THE FUNDAMENTAL FREQUENCY - SUBGLOTTAL PRESSURE RATIO

Helmer Strik and Louis Boves.

University of Nijmegen, Dept. of Language and Speech, Phonetics section,
P.O. Box 9103, 6500 HD Nijmegen, The Netherlands

ABSTRACT

It is known that subglottal pressure ($P_{sg}$) is a major factor in the control of fundamental frequency ($F_o$) in speech. Yet, the details of this relation remain unclear. Estimates of the $F_o$ to $P_{sg}$ ratio (FPR) from speech and special phonation tasks yield values between 5 and 15 Hz/cmH2O [1,2,3,4]. In another type of experiments pressure variations are induced externally, either subglottally or supraglottally. The FPR's measured in these experiments tend towards values of 2-5 Hz/cmH2O [5,6,7,8]. There seems to be no a priori reason for the FPR to be different in both kinds of experiments. Therefore we carried out experiments that aimed at resolving this discrepancy.

I. THE FPR IN EXPERIMENTS WITH INDUCED PRESSURE VARIATIONS

INTRODUCTION

The FPR in experiments with artificially induced pressure variations was studied first, because we had some ideas why estimates of the FPR in these experiments could be too low. These ideas are described below, and are formalized in three hypotheses.

Except for $P_{sg}$ there are other factors that control $F_o$. If we want to know the effect of $P_{sg}$ alone on $F_o$ then we must check whether all other factors are constant. It is known that $F_o$ is also controlled by the laryngeal muscles. Baer [5] studied the influence of the laryngeal muscles on the FPR in an experiment in which the subject is pushed on the chest to increase $P_{sg}$. He found a consistent increase in the EMG activity of vocalis (VOC) and interarytenoid 30-40 ms after each push. Even for the fastest laryngeal muscles it takes about 15-20 ms before a change in the activity of a muscle is followed by a change in $F_o$ [9,10]. So the first 45-60 ms following a push the laryngeal muscles probably do not affect $F_o$. Baer calculated the FPR during the first 30 ms and found a value of 2-4 Hz/cmH2O in the chest register, a value that did not deviate from the values reported earlier by others. We did not re-examine the effect of the laryngeal muscles on the FPR.

The first hypothesis:

- a sudden rise in $P_{sg}$ is followed by a rise in $F_o$.

In most experiments either sub- or supraglottal pressure ($P_{sg}$) is measured and varied, while the other pressure signal ($P_{tg}$ resp. $P_{pg}$) is not measured. During sustained phonation of a vowel the impedance of the glottis is high but finite. A change of the pressure on either side of the glottis could leak through the glottis. If this would happen the change in transglottal pressure ($P_{tg}$) is smaller than the change in the measured pressure signal. Because it is really $P_t$ that controls $F_o$ [11], it is also the change in $P_t$ that has to be related to a change in $F_o$. The effect would be that the estimated FPR is smaller than the ratio between change in $F_o$ and $P_t$.

The second hypothesis:

- a change in $F_o$ lags a change in $P_{sg}$.

The scatter plots of $F_o$ versus $P_{sg}$ in Baer's article [5] exhibit hysteresis. The hysteresis is already visible during the first 45 ms, so before laryngeal muscle activity could influence $F_o$.

This could be an indication that the $F_o$ change lags the $P_{sg}$ change. During the sustained vowel the vibratory system is in a steady state. When $P_{sg}$ is changed it takes some time for the vocal folds to reach a new steady state. The time constant of this adaptation process depends on the total $P_{sg}$ change. Furthermore, this lag would only show up if the time constant of the $P_{sg}$ change is less than the time constant of the adaptation process. In speech the rate of $P_{sg}$ change during an utterance of about 1-8 cmH2O/s is probably slow enough for the vocal folds to adjust almost instantaneously to the new vibratory conditions. Both Ladefoged [6] and Baer [5] used short pushes to vary $P_{sg}$. During these pushes the estimated rate of $P_{sg}$ change is substantially larger than the aforementioned rate of $P_{sg}$ change in speech. If the changes in $F_o$ would lag the changes in $P_{sg}$ then the duration of their pulsatile $P_{sg}$ changes could be too short for the vocal folds to reach a new steady state. The result would be an underestimation of $dF_o/dP_{sg}$ and hence an underestimation of $dF_o/dP_{sg}$.

The third hypothesis:

- the FPR is different in $P_{sg}$ rising and lowering.

In utterances that exhibit declination both $F_o$ and $P_{sg}$ decrease during the course of the utterance. This is most clearly seen in declarative utterances with a single accent early in the utterance. In many languages the FPR is calculated for decreasing $F_o$ and $P_{sg}$. On the other hand, in experiments where $P_{sg}$ is changed by pushing on the chest the FPR is calculated for increasing $F_o$ and $P_{sg}$. Differences between $F_o$ rising and $F_o$ falling have been reported and Breckenridge [12] summarizes them by stating that "it has been found that falling tones are more common in the world's languages than rising tones, can be produced faster, and furthermore fall more than rising tones rise." Maybe the FPR is different for $P_{sg}$ rising and lowering, i.e. the ratio in lowering is higher.

In short, three hypotheses were postulated that could explain why the estimated FPR in experiments with induced pressure changes are too low: 1 a rise in $P_{sg}$ is followed by a rise in $P_{pg}$ 2 a change in $F_o$ lags the change in $P_{sg}$ 3 the FPR is different in $P_{sg}$ rising and lowering. These three hypotheses were tested with the data of an experiment.
METHOD
An experiment was carried out in which simultaneous recordings of acoustic signal, electrogastrogram (EGG), \( P_{sb} \), \( P_{sp} \) and stenohyoid (SH) were obtained while a subject sustained a vowel /a/ at a comfortable \( F_0 \) and intensity level. During phonation he was pushed on the chest to increase \( P_{sb} \), the chest was held down to keep \( P_{sb} \) high, and finally the chest was released again to lower \( P_{sb} \). In normal speech the fall of \( P_{sb} \) during an utterance generally varies from 2 to 12 cmH\(_2\)O (see section II). In earlier experiments the magnitude of the induced pressure changes was about 1-4 cmH\(_2\)O [5,7,8]. We tried to induce larger pressure variations. The \( P_{sb} \) changes were induced as fast as possible, in order to produce a large \( P_{sb} \) gradient.

All measured signals were stored on a 14-channel instrumentation recorder (TEAC XR-510). The signals are A/D-converted off-line at a 10 kHz sampling rate. \( F_0 \) was calculated from the EGG signal with a frame rate of 200 frames/s. The pressure signals were low-pass filtered and downsamplred to 200 Hz.

RESULTS AND DISCUSSION
The results of this experiment were used to test the three hypotheses. The hypotheses were tested in the same order as they are presented in the introduction above. In Figure 1 the \( F_0 \), \( P_{sb} \) and \( P_{sp} \) signals are shown for one of the pushes.

![Figure 1. \( F_0 \), \( P_{sb} \) and \( P_{sp} \) during a chest push.](image)

The fact that a change in \( P_{sb} \) is followed immediately by a change in \( F_0 \) indicates that there must be a direct relation between these two variables. It is observed that we succeeded fairly well in keeping \( P_{sb} \) high for some time. In all cases \( P_{sb} \) decreased during the time that the chest was held down. This could be caused by a partial release of the chest, an adjustment of the respiratory muscles, or it could be a by-product of the decreasing lung volume. In the example in Fig. 1 the stepwise increase in \( P_{sb} \) was 9.2 cmH\(_2\)O, while the \( P_{sb} \) sudden decrease was 7.0 cmH\(_2\)O. This means that we also succeeded in inducing pressure variations of substantial magnitude. Both the average rate of change and the maximum rate of change are about the same during rising and lowering (±25 cmH\(_2\)O/s for the average resp. ±55 cmH\(_2\)O/s for the maximum). This value is much larger than the rate of \( P_{sb} \) change during speech utterances, that is known to be in the range of 3-6 cmH\(_2\)O/s (see section II), and therefore the \( P_{sb} \) changes seem fast enough to test whether there is a lag between \( F_0 \) and \( P_{sb} \) changes. For three chest pushes the \( P_{sb} \) variation was as intended: the rise and fall are fast and large enough, and \( P_{sb} \) is kept high for some time. Particularly the data of these pushes are used to test the hypotheses. This is discussed below.

During \( P_{sb} \) rising and lowering no significant changes in \( P_{sp} \) were observed, as can be seen from the example in Fig. 1. This was the case for the three 'successful' pushes mentioned above, but also for all other pushes. A \( P_{sb} \) rise was never followed by a \( P_{sp} \) rise, so our first hypothesis was rejected.

![Figure 2. \( F_0 \) and \( P_{sb} \) during a chest push.](image)

The \( F_0 \) and \( P_{sb} \) signals of Figure 1 are plotted together in Figure 2. \( F_0 \) changes instantaneously with \( P_{sb} \), even if the total \( P_{sb} \) change is ±9 cmH\(_2\)O and if the rate of \( P_{sb} \) change is ±55 cmH\(_2\)O/s. A lag between \( F_0 \) and \( P_{sb} \) was not found. The voice source apparently is capable of adjusting very fast to changing phonatory conditions.

![Figure 3. \( F_0(\text{P}_{sb}) \) during \( P_{sb} \) rising (1) and lowering (2).](image)

A scatter plot of \( F_0 \) versus \( P_{sb} \) is shown in Figure 3. Shown are the data during \( P_{sb} \) rising (1) and lowering (2). One can see that the FPR is almost the same during rising and lowering. A substantial difference in the FPR during rising and lowering was not observed.

CONCLUSIONS
All three postulated hypotheses were falsified. At the moment there seems to be no reason to doubt the values of the FPR found in the experiments with induced pressure variations. Therefore the values obtained from measurements on normal speech have to be questioned.

II. THE FPR IN SPEECH
INTRODUCTION
There are two mutually exclusive explanations why estimates of the FPR in speech utterances showing \( F_0 \) declination are larger than estimates in experiments with induced pressure variations: the FPR is really larger in speech, or the estimates obtained from measurements in speech are wrong. The second explanation seemed more probable to us, so we first examined the methods that are used to calculate the FPR in the experiments on declination [1,2,3,4].

Usually the \( F_0 \) and \( P_{sb} \) values are taken at two instants, one near or at the beginning (T\(_i\)) and one near or at the end (T\(_f\)) of an utterance. An estimate of the FPR is then calculated with these values:

\[
\text{FPR} = \frac{[F_0(T_i) - F_0(T_f)]}{[P_{sb}(T_i) - P_{sb}(T_f)]}
\]

EUROSPEECH '89, Paris, France, September 1989 2426
In a plot of $F_o$ as a function of $T_r$ (2), and during intermediate period (.), plot of $F_o$ versus experiment. Shown are the physiological processes that influence $F_o$. This makes it hazardous to all other factors on $F_o$ is the same at those two instants. A coefficient between $F_o$ and rate production of a short and a long Dutch sentence. The short sentence-contour pairs (4 sentences x 3 intonation contours) was different intonation contours, i.e. a 'flat hat pattern' (FH), two 'pointed hats' (PH) and question intonation (Q). Each of the 12 sentence-contour pairs (4 sentences x 3 intonation contours) was repeated at least five times to make averaging possible.

In the second experiment recordings of the supraglottal pressure ($P_{SB}$) were also made, but activity of the CT was not recorded. Near the end of the experiment the subject was asked to produce an utterance spontaneously (SU). After he spoke this sentence, he was asked to repeat the same sentence 29 times.

Preprocessing of the data was done with the Haskins Laboratories EMG data processing system. The repetitions were time aligned using line-up points. A DTW algorithm was used to correct for the differences in the temporal structure between repetitions. Median values were then calculated for all variables. The exact procedure of data measurement and data processing is described in [11].

RESULTS AND DISCUSSION

The resulting signals were used to calculate the values given in Table 1. The values for the utterances in which all syllables are replaced by /v/ were deviating. In these sentences voicing starts well before the initial peak in $F_o$ and $P_{SB}$. As a result the $F_o$ and $P_{SB}$ values are small for the first voiced sample, and $dF_o$ and $dP_{SB}$ are small too. We could have chosen another instant ($T_r$) to measure $F_o$ and $P_{SB}$, but that is beyond the scope of this paper. The total fall in $P_{SB}$ varied between 4.0 and 11.9 cmH$_2$O, and the overall rate of $P_{SB}$ change varied between 1.7 and 8.1 cmH$_2$O/s.

<table>
<thead>
<tr>
<th></th>
<th>SHORT</th>
<th>/FI/</th>
<th>/V/</th>
<th>LONG</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>297</td>
<td>1222</td>
<td>12</td>
<td>535</td>
</tr>
<tr>
<td>$T$</td>
<td>2.3</td>
<td>1.4</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>$dF_o$</td>
<td>30</td>
<td>65</td>
<td>64</td>
<td>51</td>
</tr>
<tr>
<td>$dF_o'_{T}$</td>
<td>13</td>
<td>61</td>
<td>49</td>
<td>35</td>
</tr>
<tr>
<td>$dP_{SB}$</td>
<td>4.3</td>
<td>9.5</td>
<td>11.3</td>
<td>6.2</td>
</tr>
<tr>
<td>$dP_{SB}'_{T}$</td>
<td>1.9</td>
<td>6.8</td>
<td>8.1</td>
<td>4.8</td>
</tr>
<tr>
<td>$P_{SB}$ (FPR$^1$)</td>
<td>7.0</td>
<td>8.9</td>
<td>6.1</td>
<td>7.4</td>
</tr>
<tr>
<td>$P_{SB}$ (FPR$^2$)</td>
<td>6.9</td>
<td>4.0</td>
<td>6.0</td>
<td>0.2</td>
</tr>
<tr>
<td>$P_{SB}$ (FPR$^3$)</td>
<td>3.9</td>
<td>1.5</td>
<td>3.5</td>
<td>2.5</td>
</tr>
<tr>
<td>$dF_o$</td>
<td>8.7</td>
<td>7.8</td>
<td>10.9</td>
<td>10.0</td>
</tr>
<tr>
<td>$dP_{SB}$</td>
<td>10.9</td>
<td>10.0</td>
<td>8.0</td>
<td>8.7</td>
</tr>
</tbody>
</table>

The values of FPR$^1$ and FPR$^2$ for the questions are not relevant, because $F_o$ rises markedly near the end of these sentences. The other values of FPR$^1$ vary from 6.1 to 8.9 Hz/cmH$_2$O. This is in agreement with the results of previous studies [1,2,3,4], and therefore these sentences seem suitable to test our hypothesis.

The values of FPR$^2$ for non-questions always are smaller than the values of FPR$^1$, and vary between 4.0 and 6.9 Hz/cmH$_2$O. But one has to be careful in interpreting these values. The value of a regression coefficient is dependent on the value of the correlation coefficient, and therefore a smaller correlation coefficient would result in a smaller regression coefficient. In any case, the values of FPR$^2$ are still higher than the FPR values obtained in experiments with artificially induced $P_{SB}$ variations.

The results for one sentence (long-FH) are shown in Figure 5. In most sentences CT and VOC were especially active during the first syllable, and their activity was suppressed at the end. This effect can also be often be observed in the data of previous experiments on declination in which muscle activity was measured [1,2,3,4]. The peak activity of these $F_o$ raising muscles is much
larger during a stressed syllable at the beginning than during a stressed syllable at the end. And if the first syllable is not stressed, then CT and VOC still show increased activity. On the average the FO raising muscles CT and VOC are more active at the beginning than at the end of utterances.

It is often observed that the SH is especially active just before phonation [1,13,14], and it is assumed that the SH helps in preparing the larynx for the 'speech mode.' This was also observed in some of the utterances of this experiment. Usually SH activity has dropped to its base level when phonation starts. At the end of utterances FO often falls abruptly (the so called final fall), and often this is accompanied by a rise of SH (and a lowering of the larynx). This is observed in the data of the present experiments, but also in the data of previous experiments [1,2,3,4]. On the average the FO lowering muscle SH is more active at the end than at the beginning of utterances.

Thus it seems that the laryngeal muscles participate in the declination of FO so part of the decline in FO is due to the activity of the laryngeal muscles. If we want to calculate the FPR we first have to correct FO for these influences. This is done by calculating the regression equation between FO and SH and VOC for the average signals of the spontaneous utterance, and the regression equation between FO and SH and CT for the other 12 sentences. The value of FPR3 is then calculated. Except for the long-FH-type the values vary between 1.5 and 3.3. Again we want to stress that regression coefficients do depend on the correlation between the variables. For instance in the long-FH-type the correlation was extremely low causing the value of FPR3 to be very low. Still, if we compare the values of FPR3 with those of FPR2 we see that correction for the influence of two important laryngeal muscles resulted in a lowering of the estimate of the FPR in all non-questions.

For the questions the FO rise at the end is mainly controlled by the combined activity of CT and VOC. The value of FPR3 is corrected for this increase in CT activity, and therefore the value of FPR3 is also relevant for questions. The values thus obtained are in the same range as the FPR2 values for declarative utterances.

CONCLUSIONS
The data obtained in the two experiments described above do support our hypothesis that the FPR is the same in speech and in experiments with induced pressure variations. Our data, and data of previous experiments on declination, suggest that laryngeal muscles participate in the FO decline during an utterance.

ACKNOWLEDGEMENTS
This research was supported by the foundation for language research, which is funded by the Netherlands Organization for the Advancement of Scientific Research N.W.O. Special thanks are due to Haskins Laboratories were one of the experiments was carried out; to dr. Thomas Baer who helped organizing and running the experiment at Haskins; to dr. Hiroshi Muta who inserted the EMG electrodes and the subglottal pressure sensor in the experiment at Haskins; and to dr. Philip Blok who inserted the EMG electrodes and the pressure catheter in the other two experiments.

REFERENCES