For completeness, it should be mentioned that a similar method has been used experimentally in which the oral vocal tract is closed by a mechanical valve, and the intraoral pressure recorded [9]. However, problems with the method and the necessity of speaking into a mechanical valve have kept it from being used extensively during natural speech.

However, the interpolation technique as originally implemented does not provide a real-time measurement. It has been used primarily for the analysis of previously recorded speech or singing, using a digital computer to implement the interpolation algorithm. One result of this limitation is that the interpolation technique as implemented according to the method originally proposed by Rothenberg could not be used conveniently for biofeedback in voice training exercises. In addition, the technique implemented according to this protocol is cumbersome when used in routine speech testing.

The repeated syllable algorithm used in the paper by Rothenberg [7] was designed to determine an estimate of the subglottal pressure during inverse filtered glottal airflow waveforms produced with different vowels, fundamental frequencies, degrees of vocal fold abduction and subglottal pressures. However, the application so described was not meant to exclude the use of the general method of interpolation for estimating the subglottal pressure in fluent speech or singing. The underlying principle is that subglottal and intraoral pressures equilibrate during the closure phase of an unvoiced stop consonant. This principle is valid during all speech and singing.

In this paper, it is proposed that the interpolation technique can be adapted to real time measurements in continuous speech or singing, and a practical system for doing this is outlined.

Materials and Methods

Real time interpolation of subglottal pressure during speech or singing

Figure 1 shows the basic elements of a system for estimating subglottal pressure from the intraoral pressure, assuming the intraoral pressure being used is the pressure behind the closure of a bilabial stop.

In Figure 1, the intraoral pressure is sampled by a fine tube inserted at the corner of the mouth or between the lips. The tube leads to a pressure transducer and associated preamplifier capable of measuring pressures in the range found intraorally (A miniature transducer inserted in the mouth could also be used).

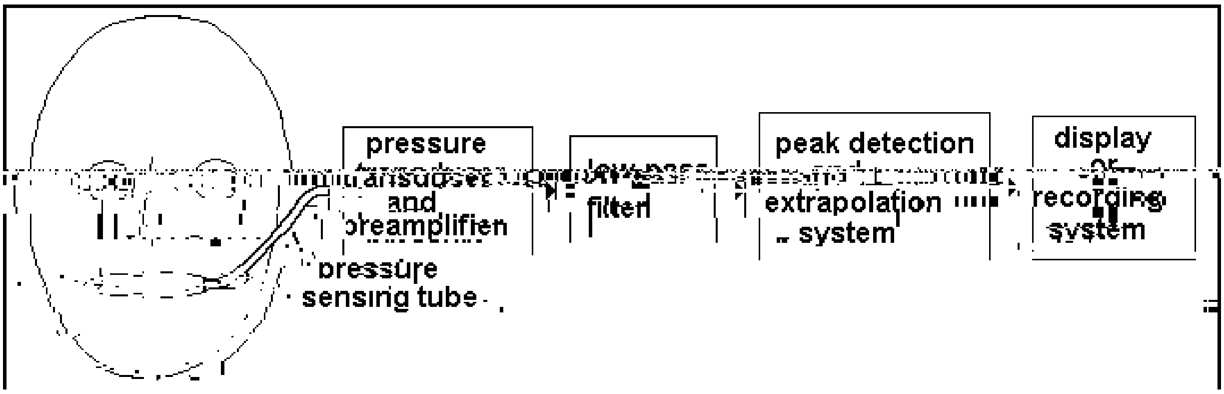

Figure 1: Block diagram of a system for sub glottal measurement and display during speech.

Since the output of the transducer/preamplifier, in addition to the low frequency pressure signal that represents the subglottal pressure, can be expected to contain unwanted acoustic pressure information, a low pass filter can be used to eliminate the acoustic pressure variations. Since the acoustic pressure variations in the mouth are generally above about 50 Hz, the low pass filter might have a cutoff frequency of approximately 30 to 40 Hz.

Since it is the peak of the oral pressure pulse during the stop closure that approaches the subglottal pressure, there must be a functionality that detects this peak and holds it for a period of time sufficient for it to be measured. This functionality can be accomplished by a digital computer program, but we suggest here an electronic circuit commonly called a peak detector, or peak-hold circuit, as shown in its simplest form in Figure 2.

In the standard peak-hold circuit in Figure 2, the diode D1 conducts whenever the input voltage Vin is greater than the voltage of C1, bringing Vc and Vout to the value of Vin, assuming an ideal diode D1.

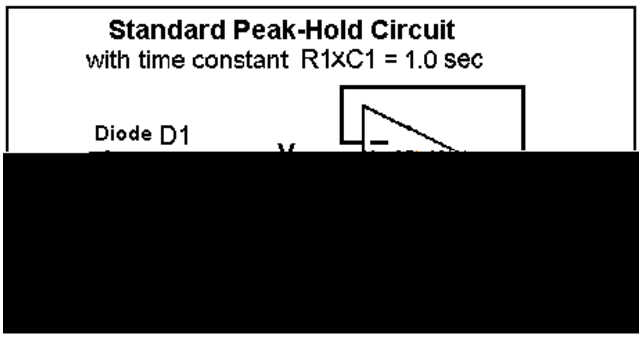

Figure 2: Electrical circuit diagram for a standard peak-hold circuit

the capacitor voltage Vc during the post-peak decay, the decay is exponential with a time constant equal to R1 × C1, as shown in Figure 3 for two values of the time constant R1 × C1.

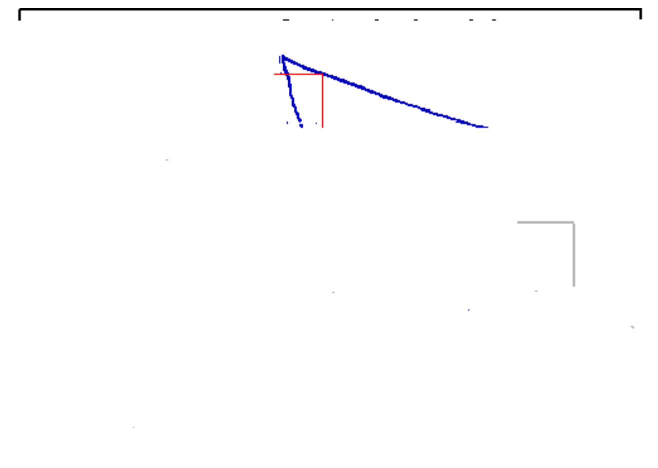

Figure 3: Chart showing the time response of a standard peak-hold circuit with a time constant of 1s and 10s to a negative step in the input of magnitude one, going from one volt to zero volts at t=0. The red vertical lines show the duration of the hold period as defined here, for each time constant. The refractory period, as defined here, is also indicated for a time constant of 1 second.

We will consider that the time during which the output of a peak-hold circuit is within 5% of the detected peak value as a 'Hold' time. The annotation in Figure 3 illustrates that the output of a standard peak-hold circuit holds the detected peak value to within 5%, for a period of time of approximately 1/20 x the time constant.

A circuit with a time constant of 1.0 seconds will take 1.0 seconds to decay from 1.0 to 1/e=0.37 sec, and a period of approximately 0.7 seconds to decay to half of its peak at t=0). During this period of exponential decay, the circuit will not detect and register other peaks that are less in amplitude than the circuit output. We will refer to this period as a refractory period.

A 'refractory period' is generally defined as 'a period immediately following stimulation during which a nerve or muscle is unresponsive to further stimulation. It is often applied to other physiological responses in which a recovery is required after a response. In other words, a peak-hold circuit requires a recovery from a given peak pressure in order to respond to a subsequent peak.

Since the subglottal pressure is a slowly changing variable, a refractory period of T seconds means that pressure peaks larger than 50% of the previous peak (our

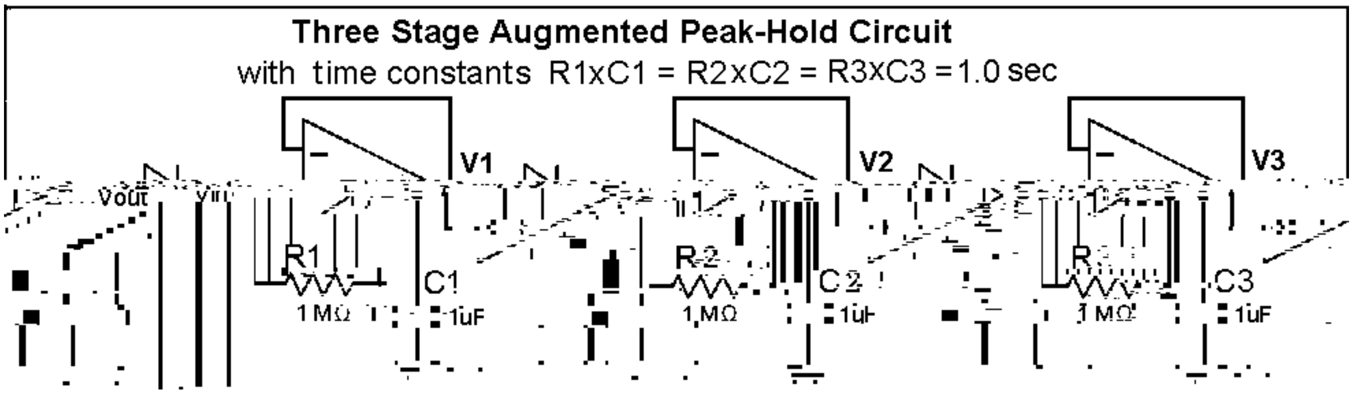

Figure 6: An electrical circuit diagram for an augmented peak-hold circuit having three stages, with a time constant of one second in each stage.

and this conclusion has been conﬁrmed by tests using a 5-stage APH circuit, though the optimal number of stages and the optimal values of

Such a decrease in the time constant may not be required in practice since the rate of subglottal pressure change is limited in actual speech.

To illustrate the response of the same system to the subglottal pressure variation found in actual speech, Figure 11 shows the low pass filter output and augmented peak-hold circuit outputs for the same adult male subject, speaking the sentence “Peter Piper picked a peck of pickled peppers” spoken at a conversational voice level.

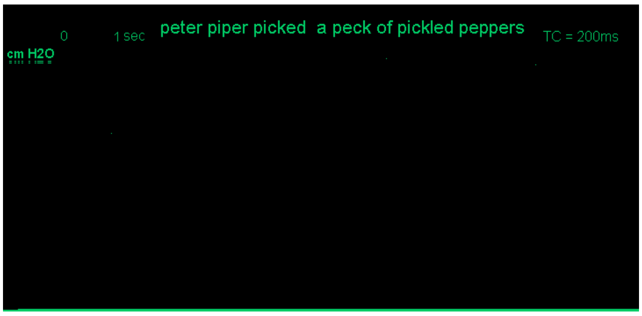

Figure 12: Measured low pass filter output and augmented peak-hold output for a spoken sentence having numerous /p/ consonants. The low pass filter output indicates that for the /p/ marked by a star, the tube was probably occluded, as by the tip touching the tongue.

In each case in Figure 13, the APH output is held long enough at the subglottal pressure for an observation to be made. However, it should be noted that the unit used for these traces was provided with a manual Hold capability triggered by a pushbutton switch, so that it was only