Effects of bandwidth on glottal airflow waveforms estimated by inverse filtering

Paavo Alku
Helsinki University of Technology, Acoustics Laboratory, Otakaari 5 A, FIN-02150 Espoo, Finland

Erkki Vilkman
University of Oulu, Department of Otolaryngology and Phoniatrics, FIN-90220 Oulu, Finland

(Received 22 March 1994; accepted for publication 17 March 1995)

The aim of this research was to examine the effect of the bandwidth on parameters that are used to quantify glottal airflow waveforms that have been obtained by inverse filtering the acoustic signals. Three bandwidths (1, 2, and 4 kHz) were compared using /a/ vowels of different phonation types. Glottal waveforms were characterized using three ratios computed from the flow (open quotient, speed quotient, and closing quotient) and two parameters extracted from the differentiated flow (the amplitude of the negative peak and the return time from the negative peak to zero level). The results show that especially in pressed phonation the bandwidth should be at least 4 kHz. Distortion caused by an insufficient frequency range was most severe in the parameters extracted from the flow derivatives. Among the parameters computed from the flow signals, the closing quotient was affected most when the bandwidth was reduced. © *1995 Acoustical Society of America.*

PACS numbers: 43.70.Aj

INTRODUCTION

Inverse filtering is a widely applied technique in the estimation of the glottal airflow waveform. Since the presentation of the idea by Miller (1959) many different versions of this technique have been developed. Rothenberg (1973) presented an approach where the effect of formants is canceled by inverse filtering the oral flow, which is recorded with a circumferentially vented wire screen mask. The glottal airflow waveform estimated by Rothenberg's method can be calibrated in the amplitude domain and the DC-flow is also obtained. Many inverse filtering methods have also been developed that are based on the application of the acoustic sound pressure wave which has been recorded with a microphone in a free field (Hunt *et al.*, 1978; Wong *et al.*, 1979; Matausek and Batalov, 1980; Alku, 1992). The resulting glottal airflow waveform is obtained on an arbitrary amplitude scale.

Glottal airflow waveforms obtained by inverse filtering are usually parametrized in order to get a quantitative presentation to describe voice production. Quantification of the estimated glottal airflow can be done, for example, by applying time-based parameters (Holmberg *et al.*, 1988; Holmberg *et al.*, 1989; Hillman *et al.*, 1989; Hillman *et al.*, 1990; Dromey *et al.*, 1992). In the glottal waveform depicted in Fig. 1(a) these parameters are defined as follows: OQ (open quotient)$=(t_{01}+t_{02})/T$, SQ (speed quotient)$=t_{01}/t_{02}$, and CQ (closing quotient)$=t_{02}/T$, where T denotes the length of the fundamental period. Amplitude domain values can be used in characterizing the glottal airflow waveform if inverse filtering is made on air flow signals (Karlsson, 1986; Hertegård *et al.*, 1992; Higgins and Saxman, 1993; Sundberg *et al.*, 1993). A widely used approach in parametric description of the glottal source is to fit certain mathematical functions to the waveform obtained by inverse filtering (Karlsson, 1990; Carlson *et al.*, 1991; Strik and Boves, 1992). One of the most widely used voice source models is the Liljencrants–Fant model (LF model), which is based on quantification of the first derivative of the glottal flow with four parameters (Fant *et al.*, 1985). According to Fant (1993), the most important parameter to characterize the glottal source is the amplitude of the negative peak of the flow derivative [$A_{\min}$ in Fig. 1(b)] since it sets formant amplitude levels. The importance of this parameter has been analyzed experimentally by Gauffin and Sundberg (1989) who showed that there is a correlation between $A_{\min}$ and the sound-pressure level (SPL). Characteristics of the glottal source have been quantified using the time that is required for the differentiated glottal airflow to return from the negative peak to zero level [$t_{\rm ret}$ in Fig. 1(b)] (Sundberg *et al.*, 1993; Price, 1989). In the LF model a slightly different approach is used, namely projection of the returning phase of the derivative at the time instant of negative peak on the zero level (Fant *et al.*, 1985). According to Fant *et al.* (1985) the return phase of the glottal flow effects on the spectral tilt of the glottal source. Even a very small increase in the return phase causes a significant spectrum roll-off in addition to the standard-12 dB/oct glottal flow spectrum.

The starting point for this study was the finding that there is a large variation in the bandwidth values in studies that have been done in the area of inverse filtering. A bandwidth of less than 1 kHz has been used in the following studies: Holmberg *et al.* (1988), Holmberg *et al.* (1989), Hillman *et al.* (1989), Hillman *et al.* (1990), Perkell *et al.* (1991), and Higgins and Saxman (1993). A bandwidth larger than 1 kHz but smaller than 4 kHz has been used, for example, by Hertegård *et al.* (1990), Hertegård *et al.* (1992), and Dromey *et al.* (1992). A bandwidth larger than or equal to 4 kHz has been used, for example, by Veeneman and BeMent (1985), Krishnamurthy and Childers (1986), Javkin *et al.* (1987), Carlson *et al.* (1991), Childers and Lee (1991),

 0001-4966/95/98(2)/763/5/$6.00

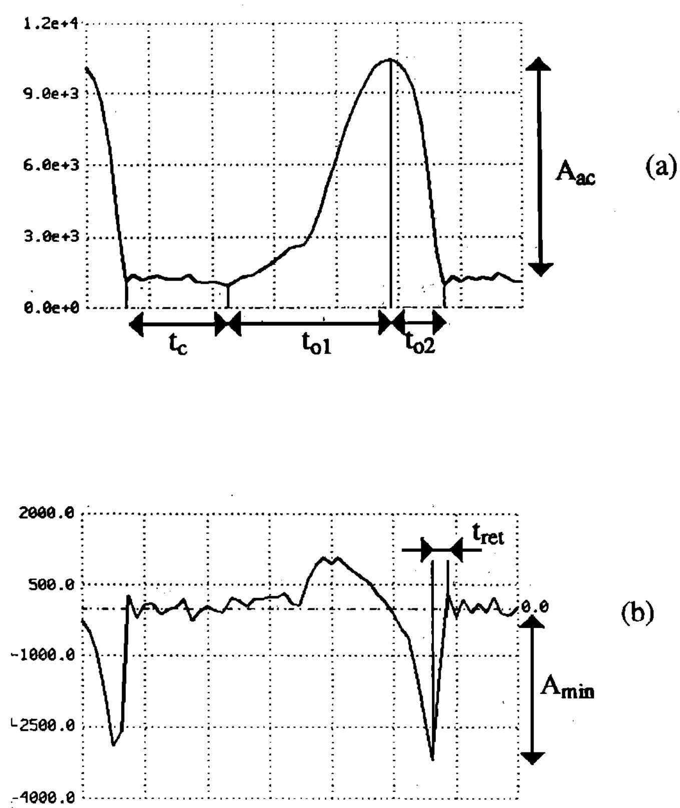

FIG. 1. (a) Glottal airflow waveform t_c=closed phase, t_{01}=opening phase, t_{02}=closing phase, A_{ac}=amplitude of AC flow. (b) Differentiated glottal airflow waveform, A_{min}=amplitude of the negative peak, t_{ret}=return time.

Howell and Williams (1992), and Alku (1992). Among the studies listed above those that belong to the first and second category (i.e., bandwidth smaller than 4 kHz) are all based on inverse filtering of the oral flow with the Rothenberg's mask. However, the studies of the last category (i.e., bandwidth larger than 4 kHz) apply the speech pressure waveform in inverse filtering.

The goal of this study was to examine the effect of the bandwidth of inverse filtering on parametrization of the glottal waveform. Inverse filtering was computed from the acoustic speech pressure waveforms. Both the glottal flow and its first derivative were used in order to quantify the voice source. We were interested in exploring the effect of the bandwidth on time-based parameters (OQ, SQ, and CQ). From the derivative of the glottal airflow waveforms we analyzed the effect of the bandwidth on A_{min} because of its importance in voice production (Fant, 1993; Gauffin and Sundberg, 1989). Extraction of the return time parameter that is used in the LF model (i.e., projection of the flow derivative waveform on the zero level) is sometimes difficult in natural speech material (Fant, 1993). Therefore, we decided to analyze the effect of the bandwidth on the time-domain properties of the differentiated flow by measuring t_{ret}.

I. MATERIAL AND METHODS

A. Speech material

The speech material of this study consisted of vowels that were produced by four female and four male speakers. All the subjects were adult native speakers of Finnish and none of them had a history of hearing or voice disorders. The voices were also perceptually within normal limits as judged by a phoniatrician. The speakers were asked to produce sustained /a/ vowel using breathy, normal, and pressed phonation types. The length of the pronunciation was 2 s. Subjects were allowed to use their natural fundamental frequency and loudness during the recording. By mimicking an experimenter the speakers were taught to produce the vowel with three different phonation types. During the recording the voice quality was assessed by an experimenter who asked the subject to repeat the task until the phonation was satisfactory. The recording of the acoustic speech pressure waveforms was made in an anechoic chamber using a condenser microphone (Brüel & Kjær 4133) which was held 40 cm from the lips of the speaker. Sound-pressure levels were measured during the recordings (Brüel & Kjær 2235). Speech data were saved onto a digital tape using a DAT-recorder (TEAC RD-200T).

Speech material was collected from voices of different phonation types because it is known that characteristics of the glottal airflow waveform vary in different modes of phonation (Hunt, 1987; Price, 1989; Childers and Lee, 1991). Breathy phonation is usually characterized by a flow that is smooth with a short closed phase of the glottal cycle. When phonation is changed to normal and to pressed, the duration of the closed phase is increased while the duration of the closing phase normally decreases. This implies that the frequency contents of the glottal airflow vary from one phonation type to another (Childers and Lee, 1991). In order to present rapid changes in the glottal flow of pressed phonation a wider frequency range is required than in the case of breathy phonation.

B. Inverse filtering

Estimation of the glottal volume velocity waveforms was performed in our study by applying an inverse filtering technique that uses as its only input the acoustic speech pressure waveform that has been recorded in a free field (i.e., *no* flow mask was used). All the recorded speech signals were first high-pass filtered with a linear phase FIR-filter in order to remove any low-frequency ambient room noise. The cutoff frequency of the filter was adjusted to be 80% of the fundamental frequency ($F0$) of the speech signal. Inverse filtering was performed with a recently developed method that estimates the glottal source automatically from the speech pressure waveform (Alku, 1992). The algorithm first computes a preliminary estimate for the contribution of the glottal source to the speech spectrum. After canceling the estimated effect of the voice source a model for the vocal tract is computed. The major innovation in our inverse filtering algorithm concerns the all-pole modeling technique that is used in estimating the vocal tract transfer function. Instead of applying conventional linear predictive coding (LPC) (Makhoul, 1975), our method takes advantage of an iterative approach where poles of the vocal tract model are determined using the Itakura–Saito distortion method (Markel and Gray, 1978). This new all-pole modeling technique yields more accurate estimates for the formants in comparison to conventional LPC analysis. Hence, the resulting glottal waveforms are less distorted by formant ripples.

The inverse filtering analysis of this study was computed using the order of the vocal tract all-pole filter which was equal to 8, 10, and 12 in the analysis of breathy, normal, and pressed phonation, respectively. The block length of the analysis was 32 ms. The position of the analysis window was varied in order to find a setting that yielded a waveform with minimum formant ripple. The original bandwidth of the analysis was 4 kHz. The obtained glottal airflow waveforms were then low-pass filtered in order to analyze pulseforms whose frequency range was 1 and 2 kHz. Low-pass filtering was computed using linear phase FIR-filters whose attenuation in the stop band was more than 70 dB. Finally, all of the estimated glottal airflow waveforms were differentiated to analyze the effect of low-pass filtering on flow derivatives.

C. Data extraction and processing

The waveforms of three different frequency ranges were analyzed on the computer screen in order to mark points in time used for the computation of parameters. All the waveforms were analyzed by one of the authors and the data was averaged over four consecutive glottal cycles. An example is given in Fig. 2 that shows the glottal airflow and its first derivative extracted during one glottal cycle. Waveforms in Fig. 2(a) and (b) were obtained using the bandwidth of 4 kHz whereas curves of Fig. 2(c) and (d) have a frequency range of 1 kHz. The rules that were used in the extraction of time and amplitude values were as follows:

(1) The time of maximal glottal opening [t_m in curves 2(a) and (c)] was first determined as the time when flow reaches its maximum value during one glottal cycle.

(2) The time of glottal closure [t_c in curves 2(a) and (c)] was then determined as the last sample during the descending portion of the flow after t_m. Extraction of t_c was checked from the derivative because t_c should be equal to the time instant when derivative returns to zero level [t_{dz} in curves 2(b) and (d)].

(3) The onset of glottal opening [t_0 in curves 2(a) and (c)] was determined as the time after glottal closure when the flow showed a clear increase. In the case of a gradual opening, t_0 was determined by searching for the first sample whose amplitude was at least 5% of the difference between the amplitude at t_m and the amplitude at t_c.

(4) The time instant of the negative peak of the differentiated flow [t_{dm} in curves 2(b) and (d)] was then determined as the minimum of the derivative between time instants of maximal glottal flow (t_m) and glottal closure (t_c).

The marked points in time were used in order to compute the values of the time-based parameters. From the flow signals, we also computed the difference between the maximum and minimum flow [A_{ac} in Fig. 1(a)]. Hence, each of the estimated glottal flow waveforms were characterized by four parameters extracted from the flow signal (OQ, SQ, CQ, and A_{ac}) and two parameters computed from the first derivative of the flow ($A_{\min}$ and t_{ret}). Mean values and standard deviations of these six parameters were then computed separately for male and female voices. Relative differences in the values of all six parameters were determined between the low-pass filtered signals and the original signals with a bandwidth of 4 kHz. (As an example, relative difference in the value of open quotient between a signal whose bandwidth was 1 kHz and a signal with a frequency range of 4 kHz was defined as follows: $100\% \cdot (\mathrm{OQ}_{1\ \mathrm{kHz}} - \mathrm{OQ}_{4\ \mathrm{kHz}})/\mathrm{OQ}_{4\ \mathrm{kHz}}$). Finally, mean values were computed over the relative differences of female and male voices.

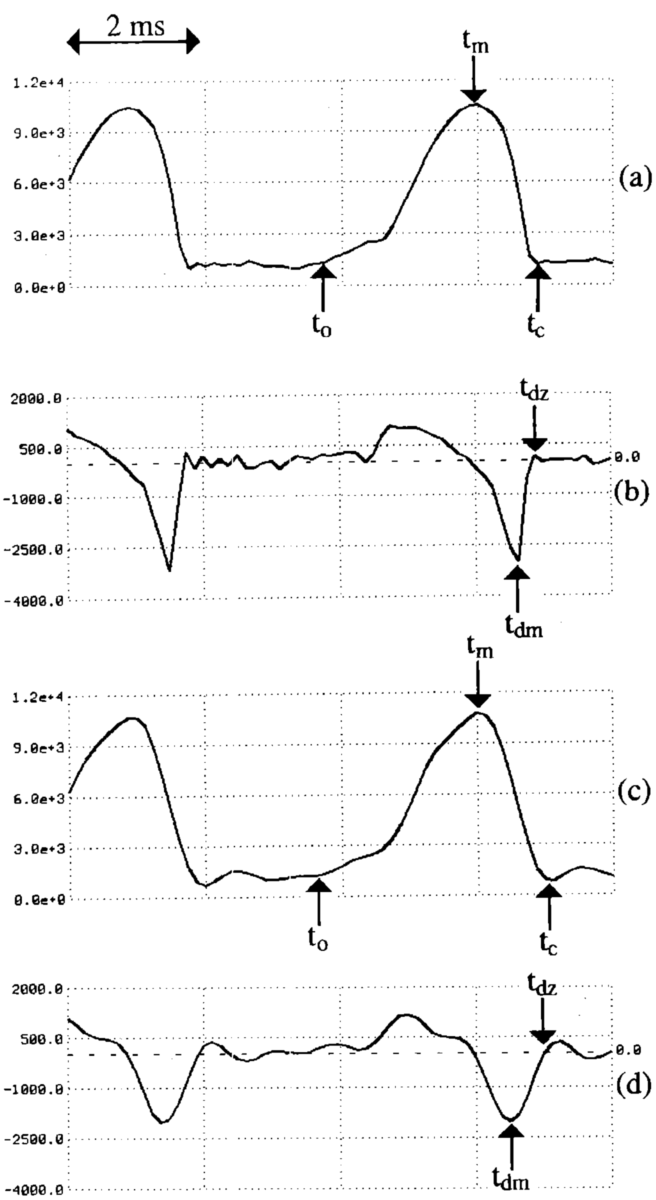

FIG. 2. Marked time instants in glottal airflow waveforms [curves (a) and (c)] and in their derivatives [curves (b) and (d)]. The bandwidth of curves (a) and (b) is 4 kHz, while curves (c) and (d) are low-pass filtered at 1 kHz. t_0: glottal opening, t_m: time instant of maximum flow, t_c: glottal closure, t_{dm}: time instant of negative peak in differentiated flow, t_{dz}: glottal closure, i.e., time instant when derivative returns to zero.

II. RESULTS

Mean values and standard deviations of all the parameters computed from the glottal airflow waveforms with a bandwidth of 4 kHz are given for male and female voices in Tables I and II, respectively. $F0$ values computed from the

TABLE I. Means (m) and standard deviations (s.d.) of parameters extracted from the glottal flow and its derivative using the bandwidth of 4 kHz, four male speakers. (Units are as follows: $F0$ in Hz, SPL in dB, t_{ret} in milliseconds, A_{min} and A_{ac} arbitrary.)

Phonation		$F0$	SPL	OQ	SQ	CQ	t_{ret}	A_{min}	A_{ac}
breathy	m	104	55	0.95	1.17	0.44	2.22	440	9270
	s.d.	12	5	0.03	0.25	0.05	1.07	252	5452
normal	m	100	59	0.88	2.13	0.28	0.63	880	9352
	s.d.	8	7	0.07	0.22	0.03	0.40	675	5471
pressed	m	101	65	0.60	2.15	0.19	0.53	1320	7276
	s.d.	7	9	0.19	0.53	0.04	0.36	971	5249

TABLE III. Mean of relative differences in parameters extracted from the glottal flow and its derivative between low-pass filtered waveforms and original signals of the 4-kHz bandwidth, four male speakers.

Phonation	bandwidth (kHz)	OQ (%)	SQ (%)	CQ (%)	t_{ret} (%)	A_{min} (%)	A_{ac} (%)
breathy	2	0.5	−1.8	1.2	−3.5	0.0	0.0
breathy	1	0.8	−2.9	1.9	−5.0	−4.8	−0.3
normal	2	2.0	−3.4	6.4	34.0	1.3	0.0
normal	1	4.1	−4.1	7.5	140.0	−6.3	0.3
pressed	2	2.6	1.8	2.5	29.3	2.0	−0.5
pressed	1	7.0	1.6	8.0	173.0	−21.8	−1.3

glottal waveforms and SPL values given by the measuring equipment are also given in these tables. Table III shows relative differences in parameter values of male voices when low-pass filtered signals were compared to glottal airflow waveforms whose bandwidth was 4 kHz. The effect of the reduced bandwidth on the parameters computed from the differentiated glottal flow was larger than on the time based parameters computed directly from the flow. This is natural since differentiation corresponds to high-pass filtering. From Table III it can be observed that especially values of t_{ret} were distorted when the bandwidth of the glottal excitations were decreased. From Table III we can also observe an important result that the amplitude of the minimum peak of the flow derivative reduced when the bandwidth was decreased. It should be observed that this happened even though A_{ac} values determined from the flow were almost constant when the bandwidth was decreased. Hence, changes in the values of A_{min} are really due to reducing the bandwidth and not, for example, by accidentally scaling the amplitude levels of the flow in low-pass filtering. As can be seen from Table III the effect of low-pass filtering on A_{min} was largest in pressed phonation where the amplitude of the negative peak of the flow derivative decreased by 21.8% when the bandwidth was reduced from 4 to 1 kHz. The results related to the time based parameters show that the effect of bandwidth is negligible in breathy phonation whereas in normal and pressed phonation the values of quotients change as a result of the reduction in bandwidth. The values of OQ and CQ increased, whereas the value of SQ decreased or was unaffected. However, the average change in each of the quotients was below 10% among the waveforms that were obtained for male speakers.

Table IV shows the relative differences in six parameters due to reducing the bandwidth for the female voices. The parameters computed from the differentiated glottal flows of the female voices were clearly affected when the bandwidth was reduced. Reducing the bandwidth from 4 to 1 kHz caused an increase of more than 50% in the value t_{ret} when phonation was normal or pressed. At the same time, the amplitude of the negative peak showed a decrease which was largest (32.5%) in pressed phonation. A_{ac} values were affected only slightly (maximally by 2.8%) when the bandwidth was decreased. An example of the effects of reducing the bandwidth is shown by Fig. 2. By comparing flow derivatives shown in Fig. 2(b) and (d) it can be observed how removing the high frequencies increases the value of t_{ret} and at the same time decreases the value of A_{min}. For the female voices, the effect of the bandwidth was more noticeable than for male voices when values of OQ, SQ, and CQ were analyzed. The effect of a narrower frequency range was largest on CQ of pressed phonation where the value was increased by 18.2% when the bandwidth was reduced from 4 to 1 kHz.

III. SUMMARY

The experiments that were performed using both male and female voices showed that decreasing the bandwidth from 4 to 1 kHz affected parameters that were extracted from the glottal flow and its first derivative. The effect of the bandwidth on the time-based parameters was greater in female than in male voices. In both female and male voices, the value of CQ was distorted most among the time-based parameters. The reason for this is that the glottal airflow waveform has the most abrupt changes during the closing phase. High frequencies are required in order to present these abrupt changes. If these frequency components are removed

TABLE II. Means (m) and standard deviations (s.d.) of parameters extracted from the glottal flow and its derivative using the bandwidth of 4 kHz, four female speakers. (Units are as follows: $F0$ in Hz, SPL in dB, t_{ret} in milliseconds, A_{min} and A_{ac} arbitrary.)

Phonation		$F0$	SPL	OQ	SQ	CQ	t_{ret}	A_{min}	A_{ac}
breathy	m	196	60	0.97	1.31	0.43	0.75	1013	8552
	s.d.	11	4	0.02	0.34	0.08	0.23	527	3814
normal	m	195	64	0.84	2.01	0.29	0.56	1149	6123
	s.d.	22	4	0.08	0.53	0.08	0.23	431	1533
pressed	m	201	74	0.77	2.57	0.24	0.56	2150	7728
	s.d.	18	5	0.09	0.85	0.08	0.32	1128	2476

TABLE IV. Mean of relative differences in parameters extracted from the glottal flow and its derivative between low-pass filtered waveforms and original signals of the 4-kHz bandwidth, four female speakers.

Phonation	bandwidth (kHz)	OQ (%)	SQ (%)	CQ (%)	t_{ret} (%)	A_{min} (%)	A_{ac} (%)
breathy	2	0.5	−2.4	2.7	5.8	−3.0	−0.5
breathy	1	1.8	1.4	2.6	14.3	−6.8	0.0
normal	2	1.6	−2.5	3.2	28.5	−6.3	−0.8
normal	1	6.6	−3.9	9.8	72.0	−18.3	0.0
pressed	2	1.1	−7.7	7.7	15.8	−14.8	0.5
pressed	1	8.5	−8.7	18.2	50.0	−32.5	2.8

the shape of the waveform will become rounded at the end of the closing phase. Hence, the value of CQ will become larger due to low-pass filtering.

It was found that the effect of the frequency range was largest on the parameters computed from the differentiated glottal flow. The parameter whose value was most severely distorted by reducing the bandwidth was t_{ret}. Decreasing the bandwidth from 4 to 1 kHz distorted (increased) the average value of t_{ret} by more than 10% for all the analyzed signals except the breathy male voices. The values of A_{min} were also shown to be greatly dependent on the bandwidth. When the bandwidth was reduced the values of A_{min} decreased especially for pressed male voices and for female voices of normal and pressed phonation types.

As a result of this study the authors recommend to apply a bandwidth which is at least 4 kHz in the analysis of glottal airflow waveforms that have been obtained by inverse filtering. The importance of the bandwidth is emphasized in the case when speech material consists of samples representing pressed phonation type especially when fundamental frequency is high. If the function of the speech production mechanism is analyzed with the help of differentiated glottal flow it is most important that the analysis is performed with a frequency range that is sufficient.

Alku, P. (**1992**). "Inverse filtering of the glottal waveform using the Itakura–Saito distortion measure," Proc. Int. Conf. on Spoken Language Processing, Banff, Canada, 12–16 Oct., pp. 847–850.

Carlson, R., Granström B., and Karlsson, I. (**1991**). "Experiments with voice modelling in speech synthesis," Speech Commun. **10** (5-6), 481–489.

Childers, D. G., and Lee, C. K. (**1991**). "Vocal quality factors Analysis, synthesis, and perception," J. Acoust. Soc. Am. **90**, 2394–2410.

Dromey, C., Stathopoulos, E. T., and Sapienza, C. M. (**1992**). "Glottal airflow and electroglottographic measures of vocal function at multiple intensities," J. Voice **6** (1), 44–54.

Fant, G., Liljencrants, J., and Lin, Q. (**1985**). "A four-parameter model of glottal flow," Speech Transmission Laboratory Quarterly Progress and Status Reports, No. 4, Royal Institute of Technology, Stockholm, Sweden, pp. 1–13.

Fant, G. (**1993**). "Some problems in voice source analysis," Speech Commun. **13**, 7–22.

Gauffin, J., and Sundberg, J. (**1989**). "Spectral correlates of glottal voice source waveform characteristics," J. Speech Hear. Res. **32**, 556–565.

Hertegård, S., Gauffin, J., and Sundberg, J. (**1990**). "Open and covered singing as studied by means of fiberoptics, inverse filtering, and spectral analysis," J. Voice **4** (3), 220–230.

Hertegård, S., Gauffin, J., and Karlsson, I. (**1992**). "Physiological correlates of the inverse filtered flow waveform," J. Voice **6** (3), 224–234.

Higgins, M. B., and Saxman, J. H. (**1993**). "Inverse-filtered air flow and EGG measures for sustained vowels and syllables," J. Voice **7** (1), 47–53.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., and Vaughan, C. (**1989**). "Objective assessment of vocal hyperfunction: An experimental framework and initial results," J. Speech Hear. Res. **32**, 373–392.

Hillman, R. E., Holmberg, E. B., Perkell, J. S., Walsh, M., and Vaughan, C. (**1990**). "Phonatory function associated with hyperfunctionally related vocal fold lesions," J. Voice **4** (1), 52–63.

Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (**1988**). "Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice," J. Acoust. Soc. Am. **84** (2), 511–529.

Holmberg, E. B., Hillman, R. E., and Perkell, J. S. (**1989**). "Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal, and high pitch," J. Voice **3** (4), 294–305.

Howell, P., and Williams, M. (**1992**). "Acoustic analysis and perception of vowels in children's and teenagers' stuttered speech," J. Acoust. Soc. Am. **91** (3), 1697–1706.

Hunt, M. J., Bridle, J. S., and Holmes, J. N. (**1978**). "Interactive digital inverse filtering and its relation to linear prediction methods," Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, Tulsa, USA, April 10–12, pp. 15–18.

Hunt, M. (**1987**). "Studies of glottal excitation using inverse filtering and an electroglottograph," Proc. of the Eleventh International Congress of Phonetic Sciences, August 1–7, Tallinn, Estonia, Vol. 3, pp. 23–26.

Javkin, H. R., Antonanzas-Barroso, N., and Maddieson, I. (**1987**). "Digital inverse filtering for linguistic research," J. Speech Hear. Res. **30**, 122–129.

Karlsson, I. (**1986**). "Glottal wave forms for normal female speakers," J. Phon. **14**, 415–419.

Karlsson, I. (**1990**). "Voice source dynamics of female speakers," Proc. Int. Conf. on Spoken Language Processing, Kobe, Japan, 19–22 Nov., Vol. 1, pp. 69–72.

Krishnamurthy, A. K., and Childers, D. G. (**1986**). "Two-channel speech analysis," IEEE Trans. Acoust. Speech Signal Process. **34** (4), 730–743.

Makhoul, J. (**1975**). "Linear prediction: A tutorial review," Proc. IEEE **63** (4), 561–580.

Markel, J. D., and Gray, A. H., Jr. (**1976**). *Linear Prediction of Speech* (Springer-Verlag, New York).

Matausek, M. R., and Batalov, V. S. (**1980**). "A new approach to the determination of the glottal waveform," IEEE Trans. Acoust. Speech Signal Process. **28** (6), 616–622.

Miller, R. L. (**1959**). "Nature of the vocal cord wave," J. Acoust. Soc. Am. **31**, 667–677.

Perkell, J. S., Holmberg, E. B., and Hillman, R. E. (**1991**). "A system for signal processing and data extraction from aerodynamic, acoustic, and electroglottographic signals in the study of voice production," J. Acoust. Soc. Am. **89**, 1777–1781.

Price, P. J. (**1989**). "Male and female voice source characteristics: Inverse filtering results," Speech Commun. **8**, 261–277.

Rothenberg, M. (**1973**). "A new inverse-filtering technique for deriving the glottal air flow waveform during voicing," J. Acoust. Soc. Am. **53**, 1632–1645.

Strik, H., and Boves, L. (**1992**). "On the relation between voice source parameters and prosodic features in connected speech," Speech Commun. **11**, 167–174.

Sundberg, J., Titze, I., and Scherer, R. (**1993**). "Phonatory control in male singing: A study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source," J. Voice **7** (1), 15–29.

Veeneman, D. E., and BeMent, S. (**1985**). "Automatic glottal inverse filtering from speech and electroglottographic signals," IEEE Trans. Acoust. Speech Signal Process. **33** (2), 369–377.

Wong, D. Y., Markel, J. D., and Gray, A. H., Jr. (**1979**). "Least squares glottal inverse filtering from the acoustic speech waveform," IEEE Trans. Acoust. Speech Signal Process. **27** (4), 350–355.