This is an expanded version of a paper of the same name in the *Proceedings of the 10th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research, June 3-4, 2013, Cincinnati, Ohio, USA*

RETHINKING THE INTERPOLATION METHOD FOR ESTIMATING SUBGLOTTAL PRESSURE
(EXPANDED VERSION)

Martin Rothenberg

Syracuse University (Emeritus), Syracuse, N.Y, USA, and Glottal Enterprises Inc.., Syracuse, New York, USA

***Keywords*: Voice; subglottal pressure; intraoral pressure; laryngeal flow resistance; interpolation method.**

ABSTRACT

The air pressure in the lungs, or more precisely, the subglottal air pressure, is the primary energy source in speech. Yet many clinicians avoid estimation of subglottal pressure in treating patients because of the difficulty in measuring it. However, since its first reported use in 1973, the so-called interpolation technique, in which the subglottal pressure during voiced speech is estimated from samples of intraoral pressure obtained during bilabial unvoiced stops, has achieved acceptance as a feasible and reliable method, if used properly. This paper describes some of the uses and misuses of the interpolation technique and questions some assumptions that have been made about its use. A new version of the method is described that is simple and economical and provides real-time pressure estimates for speech evaluation and biofeedback.

INTRODUCTION

Though it may seem obvious from elementary principles of physics that the intraoral pressure will eventually equalize with the subglottal pressure during an unvoiced stop closure, when the oral chamber is closed and the glottis is open, researchers have been rediscovering this periodically since the early 1960s. The only significant question to be answered if subglottal pressure is to be measured intraorally during the closure is 'how fast do the pressures equalize?' An analysis in the Appendix indicates that the time for equalization is approximately 25 ms. This means that equalization will occur early into the oral closure period for a stop consonant.

In the early 1970s, Rothenberg(1973) used this equalization principle in what I will call the **Interpolation Method** for estimating subglottal pressure. In the experiments reported by Rothenberg, repeated syllables containing a consonant /p/ were used to estimate the subglottal pressures during vowel segments between the consonants. The pressure measurements thus obtained were combined with simultaneous measurements of airflow using a circumferentially vented pneumotachograph mask. Here is a sample of a pressure measurement from the 1973 paper, in what appears to be the first report of the use of this technique.

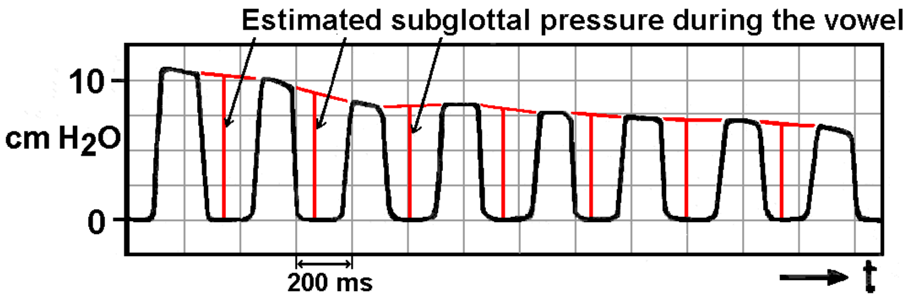


Fig. 1: Intraoral pressure during contiguous repetitions of the syllable /b æ p/. Adapted from Rothenberg (1973).

The relatively flat region at the apex of each pressure pulse could be assumed to exhibit the variation of Psg during the lip closure.

The Interpolation Method allowed Rothenberg to not only explore the effect on the voice of varying Psg but also explore the effect of varying F0 and glottal adduction "voice quality" with a fairly constant Psg.. Some of the results are accumulated in Figure 2.

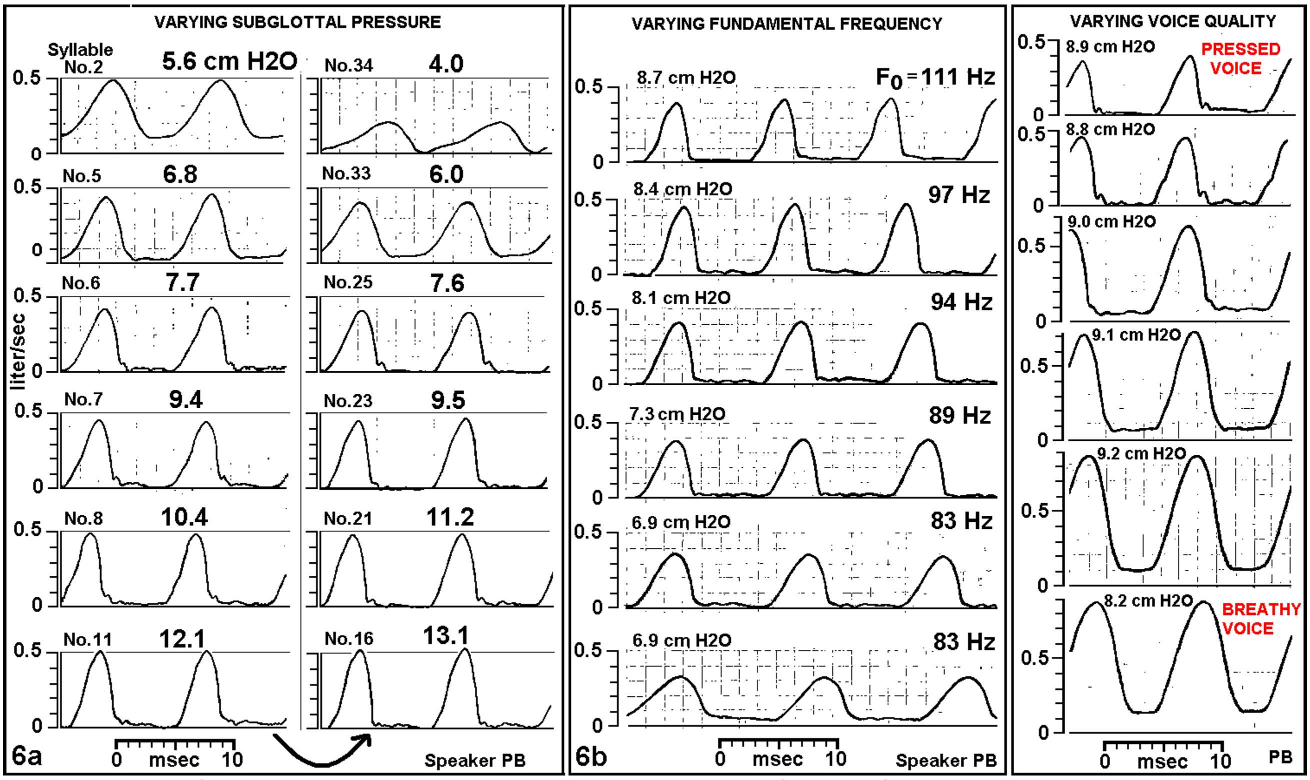


Fig. 2: Glottal airflow during the vowel /ae/ as determined by inverse filtering, for a single untrained adult male subject with a normal voice. From Rothenberg (1973).

SOME PROBLEMS WITH THE INTERPOLATION METHOD

Since its introduction in 1973, the Interpolation Method has become quite popular in research and clinical applications, and there are numerous examples in the literature of the proper use of the method. However, there have also been some problems in its use. Some of these problems are illustrated in the waveforms in Figure 2, in which the sample intraoral pressure waveform given by Rothenberg in the first use of the Interpolation Method is compared with problematic waveforms presented in the later literature in two well known reference texts and two research papers, presented in chronological order.

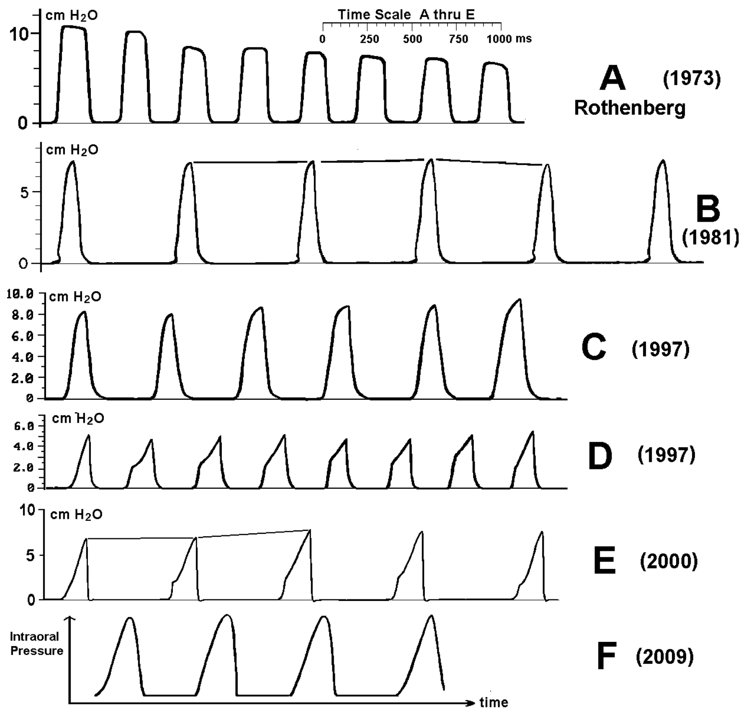


Fig. 2: Five aberrant recordings of intraoral pressure, B through F. The lines connecting peaks in B and E are recommended by the respective authors as a basis for interpolation.

In Waveform A, the relatively flat regions near the peaks of each pressure pulse show that:

(1) the subglottal pressure was approximated intraorally during the stop closure interval, and

(2) the variation of subglottal pressure over each syllable is relatively slow and thus suitable for the Interpolation Method.

Waveforms B through F in Figure 2 can be readily seen to be different than waveform A in that they lack the relatively flat peak. We will argue that this negates these waveforms as a candidate for the Interpolation Method.

Differences among the samples that were not significant in the use of the Method include the use of the different vowels. Rothenberg used /æ/ because it is easier to inverse filter. Other authors specify that a vowel /i/ should be used. However, the rationale for using /i/ is not made clear by any of the authors, and in the opinion of this author, it is being dropped in practice. Thus, the vowel appears to be not important to the technique.

TWO SOURCES OF ERROR IN THE INTERPOLATION METHOD

There are two primary sources of error that in B through F in Figure 52 that can negate the use of the Interpolation Method and lead to widely incorrect pressure estimates..

1. A measurement system that responds too slow

A "pointed" pulse in intraoral pressure can result from the use of a pressure measurement system with too long a response time. This problem appears to been the source of the aberrant waveforms in B and C of Figure 2. Intraoral pressure pulses in a /p/ consonant can have a duration of as little as 80 or 90 ms. To display adequately a pulse of intraoral pressure that has a duration of 80 ms, the pressure measurement system should have a response time of no more than about 20 ms, which is equivalent to a frequency response that extends to at least about 30 Hz, -3 dB. (See Baken and Orlikoff, 2000)[2]

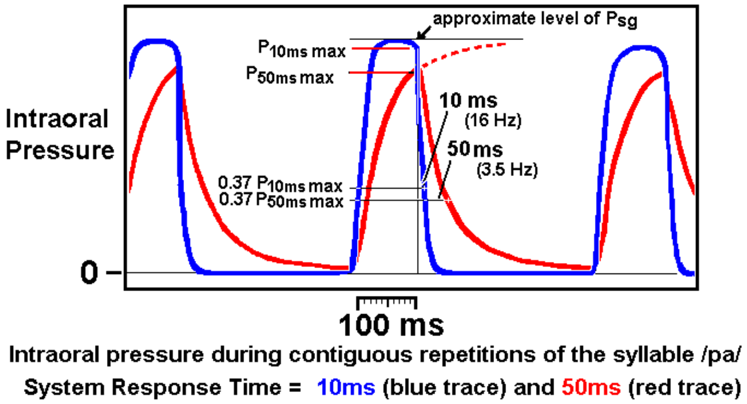


Fig. 3: Intraoral pressure measured by two systems simultaneously, to illustrate the creation of a pointed pressure pulse when the system response time is too long (red trace).

2. Intra-syllable subglottal pressure not fairly constant

A primary assumption of the Interpolation Method is that the subglottal pressure varies slowly enough so that an interpolation makes sense. If the syllables are 'punched out', each syllable with its own lung pressure impulse, the technique cannot (or should not) be used. That this is happening in waveforms D and E in Figure 2 is made clear by the analysis of Hertegård, Gauffin and Lindestad (1995) from their measurements of subglottal pressure using a tracheal puncture, simultaneous with intraoral pressure.

The sound pressure trace in Figure 4 shows that the subglottal pressure decreased during the vowel, after attaining a peak just after the aspiration interval. Moreover, the ceasing of voicing (first blue line) indicates that the subglottal pressure had decreased below the phonation threshold pressure at that point. The dashed red line in the figure shows a hypothesized variation of subglottal pressure that satisfies these constraints. Obviously the syllable was produced with its own impulse of subglottal pressure ("punched out"). It is also obvious that there is no possible interpolation of intraoral pressure that makes any sense in estimating subglottal pressure.

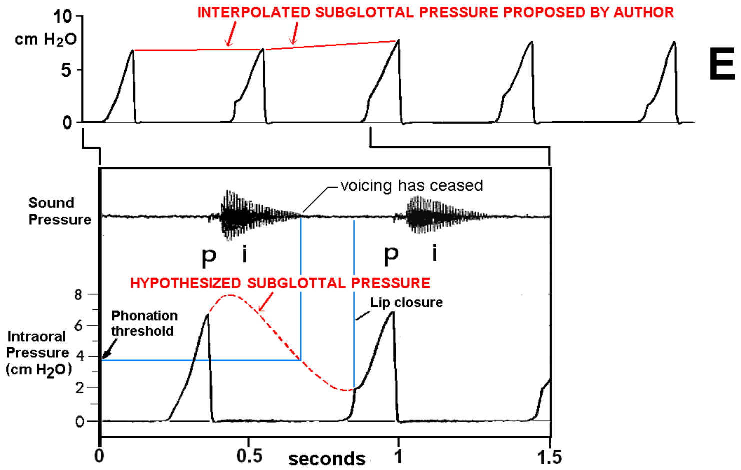


Fig. 4: A portion of the intraoral pressure trace E in Figure 2, including the sound pressure trace that accompanied the intraoral pressure in the original publication.

The author has seen instances of repeated syllables in which the punching out is not as extreme as in Figure 4, but interpolation would never-the- less be problematic. An example of this, taken with Aeroview by a user and sent to Glottal Enterprises for analysis, is shown in Figure 5.

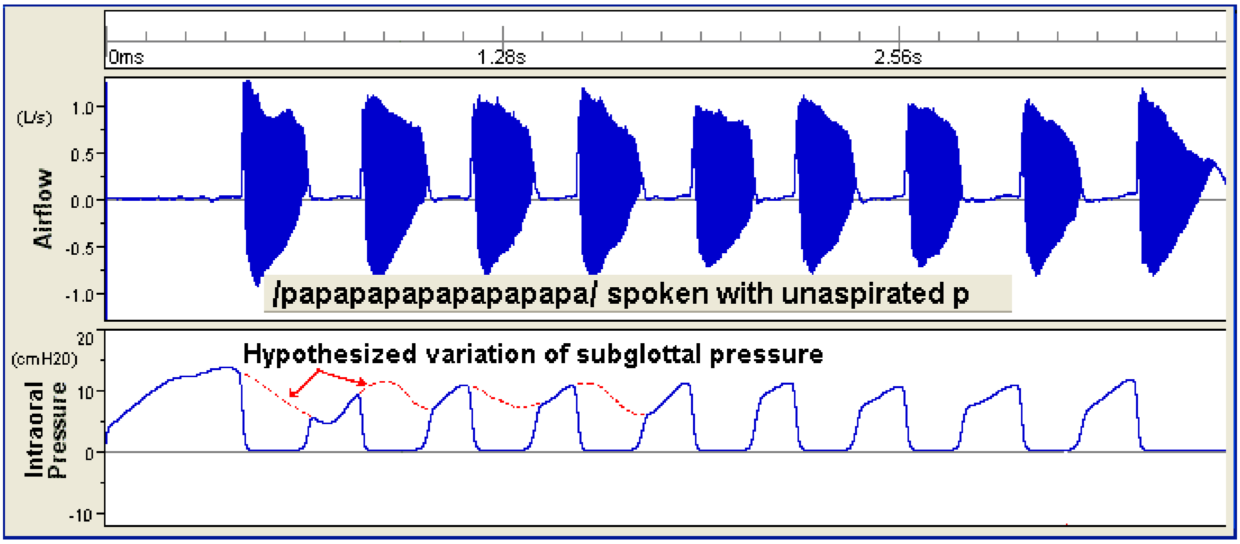


Fig. 5: A recording of a repeated /pa/ syllable in which there was evidence of a slight punching out of the individual syllables. Cropped for clarity and annotation added. The airflow trace (from a Glottal Enterprises CV mask) shows that the subject did not aspirate the release of each /p/.

The airflow trace in Figure 5 (from a Glottal Enterprises CV mask) shows that the subject did not aspirate the /p/ releases . Thus the tilting of the apex of the pressure pulses was undoubtedly caused by the uneven articulation of the syllables, and not by aspiration as described below.

ELIMINATING ASPIRATION?

A word-initial /p/ in English is usually pronounced with moderate aspiration, which will lower the subglottal pressure slightly just after the release. This drop in pressure can be expected to cause an increase in respiratory muscle activation if subglottal pressure is to be maintained. One result would be that the intraoral pressure pulses might have a tilt upwards as a result of this increased respiratory activation, as in the recording by Hertegård, et al. (1995) shown here in Figure 6 - apparently recorded with a moderate aspiration of the /p/. The subglottal pressure traces reported by Hertegård, et alwere recorded through a tracheal puncture. The error that would be caused by interpolating between the peaks of the pulses of intraoral pressure in Figure 6 would be approximately 12%.

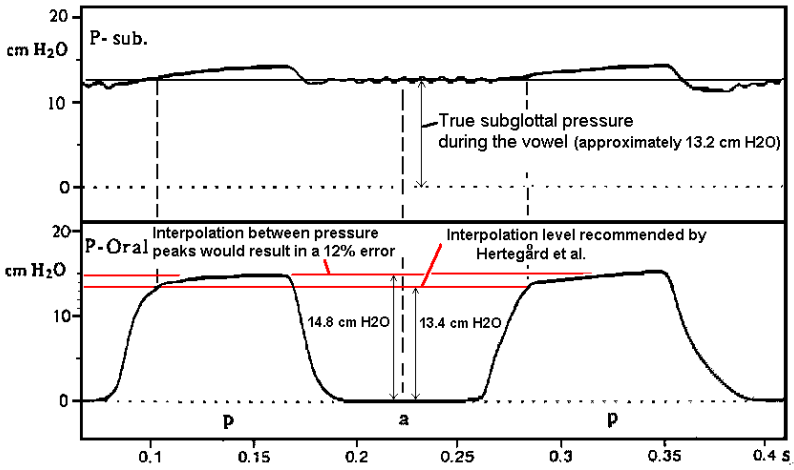


Fig. 6: Adapted from Hertegård, et al. (1995)

In the example in Fig.7, apparently produced with a high level of aspiration for the /p/, the percentage decrease in subglottal pressure caused by the aspiration in the release of the /p/ is about 32%, thus making the intraoral pressure peaks an especially poor indication of the subglottal pressure during the vowel. If the peaks of the pressure pulses, at 11.2 cm H2O, were used for a linear interpolation (as recommended by the respective authors for waveforms B and E of Fig. 2.), **the interpolated value of subglottal pressure obtained would have been 47% too high.**

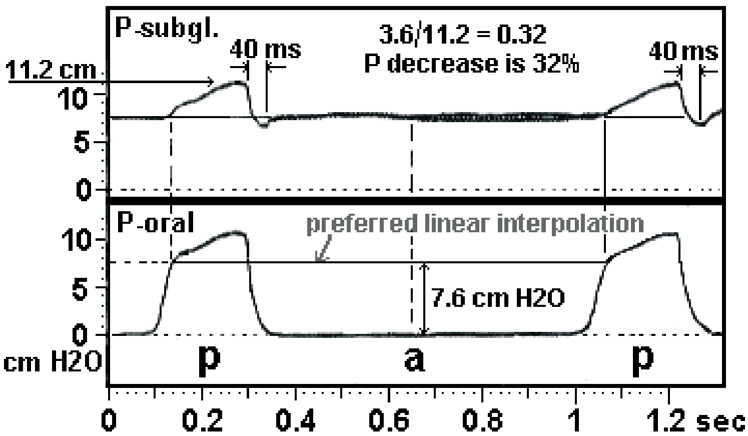


Fig. 7: Adapted from Hertegård, et al. (1995)

Eliminating aspiration - When the Interpolation Method was first used by Rothenberg, he proposed that a double or joined consonant /pb/ could be used instead of a single /p/, by having the subject repeat the syllable "b vowel p" instead of "p vowel" or "vowel p". In this way the sudden drop in subglottal pressure caused by the aspiration in the /p/ is avoided.

A second advantage of the /b vowel p/ syllable is that for the same speaking rate, the intraoral pressure pulse is approximately 20 to 25% longer in duration with a joined /bp/ consonant than with a /p/, and therefore easier to interpret.

Another advantage of the /b vowel p/ syllable is that there would be a decreased tendency for the phonetically naive subject to punch out the individual syllables with a strong variation in subglottal pressure, since the presence of the final consonant requires that subglottal pressure be maintained

Given these advantages, a /b vowel p/ syllable is preferred in the Interpolation Method, with the caveat that phonetically trained subjects, such as professional singers, can be instructed to use little aspiration for a word-initial unvoiced stop and therefore can use the /p vowel/ syllable if it is easier for them in a specific context.

Another possible solution for English speakers is to use a short word with a non-aspirated /p/, such as “spa”. Thus the subject would repeat “spa spa spa”. There are other solutions possible for speakers of other languages. For example, a Spanish speaker can repeat a short word starting with an unaspirated /p/, such as ‘pan’ (‘bread).

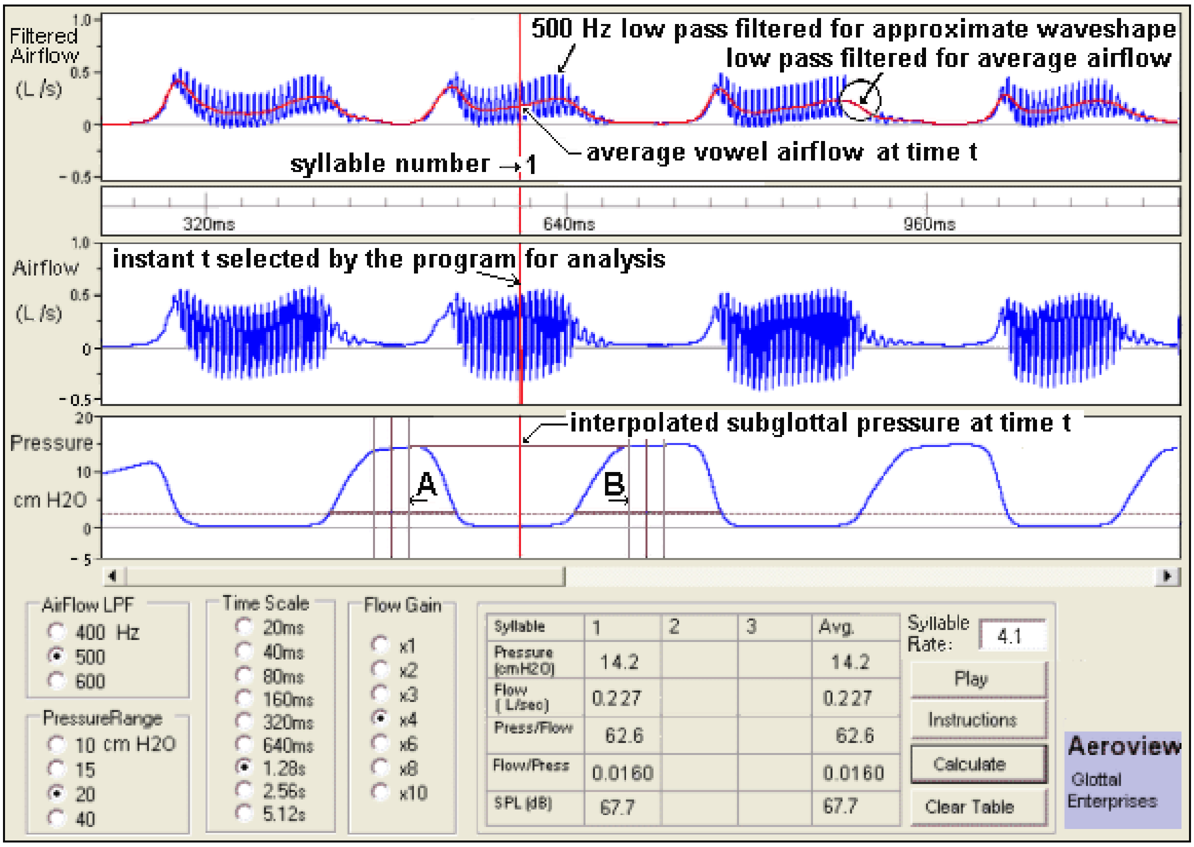


Fig. 5: A screen from a repeated syllable /baep/ as recorded using an Aeroview system from Glottal Enterprises

The Aeroview system from Glottal Enterprises, used for the recording shown in Figure 5, has an automated procedure for selecting the instants during each of the two successive oral pressure pulses that are used by the linear interpolation routine. The selection procedure for the instants used for interpolation was designed to optimize the interpolation process for syllables in which the /p/ is unaspirated or has only slight or at most moderate aspiration. More details about the algorithm used can be obtained from the company, It can be seen in the figure that the instants were not chosen at the pressure peaks, as advocated by others. The screen in Figure 5 has been cropped to show more clearly the instants that were chosen by the program, identified in the figure by vertical lines A and B.

IMPLEMENTING INTERPOLATION IN REAL-TIME

The Interpolation Method, as originally put forth by Rothenberg (1973), was not restricted to interpolating between two consonants.[1] If we are not restricted to interpolating between two consonants (the method now used widely) then it is possible to arrange some type of time-limited peak-hold electronic circuit or digital processing that will continuously display an estimate of subglottal pressure in real time. Glottal Enterprises currently markets such a real-time subglottal pressure display under the name PG System. It is a handheld, battery-operated unit that connects to a short piece of flexible tubing placed between the lips. A pushbutton, when depressed, holds the value at that time for easy reading. Though accuracy will not be as high as using a computerized non-real-time interpolation system. Testing has shown that if used correctly, the readings of the device agree well with estimates of subglottal pressure obtained with a computerized non-real-time interpolation system. Details of the test procedure and the results will be presented elsewhere.

CONCLUSIONS

1. Use a pressure measurement system with a response time of under 20 ms.

2. Consider using a concatenated /p/ and /b/, or /pb/, as the consonant instead of an aspirated /p/, by having the subject repeat the syllable /b vowel p/ instead of /p vowel/.

3. Forget about a restriction to the one vowel /i/ that is proposed by some in the literature.

4. Consider a syllable repetition rate greater than the 1.5/second proposed by some, to reduce the tendency to punch out syllables.

5. Try using natural speech. For an English speaker, preferably use words with a medial unaspirated /p/, as the medial p in “pepper in Figure A2, following a fricative consonant as in “spa”.

6. Avoid using an interpolation system that is restricted to interpolating between the peak values of the intraoral pressure pulses. (Ask the manufacturer if this is not clear.)

7. Examine captured intraoral pressure waveforms, looking for signs of punching out syllables, and delete those examples.

8. Consider using a real-time implementation of the Interpolation Method for instant readings and the possibility of biofeedback.

APPENDIX - Time Required for Pressure Equalization

A. Calculated.

To calculate the maximum time required for the intraoral pressure in a /p/ consonant to approach the subglottal pressure, we need to know the maximum compliance of the supraglottal chamber and the maximum flow resistance of the glottis when abducted for the unvoiced consonant.

Compliance of the Supraglottal Chamber during a Bilabial Stop Closure.

The compliance of the oral chamber with a bilabial closure, which would be primarily that of the walls of the chamber, for a male adult subject as measured by Rothenberg (1968, Table 7.2) is represented in the following chart.

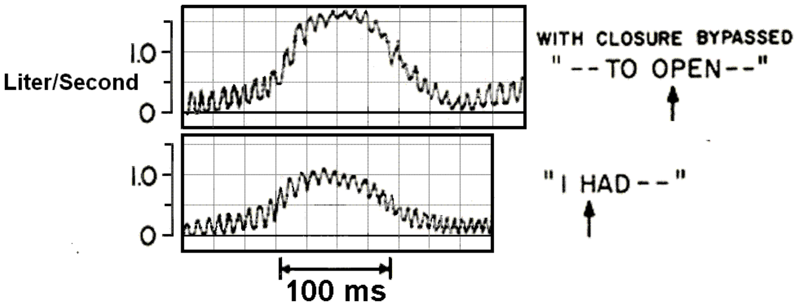
Measurements of the compliance of the superglottal cavity, C_w

Place of closure	Tension of walls	Peak to peak air volume inserted (ml)	Number of Measurements	Compliance C_w Mean	Compliance C_w Standard Deviation
Bilabial	lax (including cheeks)	5	13	0.004	0.00056
Bilabial	lax (except for tense cheeks and lips)	1	8	0.00068	0.00012
Alveolar	lax	1	7	0.00053	0.00012
Alveolar	tense	1	8	0.00038	0.00010

We can estimate from this chart that the compliance of the chamber for a bilabial stop would be likely to be equal to or less than the figure 0f .00068 liter/cm H2O, but could conceivably be as high as .004 liter./cm H2O with cheeks and lips kept very lax.

Glottal Flow Resistance During an Abductory Gesture

The approximate flow resistance of the glottis during an abductory gesture can be estimated from the following airflow records, taken from Rothenberg, 1971. (We use the term 'resistance' here for convenience, though it implies a linear relation between pressure and flow. The pressure and flow are not linearly related for the glottis.)



Volume velocity waveforms during a cyclic glottal opening movement.
[Adapted from Rothenberg, 1971, Figure 4]
top: voiced — open — voiced (minimum duration)
bottom: voiced — breathy — voiced

FIGURE A1

The airflow recordings in Figure A1 were taken using a circumferentially vented pneumotachograph mask covering the mouth and nose. In the upper trace, the words "to open" were pronounced to yield an unaspirated /p/ that had an abductory gesture of minimal duration and extent.. A short piece of tubing (~1 cm long and 1/4 or 5/16 inch ID) was inserted in the corner of the mouth (under the mask) so as to record the airflow that would have occurred if there were no closure. Because of the flow resistance of the tubing, the peak flow would have been slightly larger than the value in the figure, but we accept the flow in the figure as giving a minimum for the desired flow. The flow trace shows the flow in an /h/ pronounced with a close-to-minimal duration for the abductory gesture.

From the traces in the figure, we can estimate that the flow during the most open phase of an abductory gesture would be between approximately 1 and 1.6 liters/second. If we can assume that the subglottal pressure for these records was a typical adult value of 8 cm H2O, the flow resistance of the glottis when open would be between 5 and 8 cm H2O per liter/sec.

Multiplying glottal flow resistance times chamber compliance, we get a range of 3.4 to 5.4 ms for the likely pronunciation. In a simple resistance-compliance dynamic system, the output reaches 95% of the input at 3 times the system time constant. This rule yields a time for equalization within 5% of approximately 10 to 15 ms.. Since a brief intraoral pressure pulse is about 100 ms in duration (see the examples above), the pressures can be expected to equalize in less than 15% of the pulse duration.

B. Measured.

To test the above theoretical limitations, we can look more closely at the medial /p/ in 'pepper, in the simultaneous intraoral pressure and airflow records in Figure A2. Since the pressure trace is displayed with no smoothing filter, the rise in the pressure pulse can be assumed to reflect the time constant for the glottal air flow building oral pressure after the bilabial closure is completed.

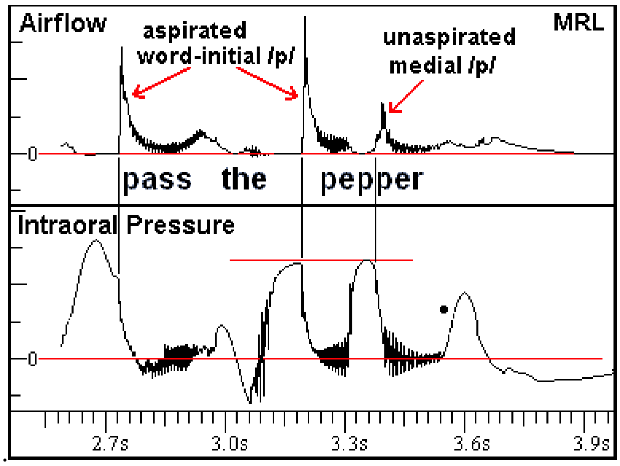


FIGURE A2

In Figure A3 we show the word “pepper” from Figure A2, but expanded horizontally by a factor of four to make the time constant more easily measurable.

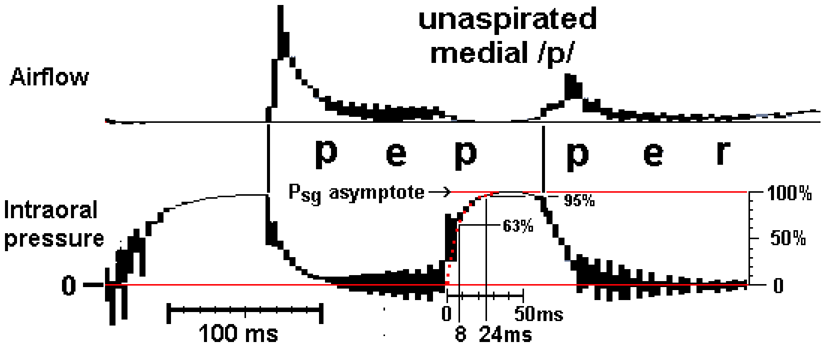


FIGURE A3

The rise in pressure, traced by red dots in the figure, exhibits a quasi-exponential form. The instants at which the rise reaches 67% (one time constant, TC) and 95% (three time constants, 3TC) are projected down to the time scale to graphically determine that the rise time constant is approximately 8 ms and 3TC is 24 ms. The theoretical and measured estimates are compared as follows:

theoretical range		**measured**	
TC	3TC	TC	3TC
3.4ms -----	10ms	8ms --------	24ms
5.4ms -----	15ms		

The time constant TC for the two estimates of glottal flow resistance would be 20ms and 32 ms for the extremely lax pronunciation, with the corresponding 3TC values at 60ms and 96ms. Thus, even in the case of the unnaturally lax pronunciation, the pressures would equalize within 5% by the end of the pressure pulse. However, we have seen no signs of equalization times that large in our work or in the literature.

These theoretical and measured values of equalization time for subglottal and intraoral pressure during a bilabial stop closure generally agree with the estimates made by Hertegård , et al. (1995)

References

Baken, RJ and Orlikoff, RF (2000), *Clinical Measurement of Speech and Voice 2e,* Singular Thompson , San Diego.

Hertegård S, Gauffin J, & Lindestad PA (1995), A Comparison of Subglottal and Intraoral Pressure Measurements during Phonation, *J Voice* 9(2):149-55.

Rothenberg, M, (1972) The Glottal Volume Velocity Waveform During Loose and Tight Voiced Glottal Adjustments, Proceedings of the Seventh International Congress of Phonetic Sciences, 22-28 August 1971; ed. by A. Rigault, and R.Charbonneau, Mouton, The Hague, 380-388.

Rothenberg, M **(1973),** A New Inverse-Filtering Technique for Deriving the Glottal Airflow Waveform during Voicing. *J. Acoust. Soc. Amer.*53: 1632-1645. (1973)

This expanded version of the paper is posted on the author's website, www.rothenberg.org.