M. Tech. Credit Seminar Report, Electronic Systems Group, EE Dept, IIT Bombay submitted Nov 02

Impedance Glottography

Anil Luthra
(Roll No. 02307413)
Supervisor: Prof P C Pandey

Abstract

Impedance Glottography is noninvasive measurement of the time variation of the degree of contact between the vibrating vocal folds during voice production. The aspect of contact being measured is called the vocal fold contact area (VFCA). To measure VFCA, the device used is called impedance glottograph. The device is also called electroglottograph or laryngograph. The principle of operation of device, the waveform obtained, an algorithm for determination of pitch period is discussed. EGG waveform for various voice qualities, drawbacks in EGG and various noises present are described. Concept of multichannel EGG and applications of electroglottography are discussed. Various commercial equipment available are compared alongwith the equipment developed by IIT Bombay.

1. Introduction

Lungs, vocal tract, and larynx are the main organs related with generation of sound. The lungs are the source of airflow. The vocal tract is an accoustic enclosure and acts as acoustic filter shaping the spectrum of the generated sound. The source of most speech occurs in the larynx. There are two folds of the muscular bundle known as vocal chords inside the larynx. These vocal chords obstruct the airflow from the lungs and produce audible vibrations that make the speech. The mechanism of generation of sound is known as phonation. The vibrations consists of three phases namely contact phase, separation phase and open phase [1].

The impedance between the vocal chords is a function of tissue path length. When the vocal chords are open, the tissue path length is maximum and hence the impedance is maximum. When the vocal chords are closed, the tissue path length is minimum and hence the impedance is minimum. Impedance glottograph measures this impedance variation [1].

Impedance glottography is, thus, a non-invasive method of measuring vocal fold contact during voicing without affecting speech production. To measure VFCA, the device used is called impedance glottograph. The device is also called electroglottograph (EGG) or laryngograph. The EGG measures the variation in impedance to a very small electrical current between the electrode pair placed across the neck as the area of vocal fold contact changes during voicing [2].

The method was first developed by Fabre (1957) and influential contributions are credited to Fourcin (1971) and Frokjaer-Jensen (1968). Commercially available devices are produced by Laryngograph Ltd., Glottal enterprises and F-J Electronics [3].

Development work for an electrogllotograph has been carried out at IIT Bombay as part of student projects by Bhagwat (1990), Sriram (1991), Thajudin (1994), Mahajan (1995), Chitnis (1998) and Patil (2000) [1].

2. Principle of Operation of Device

A high frequency electrical current of small voltage and amperage (physiologically safe) passes between two electrodes situated on the surface of the throat at the level of the thyroid cartilage. The basic configuration of the device is depicted in Fig.1. The electrodes are made of copper, silver or gold. They have the form of rings or rectangles covering an area ranging from 3 cm^2 to 9 cm^2. A third electrode is often used as a reference for impedance measurements. It may be designed as a separate electrode or as a ring electrode encircling each of the two other electrodes. The electrodes are usually mounted on a flexible band whose length may be adjusted to hold the electrodes in a steady position and to still allow the subject to comfortably speak and breathe naturally. Sometimes the electrodes are mounted on a small holder which is pressed against the throat by hand. A signal generator supplies the electrodes with an AC sinusoidal current of an alternating frequency usually ranging from 300 kHz to 5 MHz. This frequency is sufficiently high, so that the current capacitancively bypasses the less conductive skin layer without the use of additional conductive paste. The generator may produce constant voltage or constitute a constant current source. The supplied current is different for each particular device, but is not stronger than several milliamperes. The voltage between the electrodes depends on the tissue impedance but the typical value is about 0.5 V . In accordance a power dissipation of only several microwatts occurs at the level the subject's vocal folds [3].

The sensing electrode detects the current as it passes through the skin and the throat. The percentage of amplitude modulation of the received signal reflects the percentage change in tissue impedance in the current's path. The received signal is then demodulated by a signal detector circuit. The typical signal-to-noise ratio of the demodulator is about 40 dB. The demodulated waveform is then A/D converted and stored in a computer [3].

The rapid variation in the conductance is caused mainly by the movement of the vocal folds. As they are separated, the transversal electrical impedance is high due to the fact that air impedance is much higher than tissue impedance. As they approximate and the contact between them increases, the impedance decreases, which results in a relatively higher current flow through the larynx structures. At the maximum contact the decrease is about 1% (up to 2%) of the total larynx conductance. The reason for the current modulation effect is a longer tissue passage for the radio frequency current when the glottis is open, since the total impedance of the tissue is a function of the length of the tissue passage. We generally believe the impedance is least for full fold contact because under this condition there are, in effect, many parallel equally conductive resistance paths between the electrodes. The combined total parallel resistance is less than the resistance of any one path. Therefore, the tissue impedance seen by the EGG device is inversely proportional to lateral contact area of the vocal folds [3].

An integral part of the electroglottographic signal is the varying component generated by the vertical movement of the whole larynx. Therefore, the signal of rapid movements of the vocal folds is superimposed on the signal produced by the slower movements of the other structures. The name *Gx* was proposed for the waveform of larynx movement and the name *Lx* for the vibration component. The *Gx* component originates, for example, can be observed in swallowing, but it may also be caused by the vertical movement of the larynx which is related to the voice quality setting of the raised/lowered larynx. Fluctuations of its type are usually removed from further analysis. The DC offset changes (*Gx*) can be evened out because the effects of the varying larynx height are compensated by the use of additional electrodes or high pass filtering of the registered signal. The latter method may involve signal distortion, especially for low-pitched voices. The distortions may be caused by a too high cut off frequency of the filter (or a too wide filter transition band). This can cause the attenuation of the *Lx* signal component. The non-uniform phase response function of the filter can also change the shape of the filtered waveform. FIR (finite impulse response) filters should be used to prevent nonlinear phase shifts in the signal components.

It should be noted however, that even the unfiltered output of the EGG device is not free of distortion. Particularly the demodulation circuit whose frequency transfer function may influence the frequency response of the EGG device, especially in the low frequency range constitutes an additional source of signal shape deformation [3].

3. Description of the EGG Waveform

When the vocal folds are open and it is ensured that there is no lateral contact between the vocal folds, the impedance is maximal and peak glottal flow occurs (segment (e) in Fig.3). The waveform in this segment is flat, with small fluctuations. Then, the upper margins of the vocal folds make initial contact (segment (f) in Fig.3).

In the next phase of the movement (denoted as (a)) the lower margins come into contact and the vocal folds as a whole continue to close zipper-like. If the vocal folds close very rapidly and along their whole length, the phases (f) and (a) become indistinguishable and consequently the slope of the closure phase ((f)+(a)) becomes steep. The presence of this knee is typical for low to normal voice intensities and the slope of segment (f) is more gradual than the slope of (a).Glottal closure is achieved during phase (a). During the next phase (indicated as (b) in Fig.3) the vocal folds remain in contact and the airflow is blocked.

The opening and the open phase are described analogously. In the process of vocal fold separation the contact between the folds starts to diminish and subsequently the lower margins of the vocal folds begin to separate, initialising the opening. Lower margin separation proceeds gradually during phase (c) (Fig. 3). Then the upper margins also begin to separate, resulting in an acceleration in the growth of impedance (phase (d)) until the full opening is reached [4].

The time derivative of the Lx waveform is usually used in the determination of the periodicity of the signal. It can also be used to identify the distinguishable changes in the slopes during the phases of increasing and decreasing impedance which correspond to those of the simplified model of the EGG (Fig.4).

The positive peak of the derivative serves as an indication the instant of the glottal closure (CGI). This is a robust and reliable marker of the pitch period for various voice qualities and intensities. The negative peak of the EGG derivative is regarded as an indicator of glottal opening. During the opening phase the glottal area increases monotonically until it reaches its maximum. The movements of the vocal folds that are reflected in the EGG have two distinct phases. First, the EGG decreases monotonically, reflecting the decreasing in lateral contact between the vocal folds. During this interval the EGG waveform is convex. Then, as the upper margin separates the waveform of the EGG changes to concave.

EGG, thus, reflects the contact area between the vocal folds. Assuming that the depth of contact does not change during vocal fold vibration, the lateral contact area depends on the length of contact area along the upper margins of the vocal folds. There is , thus, a strong correlation between the EGG and the length of the vocal folds contact [4].

4. Pitch Period Determination

In the EGG signals the pitch period is usually defined as the duration between maximum positive peaks in the differentiated EGG waveform. This peaks are regarded as instants of glottal closure. The marking of pitch period is usually done automatically by computer programs which use a threshold values to detect the peaks of the signal derivative. The threshold is usually defined as a medium value between the minimum and the following maximum peak of the waveform [5].

In the case of pathological voices the location of the closure instant in the EGG signal is no longer as obvious as for normal voices. The adduction phase is often not smooth, and additional peaks in the differentiated EGG waveform are often observed. However, they are not significant. The waveform is irregular.The EGG signal is usually preconditioned to remove unwanted signal components. This is achieved by the use of digital linear phase high or bandpass filtering. The filtered waveform is then subjected to the method of threshold comparison [5].

The algorithm developed by A V Patil for calculating the pitch uses period-by-period analysis for detecting the pitch periods. The pitch periods are estimated from successive zero crossings, which are detected using the Schmitt trigger method as follows [1].

The *Lx* waveform is sampled at a sampling rate of 11025 or 22050 or 44100 sa/s. Now the input *Lx* waveform samples is $x(n)$. Next the mean $x_{mean}(n)$ of maximum $x_{max}(n)$ and the minimum $x_{min}(n)$ amplitude of the acquired sequence $x(n)$ from the glottoal impedance sensor is calculated. This $x_{mean}(n)$ is taken as threshold. One bit quantization of the sequence is done as follows

$$
\begin{aligned}
y(n) &= 1 && \text{if } x(n) > x_{mean}(n) + 0.01\, x_{max}(n) \\
&= -1 && \text{if } x(n) < x_{mean}(n) - 0.01\, x_{min}(n) \\
&= y(n-1) && \text{otherwise}
\end{aligned}
$$

The 1- bit quantized output values is stored in a file, and then processed again. Now calculate the number of samples *N* between the switching from one quantisation level to another for obtaining the instantaneous pitch periods. The instantaneous pitch periods are calculated as

$$F_x = f_s / 2N$$

where f_s is the sampling rate.

Next the F_x plot is obtained by plotting the instantaneous pitch values v/s time. The pitch axis is normalised from minimum pitch as 0% to max pitch as 100%, as shown in Fig.5 [1],[11].

The histogram is plotted as either a single period or a multiple (often, tripple) period histogram. The frequency range of interest is divided into a number of equal bins. Example, the frequency range is taken to be 0 to 1 KHz, and the number of bins is set to 50. Thus, each bin interval is 20 Hz. For a single period histogram, each bin is incremented when a pitch value falls within its frequency range. The content of each bin divided by the total number of pitch values is a measure of the probability density. The probability values are plotted on the log scale or linear scale. In a tripple period histogram, each bin is incremented when three successive pitch values fall within its frequency range [1].

5. The EGG Waveform and Voice Qualities

The dependence of the EGG waveform on voice quality is evident in Fig.6. Several periods of Lx recordings of */a:/* spoken by the same speaker are compared using the same amplitude and time resolution settings. The visual inspection of the waveforms leads to the following conclusions [6]:

Modal voice is characterized by a steep increase of contact in the closing phase. The closing phase is short and peak-to-peak amplitude is high. The maximum contact phase (closed phase) has a parabolic shape and the transition to the rise of the impedance is smooth. The fall of the signal lasts longer than the rise and in the increase of the impedance a knee can be observed.

Whispery voice periods are slightly longer than those of modal voice. The decrease in impedance is very fast. The signal peak is located at the beginning of the maximum contact phase which is much shorter than for modal voice. The increase in impedance is also faster and the knee in the fall is present.

The creaky voice recordings of the EGG contrast in many aspects with the other voice qualities. The shape of the waveform could be described as a triangle with rounded corners and the period duration is twice as long as for modal voice. The instant of glottal closure can be identified clearly at the sharp rise of the waveform. The maximum contact is short compared to the period duration. In the opening phase the signal gradually decrease and no knee can be identified. Due to this fact, the signal duty cycle is not easy to define and it is to be expected that the measure will be biased and imprecise.

Breathy voice is easy to distinguish from the other voices. The peak amplitude is lower, which may depend on the poor contact between the vocal folds or their incomplete closure. Also, the duty ratio of the signal is obviously greater than in other voice qualities. Considering the ratio of rise time and fall time, higher values are observed. The maximum contact phase is extremely short with the maximum located approximately in the middle of the pulse. The contour of the waveform can be described as consisting of short triangular pulses with long, straight, almost zero-valued, flat pauses between them.

Tense voice is produced with a strongly increased muscular tension in the larynx. The EGG waveform typical of tense voice differs in many aspects from the waveforms of other voice qualities. First of all, the shape is rounded, much more sinusoidal and smoothed when compared to other voices. The peak-to-peak amplitude is lower, which can be caused by incomplete glottal closure. The steepness of the increasing signal slope is significantly reduced and the parabolic part starts earlier and lasts longer during the maximum contact phase. The increase in the impedance is gradual with a slight indication of a knee, but it seems that the use of this time instant as an estimate of the opening instant is very unreliable. The maximum contact phase is relatively long. The duty cycle of the signal is comparable to that of modal voice.

6. Drawbacks in EGG

At present, EGGs are being used at many research laboratories, but except for rudimentary applications such as the measurement of vocal period, the technique has not been accepted for general clinical use. There are basically three reasons that electroglottography is not used more commonly. First, there are many subjects for whom the previously available commercial units either yield no output or one that is very noisy and/or very different from vocal fold contact area. More significantly, in cases of noisy or distorted waveforms, the user often has no clear indication that the waveform is not valid. Extremely noisy waveforms, that is, waveforms that are relatively small in amplitude and show a fine-grained irregularity that differs from cycle to cycle, can generally be considered inaccurate as a representation of the VFCA. Conversely, very strong, noise-free waveforms from a normal voice can usually be trusted [7].

Second, to obtain a waveform that represents primarily the VFCA, previous units require accurate placement of the electrodes with respect to the vocal folds. The practice of using extra guard-ring or reference electrodes for reducing noise makes accurate placement more important, since if the glottis is mistakenly placed in the electrical field going to the guard or reference electrode, the closing of the vocal folds can actually act to draw current away from the primary electrode and cause a partial signal inversion, or at least a distortion of the waveform. This can be easily tested by purposely shifting the contact during a held vowel and looking for changes in the waveform.

Third, electroglottography is not used more commonly because the various waveform features of interest to the clinician have not yet been clearly charted. This is undoubtedly due in part to the first two problems, since it would be a waste of effort to document in detail the characteristics of a device that cannot be trusted [7].

7. Noises

Low-frequency artifact

A low-frequency artifact, can result from such factors as electrode movement or the muscularly controlled non vibratory movements of the larynx and the articulators during continuous speech. Since these movements vary little during each glottal cycle, their effects on the EGG waveform are theoretically removable by means of a high-pass filter with a cutoff frequency slightly below the voice fundamental frequency.

It should be noted, however, that there are some elements of these lower-frequency components of the translaryngeal electrical impedance that are of potential interest, most notably the variation in the average value of the waveform during vocal fold adduction or abduction gestures. Thus, some manufacturers make available an output containing these lower-frequency components. The user, though, should keep in mind that these low-frequency outputs will always contain, to some degree, artifacts from other movements in or near the larynx - artifacts that are inherently not separable from the desired components [7].

Random noise

A small amount of broad-band random noise, analogous to the hiss in a weak AM broadcast transmission or the snow in a weak television signal, is always introduced by the electronics in the transmitter and receiver circuitry and by RF energy from the environment that is picked up by the receiver circuit. Random noise can be difficult to identify in an EGG signal from a very hoarse or aperiodic voice, since the noise causes cycle-to-cycle variations in the signal that may be similar in some respects to aperiodicities caused by irregular vocal fold movements. However, in most cases random noise is easy to identify in the EGG waveform by its variability between glottal cycles. In addition, if the EGG unit employs no automatic gain or level control circuitry, the level of random noise in an EGG waveform is easy to measure by merely stopping the voice, as by holding the vocal folds closed against a positive lung pressure, and measuring the resulting broad-band noise, since the random noise components tend not to depend on the presence or absence of vocal fold vibration [7].

Voice-synchronous noise

The most inherently troublesome noise sources are those that are caused by the voice itself and therefore tend to produce EGG components that are synchronous with the desired VFCA signal, that is, that are the same or similar in every glottal cycle. Such components tend to appear as a distortion of the waveform. Voice-synchronous noise can come from any voice-generated physiological vibration that can affect the electrical impedance between the EGG electrodes. Examples can include tissue vibration at the skin-electrode interface, vibration of the pharyngeal walls or tongue, or vibratory movements of the false vocal folds or adjacent structures. Because of the mass of the tissues involved, the tissue vibration causing the synchronous noise will tend to be smoothly varying at the voice fundamental frequency, and, as a result, the voice-synchronous noise components will tend to be much more smoothly varying than the VFCA waveform. For a given voice production, the amplitude of some of the voice-synchronous noise components can be estimated very roughly by moving the electrodes away from the vocal folds to reduce the VFCA component; however, any noise components generated close to the vocal folds cannot be so measured [7].

8. Multichannel EGG

The EGG system uses multielectrode arrays on each side of the neck to provide simultaneous EGG measurements at a number of neck locations. Each electrode pair, consisting of corresponding opposed electrodes, is connected to its respective transmitter and receiver, to constitute a channel. The electrodes in each array can be configured horizontally, vertically, or in a two-dimensional pattern. A horizontal array is used to identify the presence of phase differences in the vibratory pattern along the vocal folds. A vertical array can be used to track the position of the larynx as it moves vertically during speech and is referred to as a tracking multichannel EGG, or TMEGG. The following is a description of a version of the two-channel TMEGG [7].

A major problem in implementing a multichannel EGG is the noise and distortion that can be generated by interference between the RF electrical currents in the various channels. Though there are a number of methods that can be used to reduce such interference, the technique of time-synchronizing the RF signal sources is preferred.

The outputs of the two channels can be shown separately; however, it is also possible to automatically either combine the channel outputs or select between them, so as to produce one optimized signal for display or recording. If desired, amplitude normalization of this final output signal could be added, using some form of automatic gain control circuit.

When a two-electrode-pair TMEGG is used with a multichannel display device, the user would normally position the electrode array for approximately equal amplitudes. Positioning for equal waveform amplitudes would be expected to place the electrode pairs at least roughly equidistant from the level of the glottis.

In an alternate positioning procedure, a relatively simple electronic circuit can be used to compare the output amplitudes and provide the user with a meter or bar graph indication of correct position. Such a meter, labeled larynx height, could be incorporated. A center-zero position on the meter would indicate that the traces A and B were of equal amplitude, and therefore that the vocal folds were approximately centered vertically between the electrode pairs. If both traces are electronically recorded during a given segment of speech, it is possible to compensate for subsequent vertical movements of the larynx by selecting the strongest trace during any specific short period to be analyzed [7].

9. Applications of Electroglottography

Laryngeal function assessment

Electroglottography is useful for the assessment of laryngeal function, as it possible to detect the physical condition of the vocal chords using EGG signal [1].

Speech analysis

Voiced unvoiced classification. Algorithms for classification of a speech segment into voiced, unvoiced, mixed and silent have taken on new levels of accuracy and simplicity when EGG signal is used as an aid in the decision making process. For just a voiced-unvoiced decision, using the EGG signal alone is sufficient since the EGG is ideally zero during unvoiced regions and periodic and nonzero during voiced regions [1].

Pitch estimation. Using the EGG, the pitch estimation becomes simpler and accurate, since usually the pitch value is based on either zero crossings or the distance between the minima in the differentiated EGG [1].

Synthesis application. The EGG aided analysis can play an important role in obtaining parameters for synthesis of high quality speech. This is because the naturalness and intelligibility of synthesized speech are influenced by factors such as accuracy in the vocal tract modelling, voiced unvoiced classifications, pitch detection, and the nature of excitation used [1].

Improving lip reading

Lip reading is one of the alternatives for profoundly deaf person, to understand a conversation. Lip reading, sometimes becomes very difficult for differentiation of certain words, which have identical lip movements. Some of the phoneme pairs which create the discrimination problem are /g/ and /k/, /b/ and /p/, and /v/ and /f/. Words such as 'vat' and 'fat', which have identical lip movements, are very difficult to distinguish. The only difference between 'vat' and 'fat' is that the vocal folds vibrate throughout the /v/ and not during /f/. If the voice pitch is provided externally, normal listeners are lip read a person reading continuous text, at a rate two and a half times faster when they are lip reading without the provision of pitch. The provision of pitch greatly improves lip reading ability in those who have acquired an auditory knowledge of speech before becoming deaf [1].

10. Various Commercial Equipment

Laryngograph Ltd (Laryngograph Processors)

The Laryngograph Processor comprises a portable electro-laryngograph, microphone pre-amplifier, and speech or laryngograph based fundamental frequency (pitch) extractor. It links directly to a recorder (e.g. cassette or DAT) for the capture and playback of Speech (Sp) and Laryngograph (Lx) waveforms, and to the PCLX Interface card which can be fitted to any IBM PC compatible for the display, analysis and printout of the speech and laryngograph waveform and corresponding fundamental frequency. It is supplied with a small high quality microphone to pick up the speech pressure waveform (Sp) and a set of electrodes which, when positioned on the neck either side of the thyroid cartilage, detect the small, relatively rapid, variations in the conductance between the electrodes produced by changes in the nature and area of vocal fold contact during voiced speech.

The unit can be powered by battery or by a mains adapter/battery charger. A fully charged battery will give approximately thirty hours continuous use. When the mains power supply unit is plugged in the Lx Processor will be driven from this unit and the battery will be simultaneously charging. The speech (Sp) and Laryngograph waveforms (Lx) are accessible from several points on the back of the unit. A single socket allows one lead to provide the Sp and Lx waveforms for recording, and for playing back. The aux output is for linking to the PCLX card in the PC [8].

F J Electronics (Portable Electroglottograph type EG 90)

It is designed for use with a multimedia computer using standard multimedia hardware (with SoundBlaster compatible sound card) and software for speech and voice analysis

The electroglottograph is completely isolated from the registration device which normally will be a computer as the transmission passes an opto-coupler. The power supply for the opto-coupler is taken from the computer. The Electroglottograph is powered from a standard 9 volt alkaline battery which will run the system for more than 80 hours before the blue LED diode lights to indicate that the battery voltage is running low. The period of 80 hours corresponds to 4800 recordings of 1 minute each.

The output from the electroglottograph is supplied as a voltage which varies synchronously with the changes in electric impedance between the electrodes. The instrument works only when the impedance between the electrodes typically is between 20 and 250 ohms. Use of EKG jelly or salt water will ensure a good electrical contact with the skin so that the electrical resistance between the electrodes stay in the middle of the registration range. This 250 ohm limit is set by the designer of the EG90 portable electroglottograph to ensure that the operator places the electrodes in a manner which achieves good skin contact. If the electrodes are placed wrongly and the resistance is outside the registration range, the yellow LED diode on the electroglottograph lights. The lower frequency limit is 3.5 Hz, which means that low-frequency disturbances such as head movements are removed without severe disturbances of the waveform. When used together with the preamplifier, simultaneous recordings of calibrated audio signal and EGG can be made [9].

Glottal enterprises electroglottographs

These electroglottographs produce low-noise EGG waveforms and have many features. The patented two-channel models EG2 and EG2-PC indicate proper electrode placement and provides a quantitative indication of vertical laryngeal movement. The new EG2-PC model offers easy connection to a personal computer for use with today's popular speech software packages. Model EG1 is a single channel unit that provides all the advantages of the EG2 electroglottographs, except for the indication of electrode placement or vertical laryngeal movement.

These have option for selecting waveform polarity to measure either vocal fold contact area or inverse vocal fold contact area. LED signal strength indicator can be calibrated to be used to verify that the waveform is strong enough to reliably indicate vocal fold contact variation. Optional calibrator (larynx simulator) allows a verification of all aspects of EGG operation, as well as calibration of the EGG signal amplitude in terms of % variation in neck resistance. Dual rechargeable battery system assures that your measurements will never be interrupted by a low battery [10].

11. Electroglottograph developed at IIT Bombay

Development work for an electrogllotograph has been carried out at IIT Bombay as part of student projects by Bhagwat (1990), Sriram (1991), Thajudin (1994), Mahajan (1995), Chitnis (1998) and Patil (2000) [1].

The circuits developed from Bhagwat to Mahajan had noise problems and and impedance change of less than 1 ohm could not be detected properly [1].

Chitnis in 1998 was successful in developing a glottal impedance sensor, which could detect impedance change of less than 1 ohm. He also developed a stand alone data acquisition and display unit for acquiring impedance variation waveform and analyzing it, and a program for pitch extraction and display [1].

Patil in 2000 under the guidance of Prof. P.C. Pandey developed a more sensitive equipment in which excitation current was reduced from 3 mA to 1 mA and linear response was improved. A glottal impedance simulator was also developed and the glottal impedance sensor was interfaced with PC based sound / multimedia card in place of special signal acquisition hardware [1].

The glottal impedance sensor hardware consists of the module for detecting glottal impedance and a PC with sound card for signal acquisition, processing and display. The impedance detector consists of an oscillator with stabilized output, impedance sensor connected to two electrodes, and AM detector. The pair of circular electrodes is held in contact with the skin on both sides of the thyroid cartilage. Oscillator output voltage (400 KHz) is converted into current (= 1 mA) and injected into electrodes through ac coupling. The impedance across the thyroid cartilage varies in accordance with the contact area between the vocal chords, and consequently the voltage across the electrodes gets amplitude modulated. The AM demodulator circuit consists of band pass filter amplifier, precision full wave rectifier averager, and band pass filter amplifier. It provides the EGG or Lx waveform. One line input of the sound card is used for acquisition of the Lx waveform. The other input may be used for simultaneous acquisition of speech signal [11].

12. Conclusion

Electroglottography has established itself as a valuable method for evaluating laryngeal behaviour. In comparison to other methods of glottography, electroglottography allows for a better representation of the closed and closing phases of vocal fold movement, especially of the vertical contact area. Photoglottography seems particularly advantageous as far as the description of the open phase is concerned. The EGG is superior to all other methods in that it is not uncomfortable to speakers as it is completely non-invasive (it exerts no influence at all on the articulation and production of sounds)

At present, EGGs are being used at many research laboratories, but except for rudimentary applications such as the measurement of vocal period, the technique has not been accepted for general clinical use.

Multichannel techniques can be used to produce an EGG that can verify the fidelity of its own output waveform as an indicator of the time patterning of vocal fold contact and can yield a signal that helps the user properly position the electrodes and/or track vertical movements of the larynx during voiced speech or singing. The use of improved EGG units incorporating such techniques should make possible a higher level of confidence in the results of research into the use of electroglottography in the study of voice production, in voice analysis, and in voice training and vocal pedagogy.

Acknowledgement

I express my gratitude to my guide, Prof. P. C. Pandey, for his invaluable help and guidance which effectively contributed in successful completion of this seminar. He was instrumental in providing technical, moral and administrative support. It is a previlege to work under his guidance.

I am also very thankful to the members of SPI Lab for their cooperation extended.

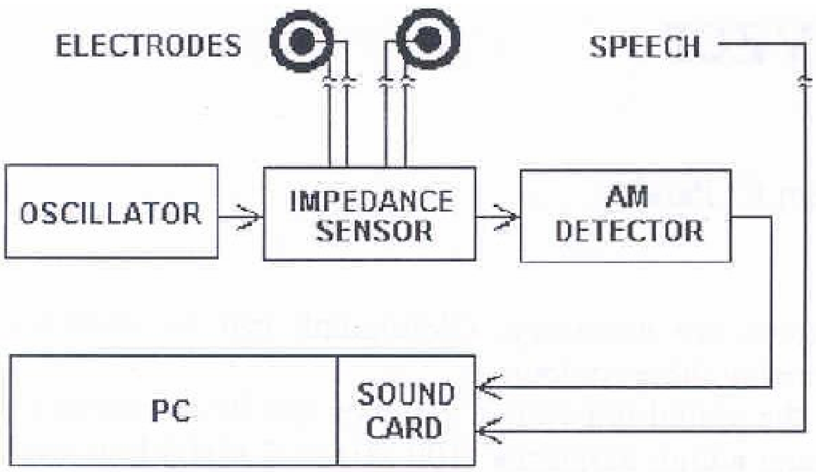

Fig 1. Configuration and principle of working of EGG device. Adapted [11].

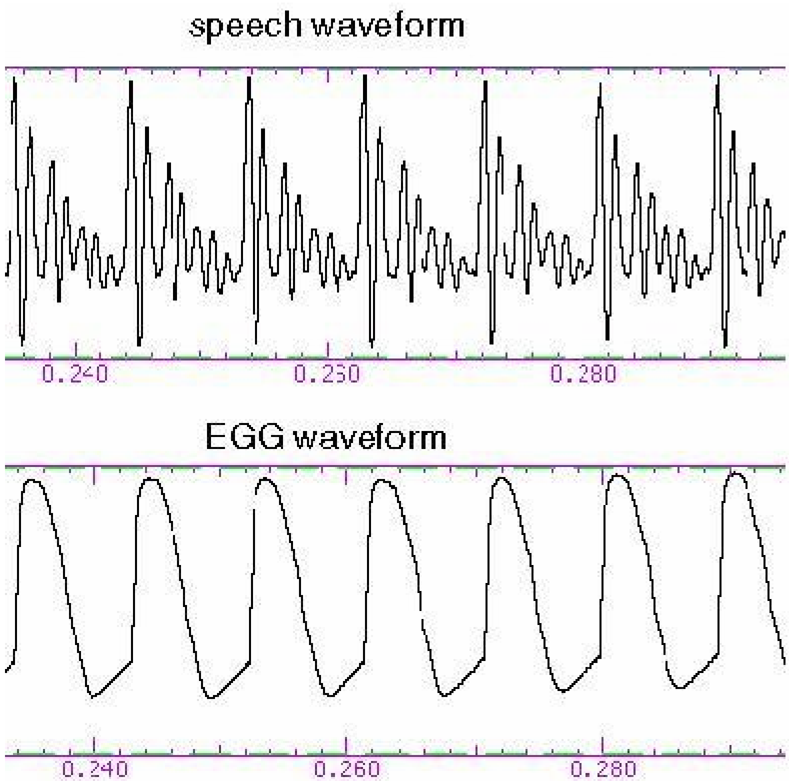

Fig 2. Audio and EGG signal of sustained vowel /a/ production and basic phases of EGG. Adapted [3].

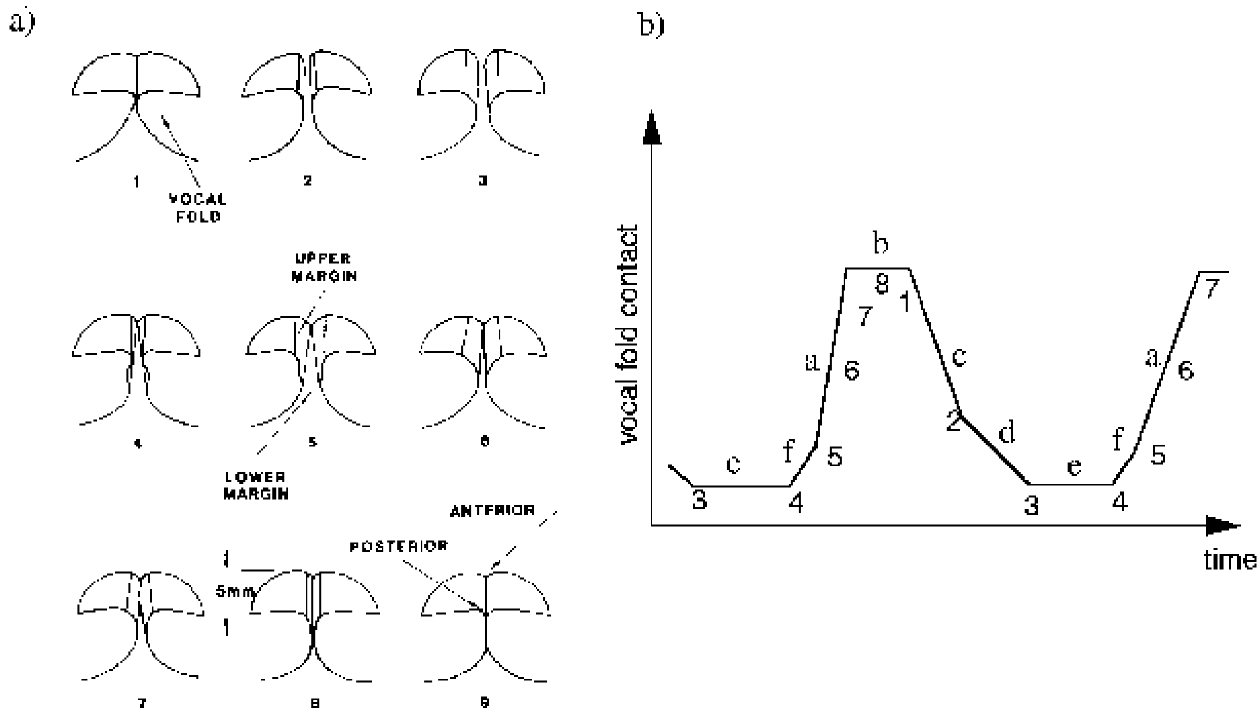

Fig 3. Representation of the vocal fold vibratory cycle. Adapted [4].

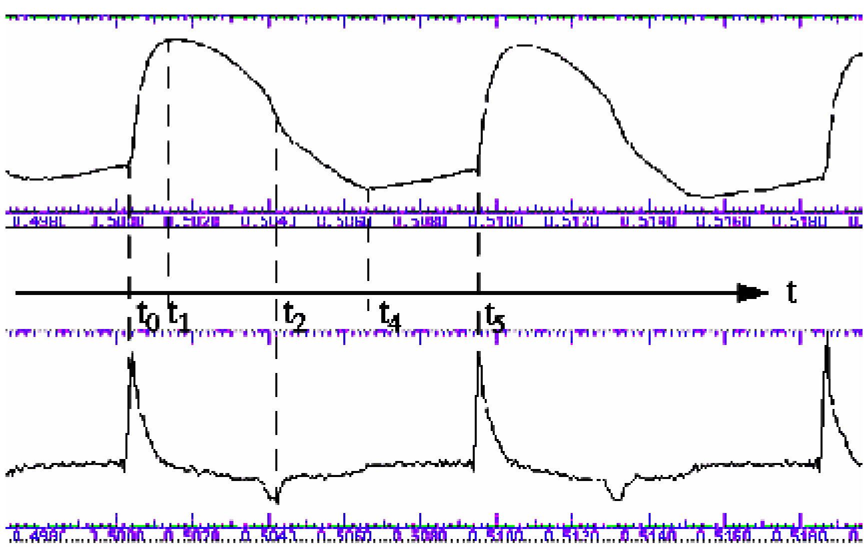

Fig 4. EGG signal and its time derivative. Adapted [4].

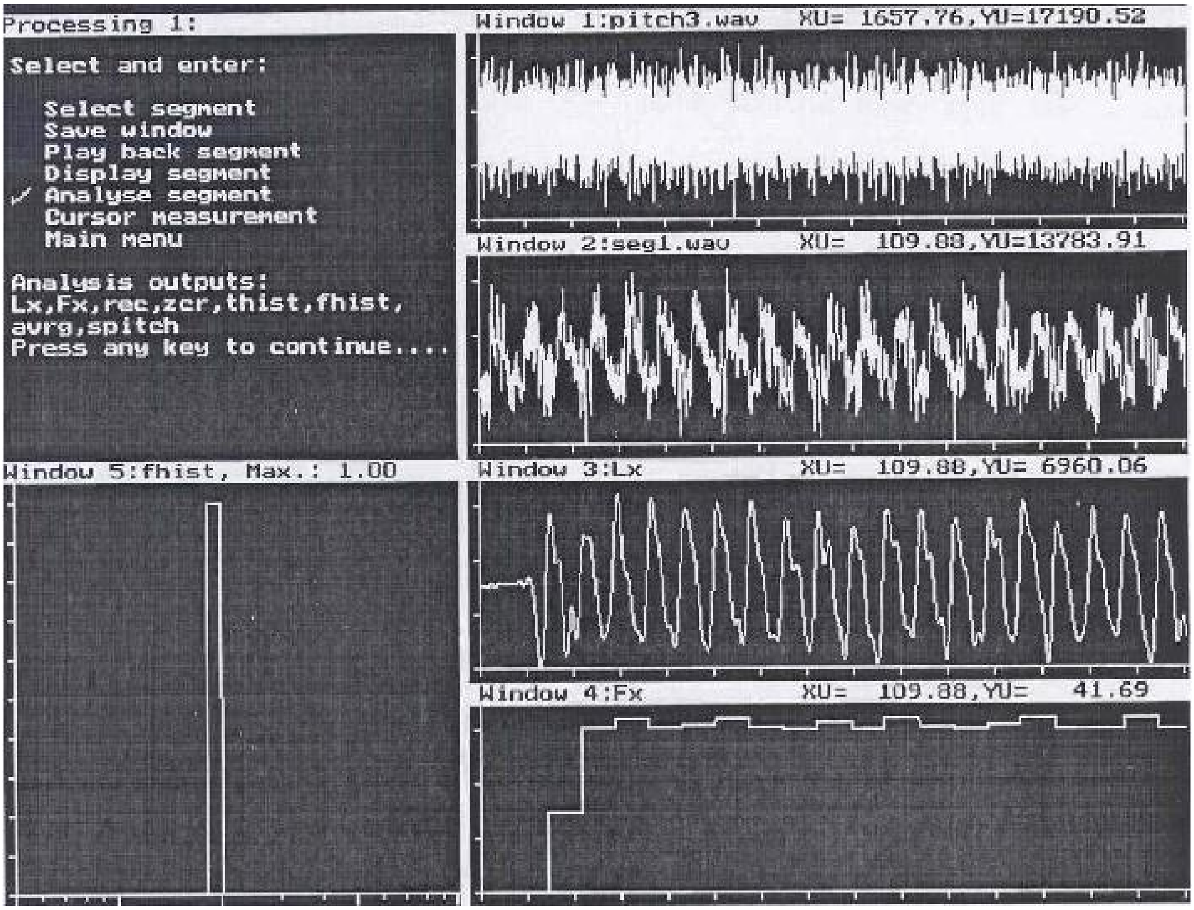

Fig 5. Lx waveform, Fx plot and Fx histogram. Adapted [11].

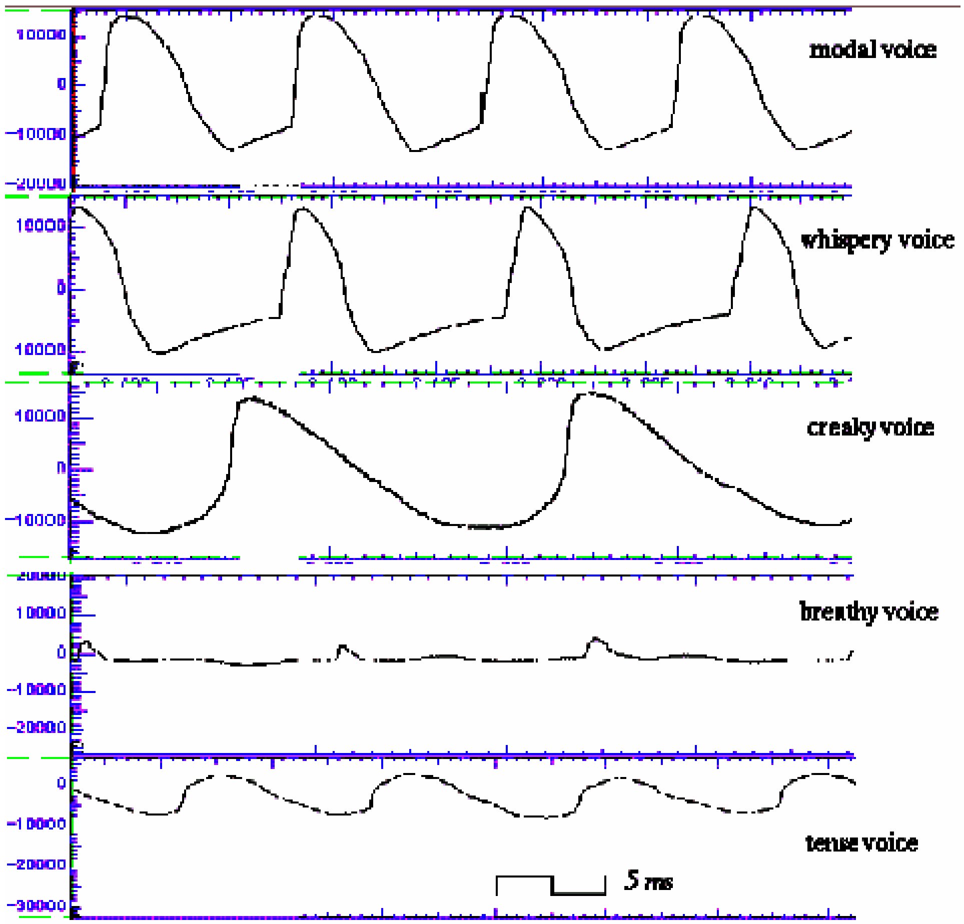

Fig 6. EGG recordings of different voice qualities. Adapted [6].

Table 1. Comparison of various products

		F J Electronics (EG 90)	Glottal Enterprises (EG 2 PC)	Developed at IIT Bombay
1.	Electrodes	Strapped or hand held electrodes	Strapped electrodes	Strapped electrodes
2.	Electrode area	24 mm in diameter.	34 mm in diameter.	30 mm in diameter.
3.	Electrode material used	Bulging circular stainless steel	Gold plated.	Gold plated
4.	Electrode voltage	420 mV	1.3 V	Below 1 V
5.	Electrode current	4.2 mA		1 mA
6.	Material used to hold electrodes	Velcro strap.	Velcro strap.	Velcro strap.
7.	Power source	Alkaline battery 9 V, size PP3, 550 mAh.	Standard dual 6 Volt lead-acid gel cell batteries.	Two nickel hydride batteries. 9V / 120 mAh.
8.	SNR	48 db.	30 db.	45 db
9.	Output impedance	600 ohms		590.10 ohms
10.	BW for impedance detection (frequeny response)	3.5 Hz - 10 KHz.	5 Hz - 6 KHz.	50 Hz - 2.4 KHz.
11.	Carrier frequency	400 KHz	2 MHz	400 KHz
12.	Impedance variation sensitivity	2 V/ohm		1 V/ohm
13.	Total current drain	9.5 mA	160 mA	44 mA
14.	Weight	170 gms	1.7 Kg	450 gms
15.	Dimensions (in mm)	120 * 60 * 30	200 * 266 * 87	200 * 145 * 37.5

References

[1] A.V. Patil, "A glottal pitch extractor", *M Tech dissertation*, Supervisor: P.C. Pandey, Dept. of Electrical Engineering, IIT Bombay, Jan 2000.

[2] http://www.drspeech.com/Electroglottograph.html, site of Dr. Speech Software Group, downloaded on 24 Sep 02.

[3] http://www.ims.uni-stuttgart.de/phonetik/EGG/pagee1.htm, site of Institute of Natural Language Processing, University of Stuttgartgart, Germany, Jul 2001, downloaded on 26 Sep 02.

[4] http://www.ims.uni-stuttgart.de/phonetik/EGG/pagee2.htm, site of Institute of Natural Language Processing, University of Stuttgartgart, Germany, Jul 2001, downloaded on 28 Sep 02.

[5] http://www.ims.uni-stuttgart.de/phonetik/EGG/pagee2.htm, site of Institute of Natural Language Processing, University of Stuttgartgart, Germany, Jul 2001, downloaded on 28 Sep 02.

[6] http://www.ims.uni-stuttgart.de/phonetik/EGG/pagee7.htm, site of Institute of Natural Language Processing, University of Stuttgartgart, Germany, Jul 2001, downloaded on 01 Oct 02.

[7] M. Rothenberg, "A multichannel electoglottograph", *Journal of Voice*, Vol. 6., No. 1, pp. 36-43, 1992.

[8] http://www.laryngograph.com/products/lxproc.htm, site of Laryngograph Ltd, downloaded on 04 Oct 02.

[9] http://www.f-jelectronics.dk/htm/electroglotto.htm, site of F-J Electronics, downloaded on 06 Oct 02.

[10] http://www.glottal.com/electroglottograph.html, site of Glottal Enterprises, downloaded on 08 Oct 02.

[11] P.K. Lehana and P.C. Pandey, "A low cost impedance glottograph and glottal pitch analyzer," *proceedings of Bio Vision 2001 International Conference on Biomedical Engineering*, Bangalore, India, pp 33-37, Dec 2001.