**). *The Efficiency of Voice Production* (Kemper, Groningen), 192 pp.

Shaughnessy, L., Lotz, W. K., and Netsell, R. (**1981**). "Laryngeal Resistance for Syllable Series and Word Productions," ASHA **23**, 745.

Smitheran, J. R., and Hixon, T. J. (**1981**). "A Clinical Method for Estimating Laryngeal Airway Resistance during Vowel Production," J. Speech Hear. Res. **46**, 138–146.

Stetson, R. H. (**1951**). *Motor Phonetics* (North-Holland, Amsterdam), 2nd ed.

Stevens, K. N. (**1987**). Personal communication.

Stevens, K. N., and Hirano, M. (Eds.) (**1981**). *Vocal Fold Physiology, Proceedings of the Vocal Fold Physiology Conference*, 1980, Kurume, Japan (Univ. Tokyo, Tokyo, Japan), 421 pp.

Stevens, K. N. (**1977**). "Physics of Laryngeal Behavior and Larynx Modes," Phonetica **34**, 264–279.

Stevens, K. N. (**1981**). "Vibration Modes and Model Parameters," in Stevens and Hirano (1981), pp. 291–304.

Sundberg, J., and Gauffin, J. (**1979**). "Waveform and Spectrum of the Glottal Voice Source," in *Frontiers of Speech Communication*, edited by S. Ohman and B. Lindblom (Academic, London), pp. 301–322.

Tabachnick, B. G., and Fidell, L. S. (**1983**). *Using Multivariate Statistics* (Harper and Row, San Francisco, CA), 507 pp.

Tanaka, S., and Gould, W. J. (**1983**). "Relationships between Vocal Intensity and Noninvasively Obtained Aerodynamic Parameters in Normal Subjects," J. Acoust. Soc. Am. **73**, 1316–1321.

Timcke, R., von Leden, H., and Moore, P. (**1958**). "Laryngeal Vibrations: Measurements of the Glottic Wave. Part 1. The Normal Vibratory Cycle," AMA Arch. Otolaryngol. **69**, 438–444.

Titze, I. R. (**1981**). "Biomechanics and Disturbed-Mass Models of Vocal Fold Vibration," in Stevens and Hirano (1981), pp. 245–270.

Titze, I. R., and Scherer, R. C. (Eds.) (**1985**). *Vocal Fold Physiology. Biomechanics, Acoustics and Phonatory Control, Proceedings of the International Conference on Physiology and Biophysics of the Voice*, Iowa City, IA, 1983 (Denver Center of the Performing Arts, Denver, CO), 527 pp.

Winer, B. J. (**1971**). *Statistical Principles in Experimental Design* (McGraw-Hill, New York), 905 pp.