Office-Based System for Voice Analysis

Gerald S. Berke, MD; David G. Hanson, MD; Terrence K. Trapp, MD;
Dennis M. Moore, MD; Bruce R. Gerratt, PhD; Manuel Natividad

• There has been recent growing interest in the analysis of various electronically recorded signals as potential tools for objective assessment of vocal dysfunction. In the past, analysis of such signals required an expensive multitrack FM recorder, mainframe computer system, customized software, and significant time commitment. This report describes an adaptation of commercially available components that allow digital recording of multiple electronic signals, storage of data, and subsequent signal analysis using an inexpensive personal microcomputer system. Commercially available software for manipulation and examination of signals is discussed as adapted for examination of glottographic and acoustic signals. The relatively inexpensive availability of similar computer systems will, hopefully, encourage assessment of the clinical applications of objective techniques of voice quality.

(*Arch Otolaryngol Head Neck Surg* 1989;115:74-77)

There has recently been a growing active interest in objective documentation and assessment of voice quality.[1] In part, this interest has occurred because of important advances in techniques and analyses that can be applied to vocal physiology in a scientific manner. Clinicians have recently explored the usefulness of various acoustic factors as well as glottographic reflections of vocal fold vibration.

For clinical documentation, acoustic voice signals are usually recorded on a tape recorder and then played back to document pretreatment and posttreatment results.[2] A number of more sophisticated and objective techniques have been proposed for evaluation of acoustic signals. Among these are spectral analysis, measures of frequency and amplitude perturbation, signal to noise comparisons, linear predictive coding analyses, inverse filtering, and Fourier analysis.[3-6] Acoustic measures can hopefully be related to the perceived quality of the voice. They are, however, more difficult to correlate with specific abnormalities at the level of voice production in the larynx. Glottographic techniques, measures that reflect the vibratory behavior of the glottis, have been advocated by a number of authors for the study of vocal fold vibratory physiology.[7-10] Glottographic measures can, it appears, be more directly related to specific pathologic vibratory behaviors. More recently, glottographic signal analyses have been applied to patient populations.[11-14] However, the application of acoustic or glottographic analyses has in the past required expensive recording equipment and computer systems that were usually limited to basic research laboratories.

Today, the personal computer (PC) is standard equipment in many clinical departments and offices. Affordable and reliable, and now with greater capacity and speed, current PC performance is often superior to that of many mainframe research computers of a decade ago. There has also been a proliferation of digital hardware and commercial software programs tailored to PC applications.

To make clinical applications of various objective voice measurement techniques more feasible and available to a greater number of clinical investigators, we have explored the adaptation of inexpensive commercially available hardware and software to the tasks required for signal analysis of acoustic and physiologic data.

METHODS

This PC system was adapted to accommodate a basic patient recording system of acoustic, glottographic, aerodynamic, and stroboscopic data that have been previously described.[11] Cross-sectional glottic area is monitored by photoglottography (PGG).[15] The technique involves a photosensor on the skin over the cricothyroid membrane that measures light transilluminated through the glottis from a light source in the mouth. The light source can be provided in a variety of ways, either transnasally or transorally. Most recently we have successfully used an inexpensive halogen flashlight directed into the oral cavity as a satisfactory light source. Simultaneous recordings are made of the recorded acoustic signal: PGG and electroglottography (EGG).[16] Impedance across the neck in the vicinity of the vocal folds is monitored using an electroglottograph (Synchrovoice). The acoustic signal is transduced via a condenser microphone (Sony) suspended 5 cm from the mouth. In addition, pressure, flow, and stroboscopic data may be recorded.

Data are acquired with an IBM-compatible, 16-bit, 640K RAM PC fitted with a 20-megabyte hard disk and an 80287 math coprocessor. For storing patient data files, we use a 1.2-megabyte floppy disk. Conversion of the signal from analog to digital format is accomplished with a commercially available 12-bit A to D board (Lab Master, Scientific Solutions Inc, Solon, Ohio).

Figure 1 demonstrates the components of a data acquisition system that parallels the recording system of the UCLA Voice Research Laboratory, which is based on a DEC PDP 11/73 minicomputer system.[11] For the PC system, a commercially available, menu-driven software system ("C-Speech," designed by Paul Milenkovic, University of Wisconsin, Madison) is used to perform on-line digital recording and analysis of the glottographic and acoustic signals. An example of a three-channel dis-

Accepted for publication Aug 15, 1988.
From the Division of Head and Neck Surgery, UCLA School of Medicine, Los Angeles.
Read before the Southern California Section Meeting of the American College of Surgeons, Newport Beach, Calif, Jan 23, 1988.
Reprint requests to CHS 62-132, 10833 Le Conte Ave, Los Angeles, CA 90024 (Dr Berke).

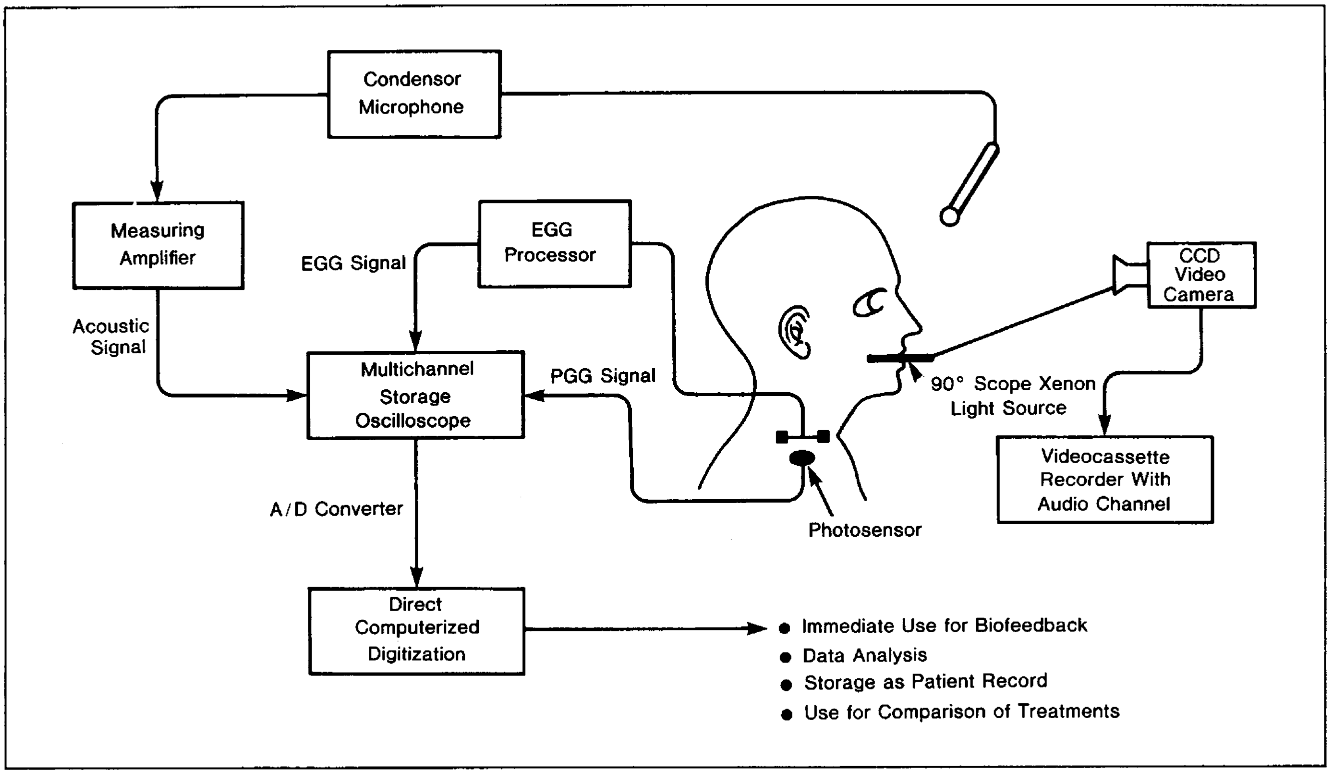

Fig 1.—Diagrammatic representation of voice analysis system that is based on IBM compatible personal computer. EGG indicates electroglottography; PGG, photoglottography; CCD, charge-coupled device; and A/D, analog/digital.

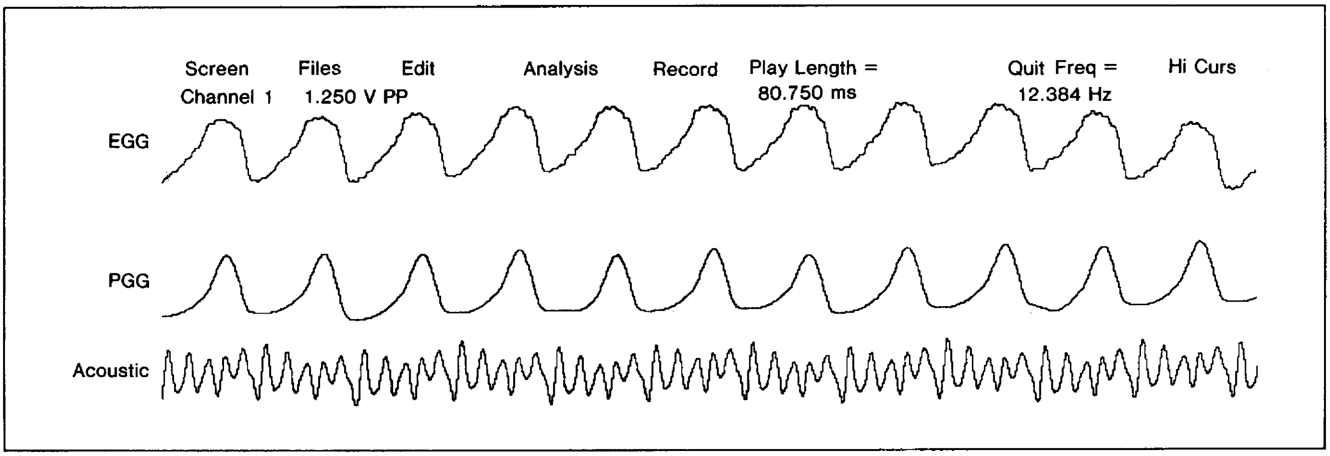

Fig 2.—Computer display of data of 80-ms recording of electroglottography (EGG), photoglottography (PGG), and acoustic signal from normal male speaker. *y*-Axis for EGG represents increasing impedance (less vocal fold contact). *y*-Axis for PGG depicts increasing transillumination (increasing glottal aperture). VPP indicates volts peak-to-peak.

play using the C-Speech program is shown in Fig 1. A multichannel storage oscilloscope is used to monitor the signals generated by the patient during recording. In addition, most software supports on-line preview of the signals prior to record sampling. Most of the current speech analysis applications are menu driven, as shown on top of Fig 2, allowing commands to be initiated by simply choosing from a list. Beyond the simple facility of acquiring and storing data, most software packages allow for waveform editing, time and amplitude measurements, and waveform analysis with the help of a cursor. For instance, the waveform processing menu of C-Speech contains Fourier analysis, LPC analysis, differentiation, integration, low-pass filtering, jitter, and shimmer. In addition, data files can be plotted immediately, if desired, or template modified to fit other software systems.

CLINICAL APPLICATIONS

Figure 2 shows an example of data display for PGG, EGG, and acoustic signals simultaneously recorded from a normal male speaker. The PGG shows increasing voltage, reflecting greater glottic area, in an upward

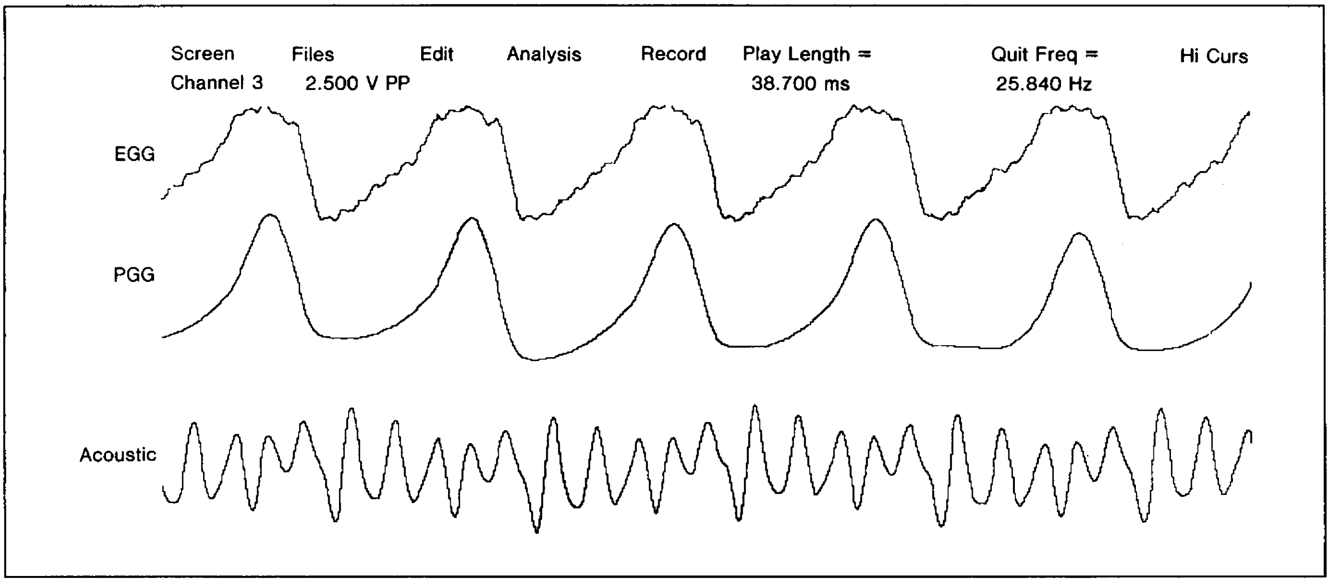

Fig 3.—Computer display of signals can be easily enlarged. This screen represents focus on 38-ms window of data shown in Fig 2. EGG indicates electroglottography; PGG, photoglottography; and VPP, volts peak-to-peak.

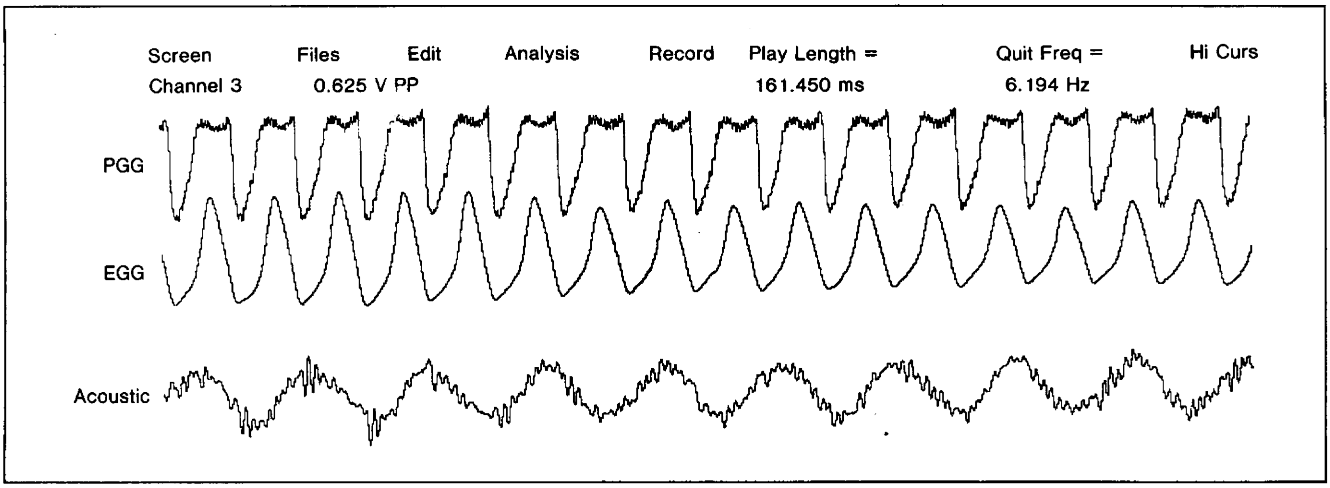

Fig 4.—Electroglottography (EGG), photoglottography (PGG), and acoustic signal data are shown for patient with recurrent nerve paralysis. VPP indicates volts peak-to-peak.

direction. The EGG is oriented with greater impedance, reflecting less vocal fold contact, also in an upward direction. It can be observed that this normal PGG signal shows a regularly recurring pattern of glottic area projection with relatively flat closed periods. The opening and closing slopes of the PGG are comparatively symmetric, and the peak glottic area is centered in relation to the plateau of impedance (representing complete separation of the folds) seen in the EGG signal. Using the commercially available C-Speech program, waveforms from a segment of the data seen in Fig 2 were graphically enlarged for a closer look at the waveform characteristics. The result of this manipulation can be seen in Fig 3. The increasing impedance (decreasing vocal fold contact) on the *y*-axis of the EGG is complemented by the increasing transillumination (increasing glottal aperture) of the PGG during each vocal fold vibratory cycle. These signal characteristics are usual for normal male voices in modal range of phonation.[8] For comparison, Fig 4 shows signals recorded with this system from a patient with a recurrent laryngeal nerve paralysis. Hanson et al[14] preliminarily reported that glottographic waveform configurations dif-

Computer Equipment Expenses

Equipment	Cost, $
Personal computer	1000-3000
Analog/digital converter	500-1500
Software	500-1000
Total Cost	**2000-5500**

ferentiate at statistically significant confidence levels among different lesions to the innervation of the larynx. The PGG signal seen in Fig 4 documents lack of a closed period. The baseline varies because there is always an opening between the folds. The opening phase of the glottic cycle is more rapid than the closing phase. The peak of the PGG signal relative to the plateau phase of the EGG signal is shifted to the left in comparison to the normal signals (Figs 2 and 3). These characteristics of the glottographic waveform pattern are characteristic of flaccid laryngeal paralysis associated with section of the recurrent laryngeal nerve.[12,14] That this pattern results from the physiologic sequelae of recurrent laryngeal nerve section has been confirmed in a canine model of vibratory behavior during experimentally produced phonation.[17]

The acoustic signal demonstrated in Fig 4 documents a lack of normal harmonic structure associated with perception of a breathy, weak voice. The breathy quality of the voice occurs because of a lack of vocal fold closure, which is documented in the glottic area projection signal of the PGG. This lack of a firm period of closure in the glottal cycle, due to the flaccid paralysis of the thyroarytenoid/vocalis muscle, occurred even though the vocal processes were completely approximated in this compensated recurrent paralysis.

COMMENT

In our practice, we find glottographic, airflow, and acoustic measures helpful in differentiating among neuromuscular disorders that can affect laryngeal function. Objective measures of vocal fold function may show distinctive characteristics that can be associated with particular movement disorders such as Parkinson's disease, Shy-Drager syndrome, spastic dystonia, various types of tremor, myoclonus, and ataxias. Neuromuscular dysfunction is more common in the elderly patient, and objective measures are particularly helpful in the diagnosis of voice disorders in that population.[18] In several cases, glottographic measures have, in our experience, correctly predicted the appropriate neurologic diagnosis on voice analysis alone.[18]

There appears to be considerable potential in objective physiologic measures for documentation and diagnosis of vocal dysfunction. However, considerable experience will be necessary to determine the clinical value of encouraging observations that have been made and reported to date. The development and clinical experience that will be necessary to make objective voice measures truly clinically applicable will occur more rapidly if instrumentation can be refined so that it is as simple and economical as possible. It is now possible to have an inexpensive personal microcomputer system that can be used for digital recording and signal analysis.

The Table shows a representative range of the cost involved with setting up a PC-based voice signal recording and analysis system. In comparison to the situation just five years ago, the equipment is relatively inexpensive. In addition to the C-Speech program, there are now a number of good commercial programs on the market that can be used to digitize and analyze voice-related data (Asystant +, Asyst Software Technologies Inc, Rochester, NY; Lab View, National Instruments Inc).

The development and understanding of objective methods for documenting physiologic characteristics of vocal function that can be related to abnormalities is an important goal of current clinically oriented voice research. (The potential value of this area of clinical investigation has recently been highlighted.[19])

The simplicity of adaptation of inexpensive hardware and software computer systems that are presently available in most clinical situations is encouraging for eventual clinical applications of voice analysis techniques. For the present, the availability of this technology will hopefully encourage further clinical investigation of objective measures that may be used to quantify and document voice pathophysiologic conditions.

References

1. Daniloff R: Laryngeal function in phonation and respiration. *ASHA* 1988;30:60-61.
2. Rontal E, Rontal M, Rolnick MI: The use of spectrograms in the evaluation of vocal cord injection. *Laryngoscope* 1975;85:47-56.
3. Horii Y: Fundamental frequency perturbation observed in sustained phonation. *J Speech Hear Res* 1979;22:5-9.
4. Yumoto E, Sasaki Y, Okamura H: Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness. *J Speech Hear Res* 1984;27:2-6.
5. Davis SB: Acoustic characteristics of normal and pathological voices, in Ludlow CL, Hart MS (eds): *Proceedings of the Conference on the Assessment of Vocal Pathology.* American Speech and Hearing Association (ASHA) Reports 11. Rockville, Md, ASHA, 1981, pp 97-115.
6. Berke GS, Gerratt BR, Hanson DG: An acoustic analysis of the effects of surgical therapy on voice quality. *Otolaryngol Head Neck Surg* 1983;92:502-508.
7. Baer T, Lofqvist A, McGarr N: Laryngeal vibrations: A comparison between high speed filming and glottographic techniques. *J Acoust Soc Am* 1983;73:1304-1307.
8. Baer T, Titze I, Yoshioka H: Multiple simultaneous measures of vocal activity, in Bless D, Abbs J (eds): *Vocal Fold Physiology: Contemporary Research and Clinical Issues.* San Diego, College-Hill Press Inc, 1983, pp 237-247.
9. Harden JR: Comparison of glottal area changes as measured from ultra high speed photographs and photoelectric glottographs. *J Speech Hear Res* 1975;18:728-738.
10. Kitzing P, Sonesson B: A photoglottographical study of the female vocal folds during phonation. *Folia Phoniatr* 1974;26:138-149.
11. Hanson DG, Gerratt BR, Ward PH: Glottographic measurement of vocal dysfunction. *Ann Otol Rhinol Laryngol* 1984;92:413-419.
12. Gerratt BR, Hanson DG, Berke GS: Glottographic measures of laryngeal function in individuals with abnormal motor control, in Harris K, Sasaki C, Baer T (eds): *Vocal Fold Physiology: Laryngeal Function in Phonation and Respiration.* San Diego, College-Hill Press Inc, 1987, pp 521-532.
13. Kitzing P: Clinical application of combined electro- and photo-glottography, in *Proceedings of the International Association of Logopedics and Phoniatrics Conference, Copenhagen.* Copenhagen, International Association of Logopedics and Phoniatrics, 1977, vol 1, pp 528-539.
14. Hanson DG, Gerratt BR, Berke GS: Glottographic measures of vocal fold vibration: An examination of laryngeal paralysis. *Laryngoscope* 1988;98:480-485.
15. Sonesson B: A method for studying the vibratory movements of the vocal cords. *J Laryngol Otol* 1959;73:732-737.
16. Fabre P: Un procede electrique percutane d'inscription de l'accolement glottique au cours de la phonation: Glottographic a de haute frequency. *Bull Acad Natl Med* 1957;121:66-69.
17. Trapp TK, Berke GS: Photoelectric measurement of laryngeal paralysis correlated with videostroboscopy. *Laryngoscope* 1988;98:486-492.
18. Hanson DG, Ward PH, Gerratt BR, et al: Geriatric otolaryngology: Diagnosis of neuromuscular voice impairment, in Goldstein J, Kashima H (eds): *Geriatric Otolaryngology.* Toronto, BC Decker Publishers, 1989, pp 71-78.
19. Ward PH, Hanson DG, Gerratt BR, et al: Current and future horizons in laryngeal and voice research. *Ann Otol Rhinol Laryngol,* in press.