

# A study of the two-mass model of the vocal cords from a fluid dynamical point of view

**Citation for published version (APA):**

Pelorsen, X. (1992). *A study of the two-mass model of the vocal cords from a fluid dynamical point of view*. (IPO rapport; Vol. 873). Instituut voor Perceptie Onderzoek (IPO).

**Document status and date:**

Published: 28/10/1992

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

Rapport no. 873

A study of the two-mass model  
of the vocal cords from a  
fluid dynamical point of view

X. Pelorson

**A STUDY OF THE TWO-MASS MODEL OF  
THE VOCAL CORDS FROM A FLUID  
DYNAMICAL POINT OF VIEW.**

**X. PELORSON**

## Table of contents.

### Introduction.

#### I. Characteristics of the vocal folds.

I.1. The sub-glottic part.

I.2. The larynx and the vocal folds.

I.3. Voicing.

I.3.1. Physiology of phonation.

I.3.2. Vibratory pattern of the vocal folds, examples for different kinds of phonation.

I.3.3. Characteristics of the movement.

#### II. A numerical model for phonation.

II.1. Overview of existing models.

II.2. The two mass model.

II.2.1. Description.

II.2.2. Mechanical equations.

II.2.3. Hydrodynamical equations

II.3. Results.

II.3.1. Choice of the parameters.

II.3.2. Examples of results.

II.3.3. Speech production, a model for the vocal tract.

II.4. Conclusion.

#### III Modifications and discussion.

III.1. The collision problem.

III.1.1. Problems.

III.1.2. Spring models.

III.1.2.1. Damped-spring model.

III.1.2.2. Introduction of additional displacement

III.2. A modified description of the flow within the two-mass model.

III.2.1. Flow at the entrance of the glottis (first channel).

III.2.2. Flow through the first channel of the glottis.

III.2.3. Flow in the second channel.

III.2.4. Flow in the larynx.

III.3. Introduction of flow separation in the two-mass model.

III.3.1. Discussion about the relative importance of separation process.

III.3.2. Unsteady separation.

III.3.3. Conclusion and further research.

### Conclusion.

## Acknowledgements

This work was supported by the Ambassade de France in Den Haag, the Institute for Perception Research (IPO) and the Technical University of Eindhoven (TUE).

First I would like to thank Pr. Bouma, Pr. Houtsma and Dr. Collier for welcoming me at IPO for five months. I wish also to thanks all the people who helped to make this staying so pleasant and interesting. Among many others thanks to B Eggen, L. Ten Bosch, L. Vogten, M. Swerts, K. van Deemter...

Finally I thank also Dr Hirschberg from TUE and Dr. Badin from the Institut de la Communication Parlée for their support during this work.

## **Units.**

International Standard Organization (ISO) recommends the use of MKSA system of units (Système International). However in speech research, traditionally, the CGS system is more likely used. For this reason and in order to simplify comparison between our results and literature, we will use this system of units. In particular, the pressures will be expressed in centimetre of water (cm H<sub>2</sub>O) instead of Pascal (1 cm H<sub>2</sub>O = 980 Pa) while the forces will be expressed in Dyne instead of Newton (1 Dyne =  $10^{-5}$  N).

## INTRODUCTION

The production of voiced sounds, also called phonation, is due to the vocal cord movements, converting subglottic pressure (pressure from the lungs) into acoustic pressure. The vocal fold oscillatory movement is essential not only because its fundamental frequency is highly correlated with perceived pitch, but also because other parameters directly related with the vocal cords movement are important for speaker recognition [1]. Lastly, the typical computer accent of synthesized speech seems to be caused to a considerable extent by an inaccurate estimate of the flow in the glottis.

For these reasons an increasing effort has been made to modelize and to predict a more realistic behaviour of the vocal cord movements. A reasonable compromise between theoretical and computational complexity and efficiency seems to be the two-mass model [2], [3]. Indeed this model was found enough accurate to describe specific physiological behaviour and to produce an excitation for speech-like signals. On the other hand, it is a model simple enough that the influence of all the parameters that are taken in account can be evaluated and that new effects can easily be added.

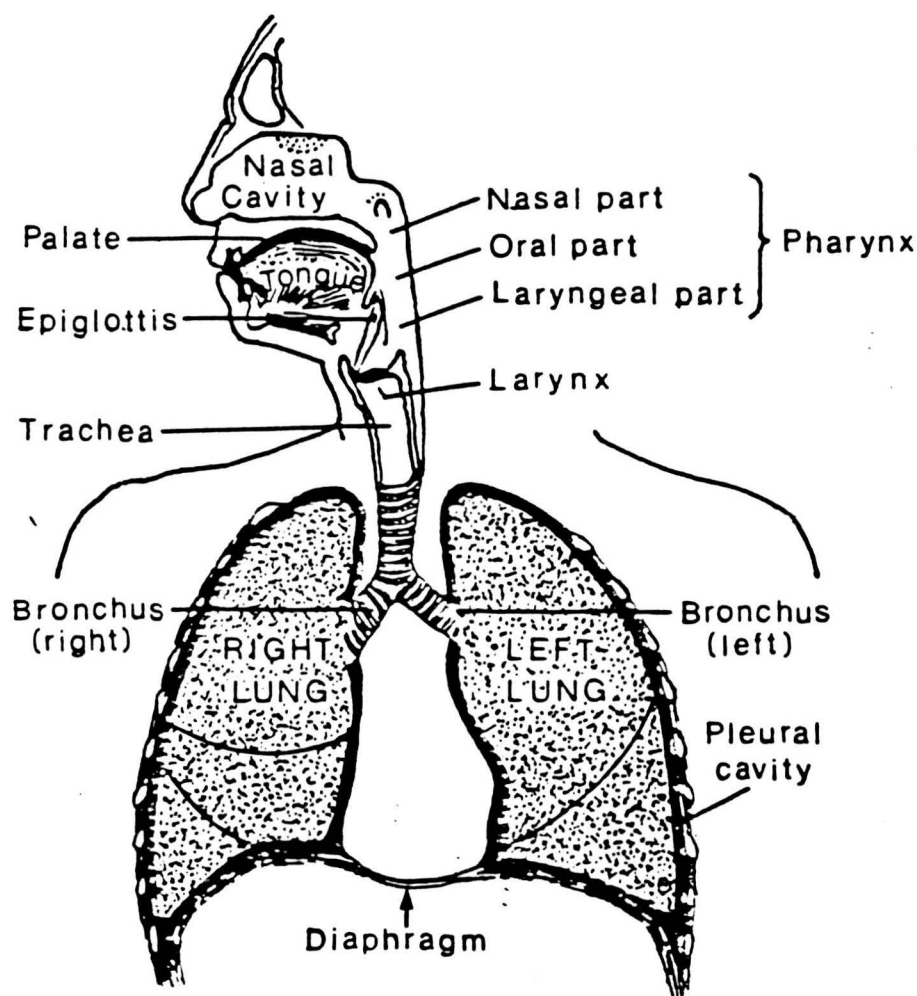
However, more and more studies pointed out some limitations of this simple model. Some of them are based on measurements upon models of the vocal folds and exhibit disagreements with simulation results [4]. Others are much more theoretical and systematic [5]. We believe that it is not necessary to increase the complexity of the model nor to reconsider all the basic assumptions or theory in order to improve the quality of synthesized speech. We expect that the introduction of fluid dynamical effects, in addition to the modification of the two-mass model, will be sufficient to explain and take into account most of the problems expressed.

This present report should be understood as a starting point for further research. Our purpose is first to study basic literature about speech and more precisely about voiced speech. We then turn to the construction and the validation of a two-mass model, similar to the original one, presented by Ishizaka and Flanagan in 1972 [2]. A few results will be presented and compared with real speech effects. The last part of this report is dedicated to a discussion about the limitations of the model. Some modifications of the two-mass model will be presented and their possible effects will be discussed. Lastly, we will present the aim and purpose of our further research.

## I. Characteristics of the vocal folds.

The speech organs, the structures used for speech production, can be divided into three parts (figure 1) :

- The respiratory system supplies air flow to the larynx and the vocal tract.
- The larynx, and more precisely the vibration of the vocal folds, is the source for voiced sounds. The conversion process of air pressure into sound in the larynx is called phonation.
- The vocal tract, which is the place for vowel articulation, is also responsible for production of some consonant sounds which can be voiced or unvoiced (fricatives, affricatives and stops).

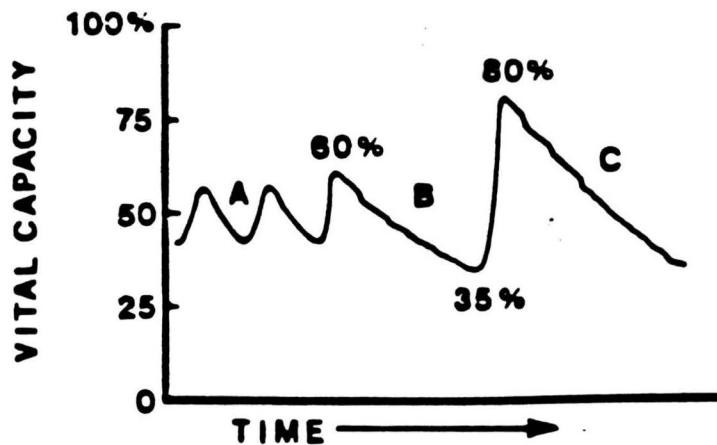


**Figure 1:** Parts of the speech organs.(from [6])



## I.1. The sub-glottic part.

Although it is also possible during inspiration [7], the most important part of speech is done during the expiration phase. In quiet expiration, the air exchange is small (about 0.5 l) and its cycle duration is short (2 - 5 s). For speech, the expiration is a much longer process (up to 40 s) and the volume of air exhausted is considerably higher, as shown figure 2.

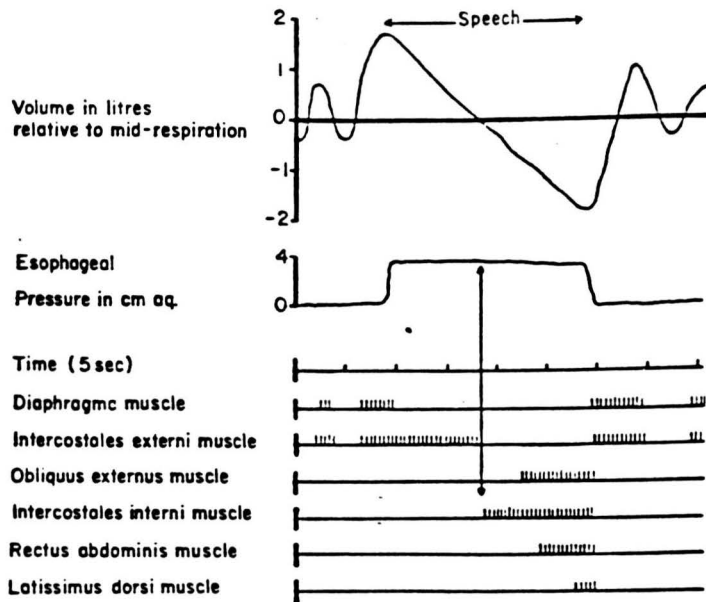


**Figure 2** : Relative lung volume changes during :

- A : tidal breathing,
- B : conversational speech,
- C : loud speech.

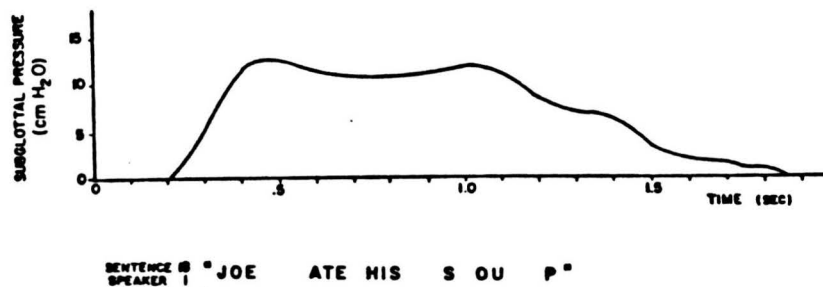
(from [6]).

For our purpose it is more convenient to consider the respiratory system as a source of pressure. Figure 3 presents the subglottic pressure measured in a subject counting from 1 to 32 at a conversational loudness.



**Figure 3:** Relative lung volume and subglottic pressure during an utterance spoken after a deep breath. The lower curves are showing the activity of several respiratory muscles. (from [8]).

As shown in this figure, speech involves several muscles to maintain the subglottic pressure relatively constant. In conversational speech, however, the subglottic pressure presents time variations related to the differences of intensity (emphasis) or pitch (figure 4).



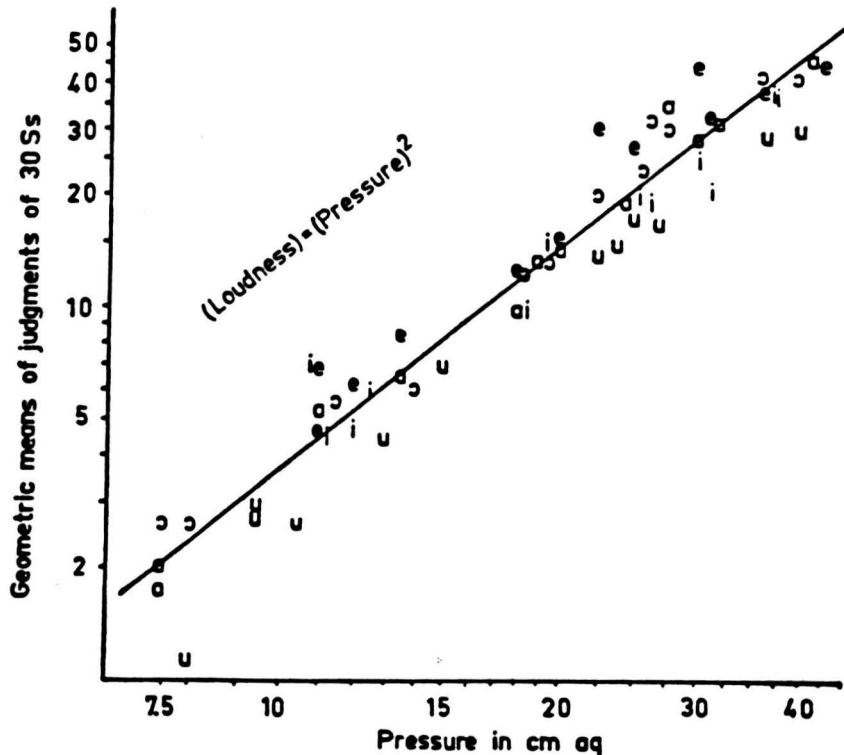
**Figure 4:** Subglottal pressure during the pronunciation of the sentence "Joe ate his soup". (from [9])

Although difficult to realise, and sometimes harmful, several quantitative measurements of subglottic pressure during speech have been made using different techniques [10] :

- insertion of a catheter with pressure transducers through the nasal passage and the glottis [11],[12],
- direct measurements by means of tracheotomy,
- Insertion of a small latex balloon in the oesophagus [13] and estimation of the subglottal pressure.

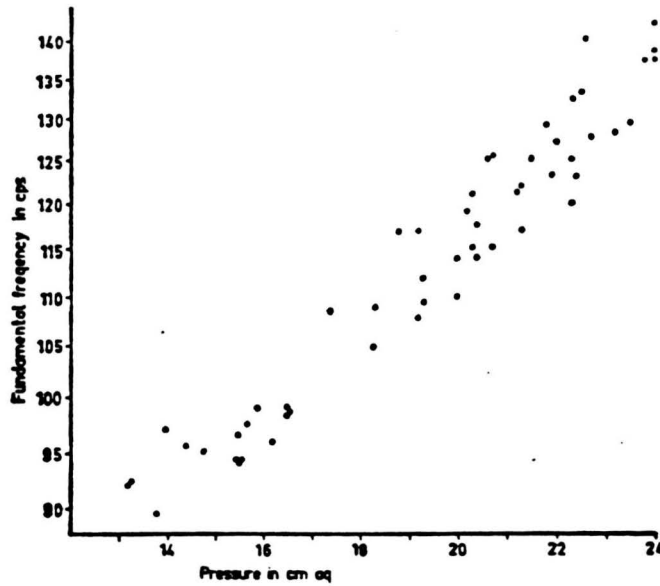
Results obtained by these techniques showed that the range of pressure comes from 3 cm H<sub>2</sub>O (very weak speech) up to 20 cm H<sub>2</sub>O (loud speech). For normal conversational speech (chest register, output level of approximately 60 dB) the subglottal pressure is of about 6 to 10 cm H<sub>2</sub>O.

From the point of view of speech perception, the subglottic pressure is of course related with speech loudness as shown figure 5.



**Figure 5** : Correlation between sub-glottic pressure and perceived Loudness. (from [14])

In addition, other studies [7], [15] also pointed out a strong correlation between subglottic pressure and pitch. Figure 6 is an example of such result exhibiting a typical slope of 4 to 5 Hz / cm H<sub>2</sub>O.

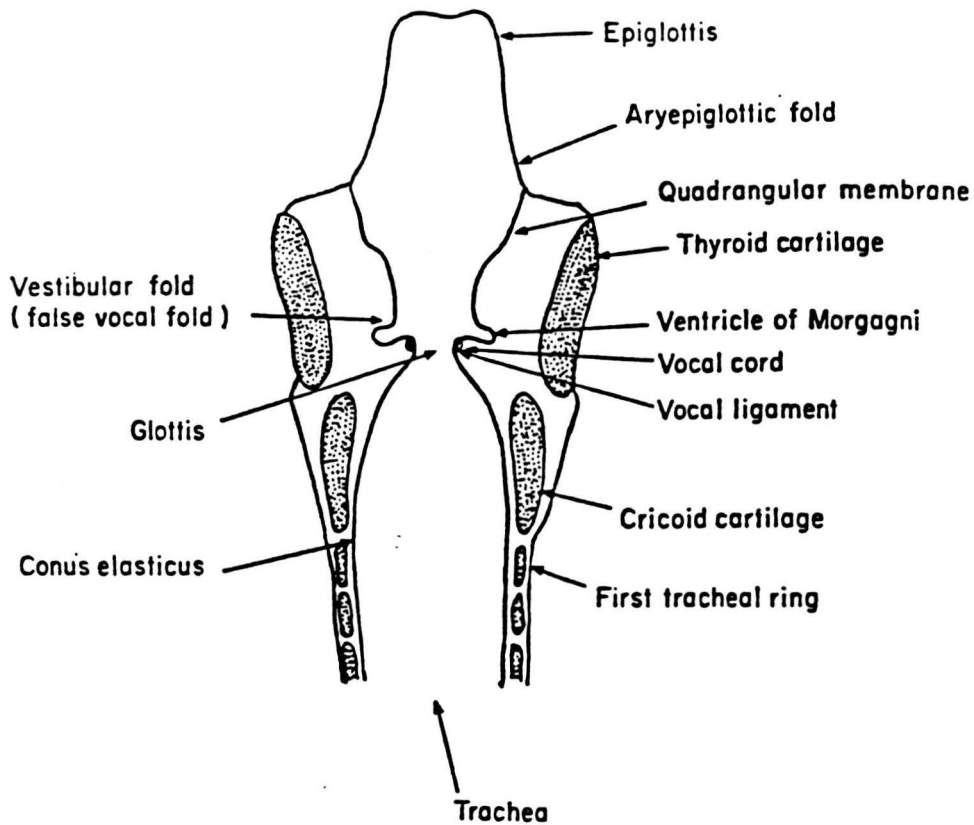


**Figure 6:** Correlation between Subglottic pressure and fundamental frequency.

## I.2.The larynx and the vocal folds.

In addition to phonation, the larynx has other important biological functions such as preventing foreign substances from entering the lungs, aiding for swallowing, controlling the flow of air into and out of the lungs . . .

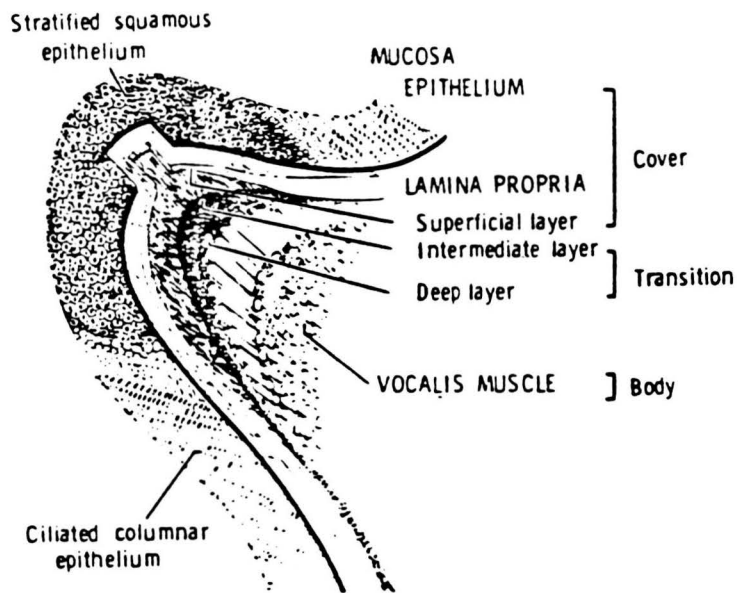
Schematically, the larynx can be represented as a tube composed of nine cartilages connected by ligaments and connecting membranes and covered by a mucous membrane. Figure 7 is an usual representation of the human larynx.



**Figure 7 :** A representation of the human larynx.(from [6])

The structure including the vocal ligament (vocalis), the muscles that are attached to the ligament and the internal part of the thyroarytenoid muscles and the mucous membrane is called the true vocal folds or vocal cords. From a clinical point of view, we should also include the arytenoid cartilage although it has no direct function in phonation. The glottis is the narrow slit between the vocal folds.

Thus, the vocal fold is not a uniform structure, it consists in fact of multiple layers [16], [17], each having its own mechanical behaviour. This layered structure varies along the length of the vocal fold as can be seen in figure 8.

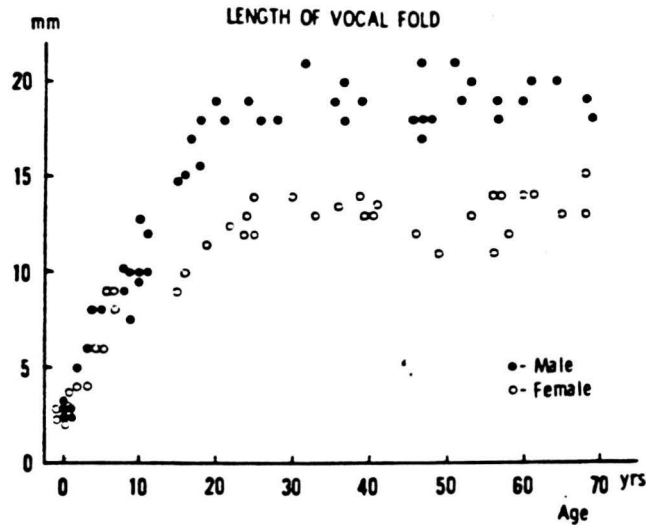


**Figure 8** : The layered structure of the vocal folds (from [17]).

In the following sections we are going to present a few results about the structure of human vocal folds. Of course, since voicing varies a lot with age, especially at puberty for men, and with sex, the characteristics of the vocal cords are also a function of age and sex. Several studies have been made for deriving quantitative measures of relevant characteristics. Most of them are based upon X-rays measurements, High Speed or Stroboscopic Movies or direct measures upon excised larynges.

### **Length of the vocal folds.**

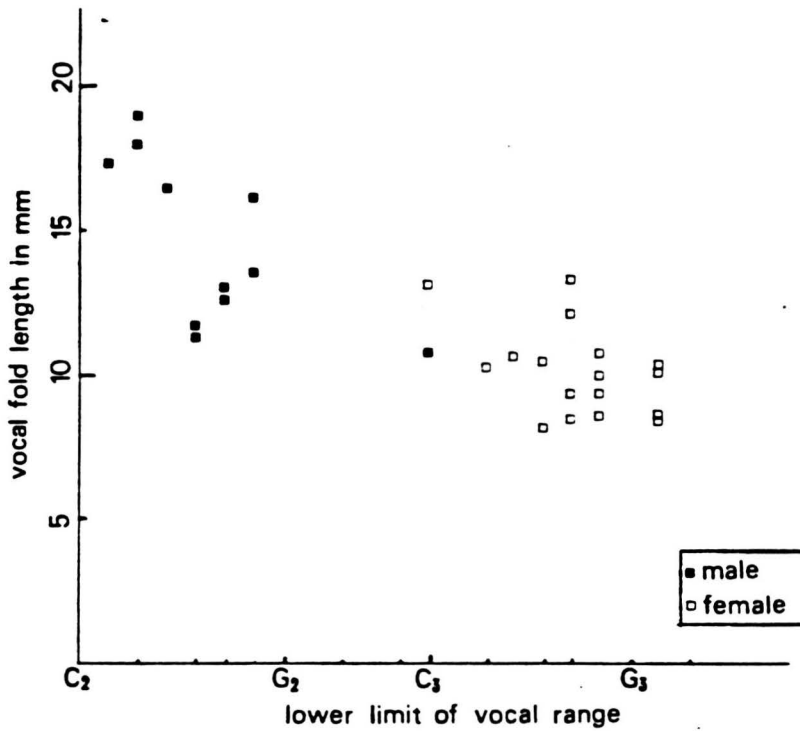
As previously stated, the length of the vocal folds varies strongly with age [17]. Figure 9 shows the length of the vocal folds including the cartilaginous part as a function of age.



**Figure 9:** Total length of the vocal fold measured on 48 males and 40 females as a function of age. (from [17])

As it can be clearly seen of this figure the entire length of the vocal cords is an increasing function of age up to about 20 years of age. In adults (i.e older than 20), the length is more or less constant and values 17 to 21 mm in males and 11 to 15 mm in females. Since the cartilaginous part of the vocal folds is not active during phonation, we rather consider as relevant the length of the membranous portion of the vocal fold (14 to 18 mm in males, 8.5 to 12 mm in females).

A consistent correlation can be found between vocal fold length and voice register. As an example we present on figure 10 the stereoscopic measures derived by Sawashima et al.[18] showing the relationship between vocal fold length, measured during inspiration, and the lower limit of the vocal range.



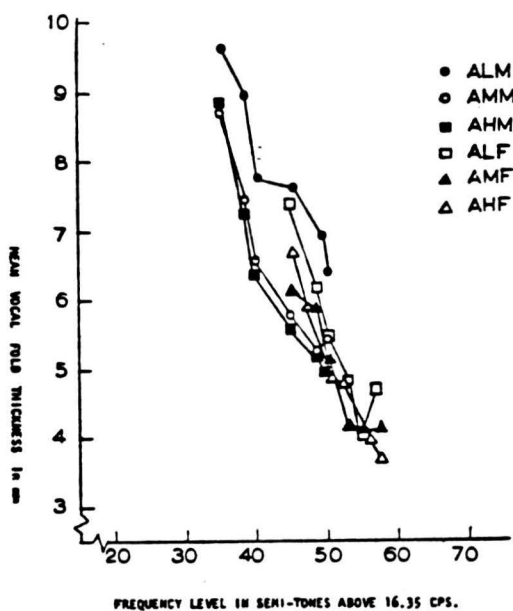
**Figure 10** : Relationship between the vocal fold length and the lower limit of the vocal range ( from [18]). Musical scale (logarithmic) is used as abscissa (C<sub>2</sub> = 130.8 Hz, G<sub>3</sub> = 392 Hz).

### Thickness of the vocal folds.

The thickness of the vocal folds also depend strongly upon the voice register and the sound to be produced [19],[20]. The overall thickness varies from about 5 to 7 mm for females and 6 to 9 mm for males.



The thickness of the vocal folds is negatively correlated with the fundamental frequency as shown figure 11.



**Figure 11** : Relationship between the thickness of the vocal folds and the fundamental frequency (from [19]).

### **I.3. Voicing.**

#### **I.3.1. Physiology of phonation.**

Qualitative and quantitative studies of the phonation process are derived from several different techniques :

- Ultra-speed cinematographic observations.
- Stroboscopic cinematography.
- Ultrasonic pulse trains.
- X ray stroboscopy,
- Inverse filtering.

Other promising techniques such as Ultrasound Echo System [21] are still under development or improvement.

Phonation is coordinated with respiration : before speaking, one must take his breath. Then, the glottis opens quickly before the thorax expands. The vocal folds are then adducted and are strengthened.

When air flow first enters the glottis, an aerostatic force caused by the subglottic pressure is exerted against the vocal cords. This force acts to displace the vocal cords outwards their rest position. As the vocal folds are opening two forces are opposing this movement :

- The first one is due to the elasticity of the tissues and tends to restore the vocal folds to their rest position.
- The second one is the Bernoulli force and can be explained as follows :

As the velocity of the air passing through the glottis increases the local pressure decreases due to the principle of the conservation of energy. Due to this pressure difference the vocal cords tend to be sucked together (Bernoulli force). For a more formal description see appendix A.

The opening phase proceeds until the aerostatic force is counterbalanced by the sum of the Bernoulli and restoring forces. The vocal folds start then to move inward and tend to close.

As the glottis is made narrower the velocity (and thus the Bernoulli force) increases up to a certain point for which the friction forces (proportional to the inverse third power of the aperture) become predominant.

When the vocal folds close the Bernoulli force ceases. The restoring force and the aerostatic force then reopens the glottis and the cycle is repeated.

To summarize, phonation occurs when :

- The pressure difference across the glottis is high enough to move the vocal cords outward from their rest position,
- The vocal folds are adduced and made elastic enough so that vibrations can take place.

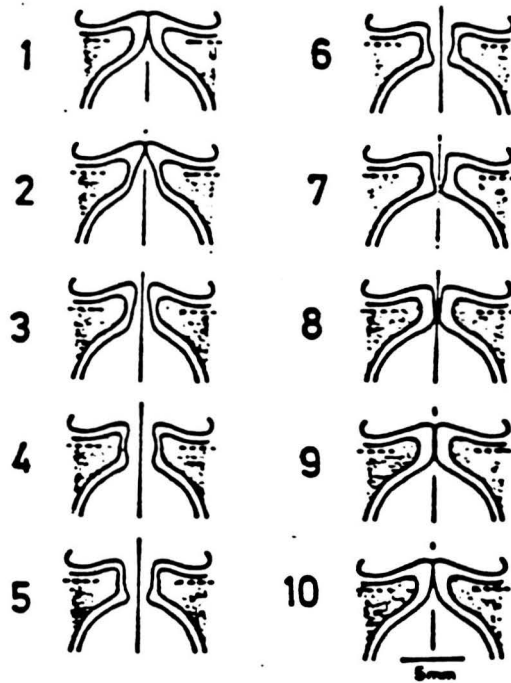
### **I.3.2. Vibratory pattern of the vocal folds, examples for different kinds of phonation.**

#### **Normal voice (chest register).**

When initial conditions are satisfied, the vocal folds start to vibrate and collide quasi periodically with a fundamental frequency  $F_0$ . In normal voicing (chest register) one cycle of vibration of the vocal fold consists of three phases :

- Opening phase.
- Closing phase
- Closed phase

The vibratory pattern of the vocal folds, especially at closure, often exhibits a different behaviour of the lower and the upper part (lip) of the vocal folds. Schematically, the vocal cords are opening by their upper part and closing by the anterior one (figure 12).



**Figure 12** : Vibratory pattern of the vocal fold during one cycle of vibration.(from [22])

### **Breathy phonation (murmur).**

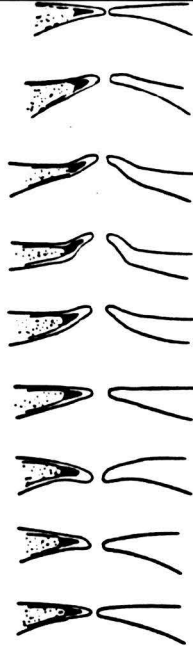
In this kind of phonation, only one part of the vocal folds is vibrating while the other part is kept open allowing air flow to pass continuously. Thus, a typical glottal airflow exhibits periodic pulses in addition to noise produced by the leak in the glottis.

### **Creaky Voice**

In creaky voicing, the vocal folds are closely adduced and only a small part of the vocal cords is allowed to vibrate. The fundamental frequency of such sounds is low and irregular (within the range 3 - 50 Hz).

### **Falsetto voice (second register).**

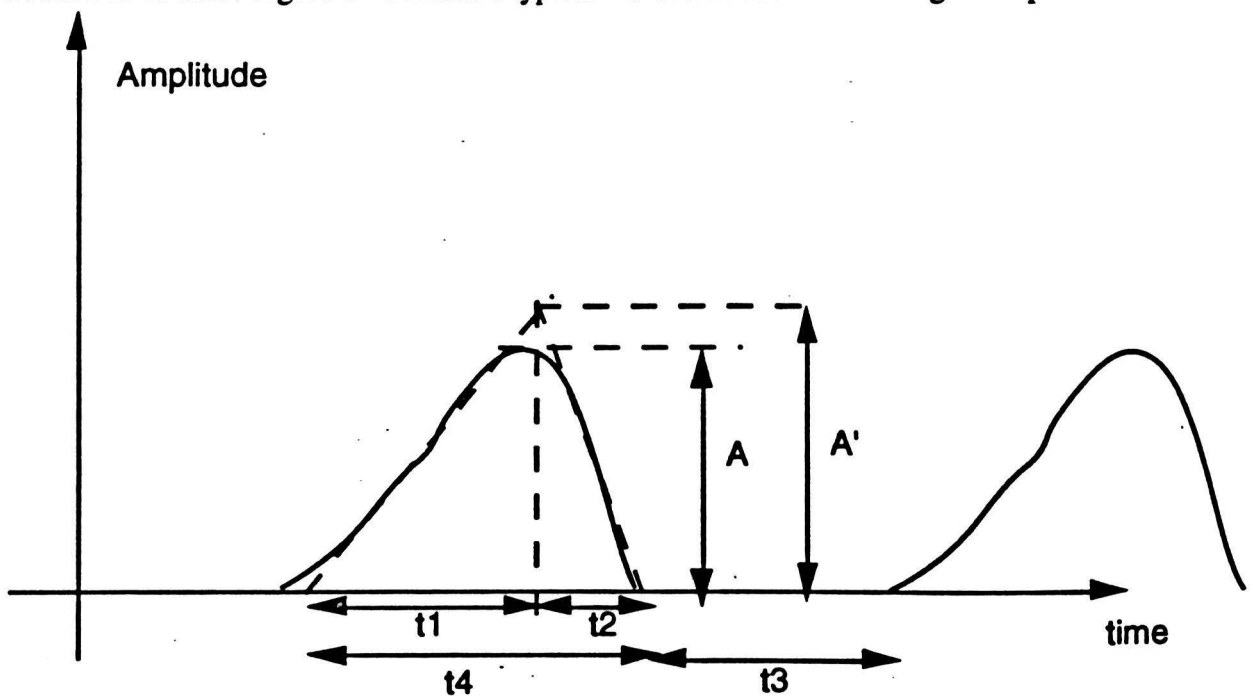
Falsetto voice is a type of phonation in which the vocal folds are made thin and do not present closed phase. Figure 13 presents a sketch of the vibratory pattern of the vocal folds.



**Figure 13** : Vibratory pattern of the vocal folds during falsetto register speech.(from [22])

### I.3.3. Characteristics of the movement.

In addition to the fundamental frequency  $F_0$ , other parameters are sometimes used to describe the vibratory characteristics of the vocal cords. Usually, these parameters are calculated by means of the glottal pulses that is the waveform of the volume velocity of air as a function of time. Figure 14 exhibits a typical waveform observed during chest phonation.



**Figure 14** : Glottal pulses and parameters of the waveform.

The most commonly used parameters are [23],[24],[25] :

- The speed quotient also called slope quotient or dissymmetry ratio :  $Sq = t1/t2$  expresses the dissymmetry of the glottal waveform,
- the open quotient  $OQ = t4/(t3+t4)$  is a measure of the time the glottis is opened,
- the peakedness (peak sharpness)  $P = A / A'$ ,
- The closed-open ratio  $C = t3 / t4$  is a comparison between the opening time and the closing time, and can be expressed as a function of OQ :  $C = 1/OQ - 1$ .
- The closed-speed ratio  $Cs = t2/(t3+t4)$ .

The following values were measured on natural speech and are generally accepted [23]

|          | Minimum value | Maximum value |
|----------|---------------|---------------|
| F0 males | 60 Hz         | 200 Hz        |
| females  | 130 Hz        | 400 Hz        |
| Sq       | 1.1           | 3.5           |
| Oq       | 0.4           | 0.83          |
| P        | 0.7           | 0.9           |
| C        | 0.2           | 1.5           |

**Table 1** : Range of values for some parameters of the glottal waveform.

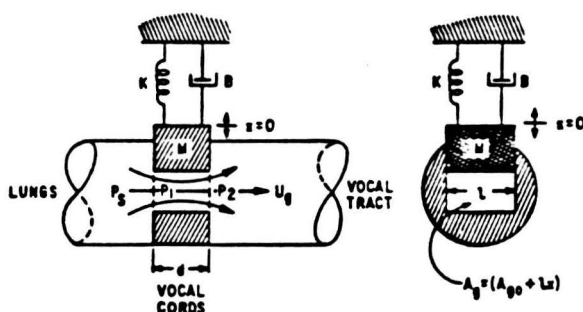
Perceptually these parameters can be related to the spectral distribution of speech [25], [26].

## II. A numerical model for phonation.

### II.1. Overview of existing models.

#### - The one-mass model [27] : 1 degree of freedom.

This model constitutes probably the most simple description of the vocal folds. Indeed this idea is not new since similar descriptions were reported in the very beginning of the century [28]. In this model, the vocal cords are approximated by a pair of opposing masses and springs as shown figure 15.

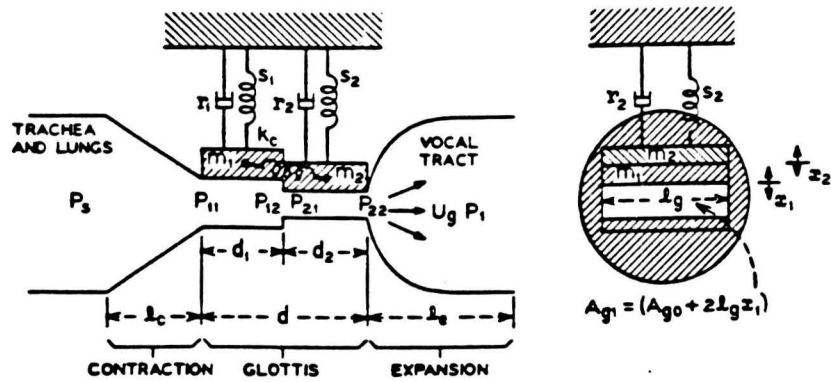


**Figure 15** : One mass model of the vocal cords [27].

This model however is no longer used since it was found unable to produce self sustained oscillations (and to produce sound) for realistic conditions.

#### - The two-mass model [2] : 2 degrees of freedom.

To remedy the limitations of the one-mass model, Ishizaka et al. proposed a more complex description as shown figure 16.



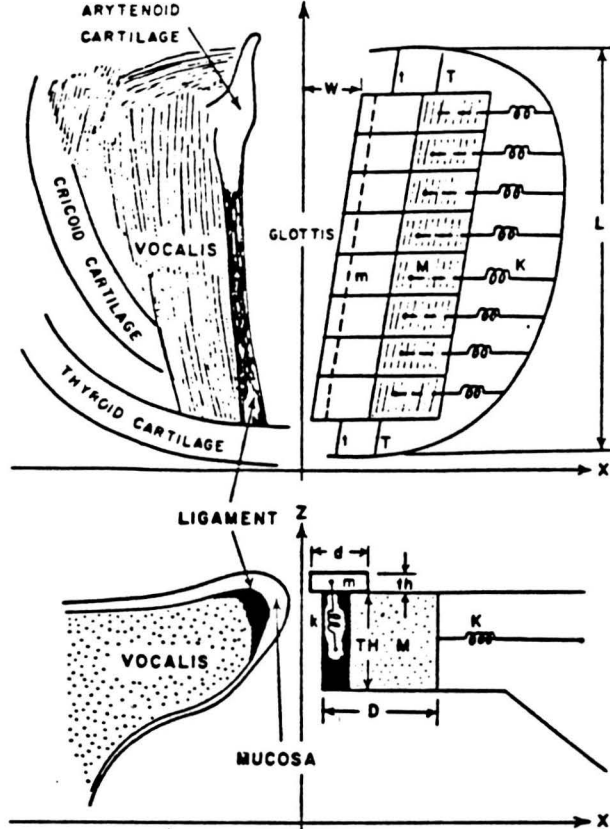
**Figure 16** : Two-mass model of the vocal cords (after [2]).

The second mass-spring system introduces an additional degree of freedom. As a consequence, self-sustained oscillation of the model were found possible for mostly all typical speech conditions.

**- Multi-mass models : more than 2 degrees of freedom.**

Originally proposed by Titze [29] this model has an increasing complexity. With these models, the goal is not only to produce sounds but also to capture the fine structure of the vocal folds (in order to study pathological behaviours for instance [30],[31]). In this sense, the 16-mass model of Titze presented figure 17 is a discretisation of the structure of the vocal folds including the layered structure.

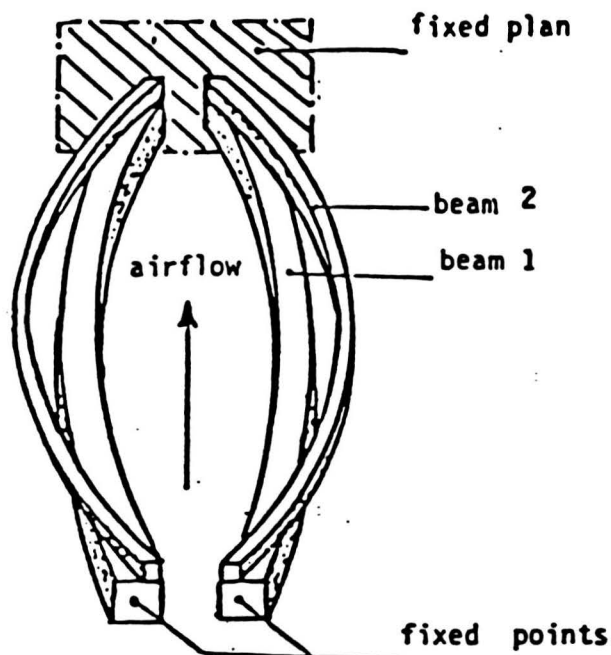




**Figure 17** : 16-mass model for the vocal folds [30].

**- Beam-model [32] : 2 and infinite degrees of freedom.**

In this model, the vocal folds are represented by a pair of beams articulated at their extremities (figure 18).



**Figure 18** : Beam-model for the vocal folds (from [32]).

As one can see, this model does not constitute a continuous model since there is still a discontinuity between two sections.

Interesting results have been obtained with such model. Unfortunately, due to the high complexity of both mechanical and hydromechanical behaviours (in addition probably with computation difficulties), this model is still to be improved.

The models listed here are of course not the only ones that can be found in the literature. We could for instance add the collapsing tube model [33],[34] which is an analytical model based on a linear approximation allowing only to describe the onset of phonation. As one can see from this brief description, the two-mass model seems to be, so far, the most simple efficient computer model. For this reason we chose it as a tool for further discussion. The following part describes the two-mass model we constructed, which is formally similar to the original formulation of Ishizaka and Flanagan [2].

## II.2. The two mass model.

### II.2.1. Description.

Figure 19 presents the original model derived by Ishizaka and Flanagan [2]. Each vocal cord is approximated by two stiffness-coupled masses. Since the vocal cords are assumed to be symmetric it is only necessary to study the behaviour of one system.

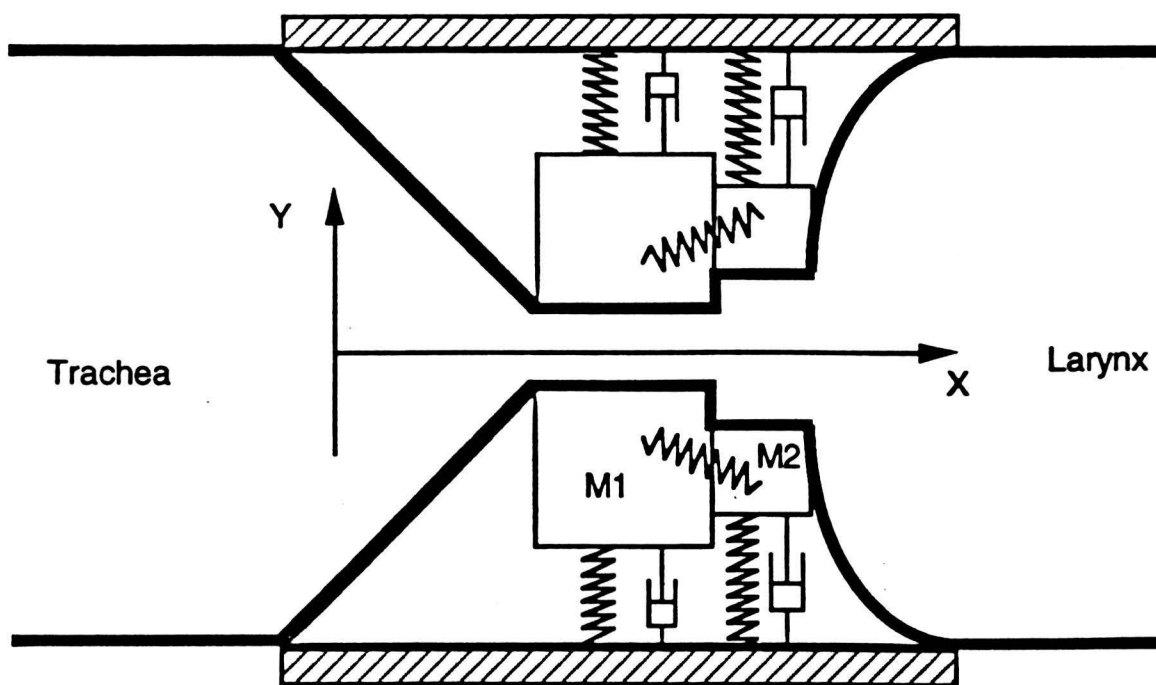


Figure 18 : Two mass approximation of the vocal cords.

Description for the behaviour of the model requires two sets of equations. The first set describes the oscillatory movement of the masses submitted to an external hydrodynamic force. The second set describes the fluid behaviour within the channel formed by the masses.

### II.2.2. Mechanical equations.

Let  $y_1(t)$  and  $y_2(t)$  be the displacement of mass 1 and 2 from their rest positions  $y_{01}$  and  $y_{02}$ . For a given time  $t$ , the coordinate of mass number  $i$  is then  $y_{0i}+y_i(t)$ .

The general law of dynamics can be expressed as :

$$M_i \vec{\gamma}_i = \vec{F}_i + \vec{F}_{s_i} + \vec{F}_{ij} \quad i = 1, 2 \quad j = 2, 1 \quad (1)$$

With :

- $\vec{F}_i$  : hydrodynamic force imposed by the flow,
- $\vec{F}_{s_i}$  : restoring force imposed by the spring  $i$  on mass  $i$ ,
- $\vec{F}_{ij}$  : force exercised by mass  $j$  on mass  $i$  due to the coupling,
- $\vec{\gamma}$  : is the acceleration vector.

Since we assume all the forces to be directed in the (Oy) direction, equation (1) can be expressed as :

$$M_i \frac{d^2 y_i(t)}{dt^2} = F_i + F_{s_i} + F_{ij} \quad i = 1, 2 \quad j = 2, 1 \quad (2)$$

- Expression of  $F_{s_i}$  :

This force is a representation of the tension of the vocal cords during phonation. In a simple spring-mass model,  $F_{s_i}$  can be expressed as being the sum of two terms : the first one expresses a restoring force, the second one the losses due to viscous damping. Ishizaka and al. [36] did measurements of the stiffness of fresh excised human vocal folds. They found a non-linear relationship between the deflection from rest position and the force exercised. According to their results, the restoring force can be approximately expressed as :

$$F_R = - k_i \cdot y_i(t) \cdot (1 + \eta_{ki} \cdot y_i^2(t)) \quad i = 1, 2 \quad (3)$$

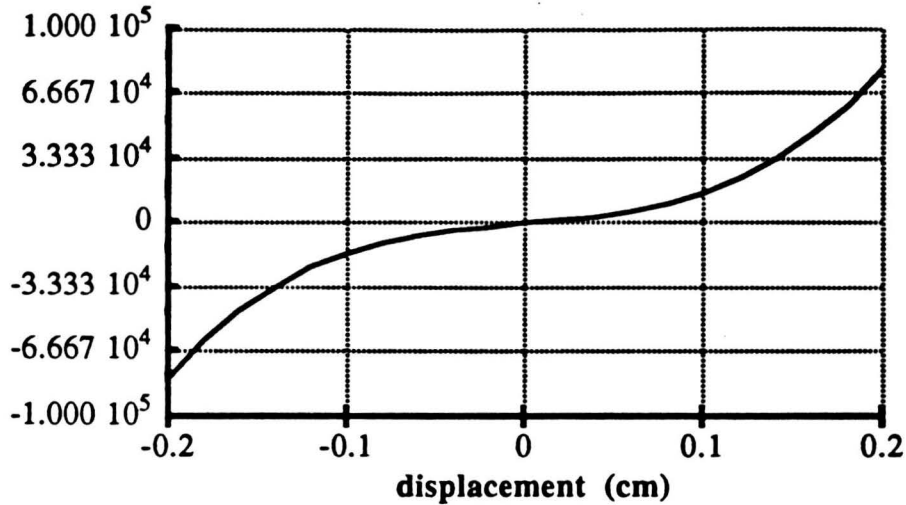
where :

$k_i$  : is the linear stiffness of the equivalent spring,

$\eta_{ki}$  : is a second order non-linear coefficient

Figure 19 is a representation of the restoring force as a function of the displacement  $y_i(t)$ .

**Restoring force (dynes)**



**Figure 19** : Non-linear restoring force as a function of displacement. Calculation made for  $k_1 = 80 \text{ kDynes.cm-1}$ ,  $\eta_{k1} = 100$ .

The viscous losses can be simply described by an equivalent force  $F_v$  :

$$F_v = -r_i \cdot \frac{dy_i(t)}{dt} \quad i = 1,2 \quad (4)$$

Thus,  $F_{si}$  can be expressed as :

$$F_{si} = -k_i \cdot y_i(t) \cdot (1 + \eta_{ki} \cdot y_i^2(t)) - r_i \cdot \frac{dy_i(t)}{dt} \quad i = 1,2 \quad (5)$$

### - Expression of $F_{ij}$

The coupling force  $F_{ij}$  expresses the stiffness of the vocal cords in the lateral direction. For this purpose a simple linear spring model is often used. Thus,  $F_{ij}$  can be expressed as :

$$F_{ij} = -k_{ij} \cdot (y_i(t) - y_j(t)) \quad i = 1,2 \quad j = 2,1 \quad (6)$$

### II.2.3. Hydrodynamic equations

In order to derive the expression of the driving force  $F_i$ , we need to make a complete description of the flow through the glottis. In the following section we will only make a quick presentation of the original theory proposed by Ishizaka et al. [37]. A more detailed presentation and discussion will be made in the third part of this report.

#### Basic assumptions :

Although very often they are not explicitly expressed in the literature, four basic assumptions are made :

A1)- Since the width of the glottis is small compared to its length (0.3 cm at the most compared to 1.4 cm in our model), we can make the assumption for the flow to occur between two parallel infinite plates.

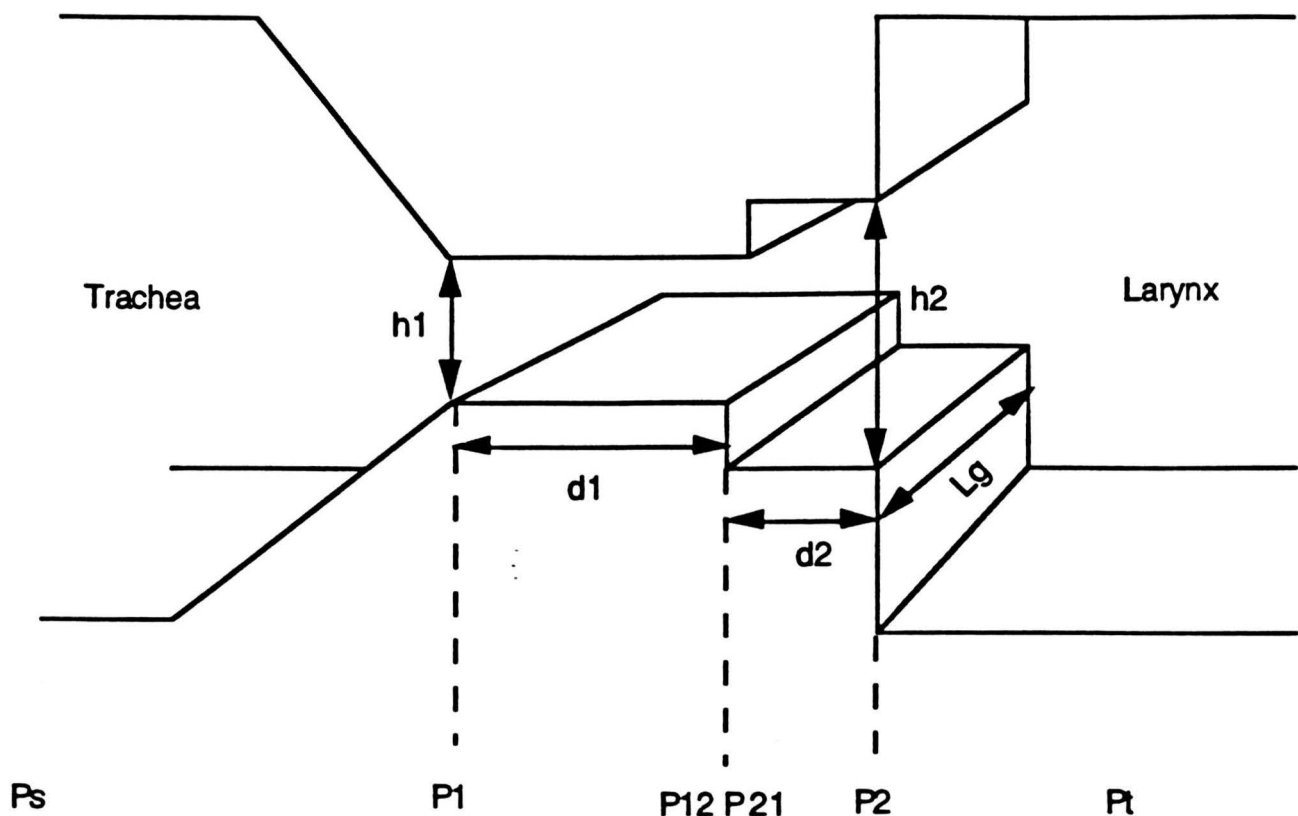
A2)- The gravitational force is negligible.

A3)- Because the Strouhal number  $Sr = f \cdot h / U_g$  is low (order of  $10^{-2}$ ), the glottal flow can be considered as quasi-steady.

A4)- Experiments on flow behaviour through orifices and nozzles showed that the compressibility can be neglected as long as the ratio of the applied pressure to the atmospheric pressure is less than about 0.01 [36]. Since atmospheric pressure at sea level is about 976 cm H<sub>2</sub>O, the assumption of incompressibility is a consequence of the relatively low subglottic pressures (10 cm H<sub>2</sub>O at conversational loudness levels).

#### Static considerations.

Within the assumption A1, the model can be considered as a succession of ducts delimited by parallel planes (figure 20).



**Figure 20** : View of the two mass model as a channel.

Let :

$U_g$  be the volume velocity,

$d_1, d_2$  the length of region 1 and 2,

$h_1, h_2$  the instantaneous opening of region 1 and 2,

$P_s, P_1, P_{12}, P_{21}, P_2, P_t$  are the pressures (referenced to the atmospheric pressure) at several positions among the glottis.

Starting from the trachea, the pressure  $P_s$  is imposed by the lungs. Since the dimensions of the trachea are large and wide, the viscosity in this part is confined in a very thin boundary layer and can be neglected for the main bulk of flow. The pressure drop between the trachea and the first part of the vocal cords model is, according to Bernouilli's law (see appendix A) :

$$P_s - P_1 = \frac{1}{2} \rho \cdot (1 + K) \left( \frac{1}{h_1 \cdot L_g} \right)^2 U_g^2 \quad (7)$$

Where  $\rho$  is the air density.

The coefficient K expresses the pressure loss due to the abrupt contraction, assuming a vena contracta effect. Van den Berg et al.[35] reported a pressure loss  $K = 0.37$ .

In the channel formed by the first masses, the flow is essentially governed by friction forces. According to Poiseuille's formulation of a flow between two parallel plates (see appendix A) :

$$P_1 - P_{12} = \frac{12 \cdot \mu \cdot d_1}{h_1^3 \cdot L_g} \cdot U_g \quad (8)$$

where  $\mu$  is the kinematic viscosity of air.

The effects of the varying shape of the glottis are concentrated in the abrupt discontinuity between the first and the second mass (step-wise approximation). Assuming that the discontinuity is not too abrupt, we can write the pressure drop at this point to be:

$$P_{12} - P_{21} = \frac{\rho}{2} \left( \left( \frac{1}{h_2 \cdot L_g} \right)^2 - \left( \frac{1}{h_1 \cdot L_g} \right)^2 \right) U_g^2 \quad (9)$$

In the channel formed by the second masses, the flow is governed by a Poiseuille's law namely :

$$P_{21} - P_2 = \frac{12 \cdot \mu \cdot d_2}{h_2^3 \cdot L_g} U_g \quad (10)$$

At the end of the glottis, the trachea forms a sudden expansion. In such a situation, experiments showed that the flow separates from the walls, forming a jet bounded by shear layers. This phenomenon induces also formation of vortices and thus dissipation of energy. We assume that at a certain distance (of the order of a few diameters) the jet reattaches to the walls of the larynx. If PL is the pressure in this region, the pressure recovery is :

$$P_2 - P_L = - \frac{\rho}{2} \left( \frac{U_g}{L_g \cdot h_2} \right)^2 \cdot 2 \cdot \frac{L_g \cdot h_2}{AL} \left( 1 - \frac{L_g \cdot h_2}{AL} \right) \quad (11)$$

Where AL is the cross section area of the larynx.

## Dynamic considerations.

Under the assumption A3 (low Strouhal number), the time varying behaviour of the vocal-folds can be considered as a succession of steady states. However, in second approximation, the inertance effects of air can be taken in account in equation (8) and (10) by considering the additional factors respectively :

$$\frac{\rho.d1}{Lg.h_1} \cdot \frac{dU_g}{dt}$$

and

$$\frac{\rho.d2}{Lg.h_2} \cdot \frac{dU_g}{dt}$$

representing the inertia of air in the glottis.

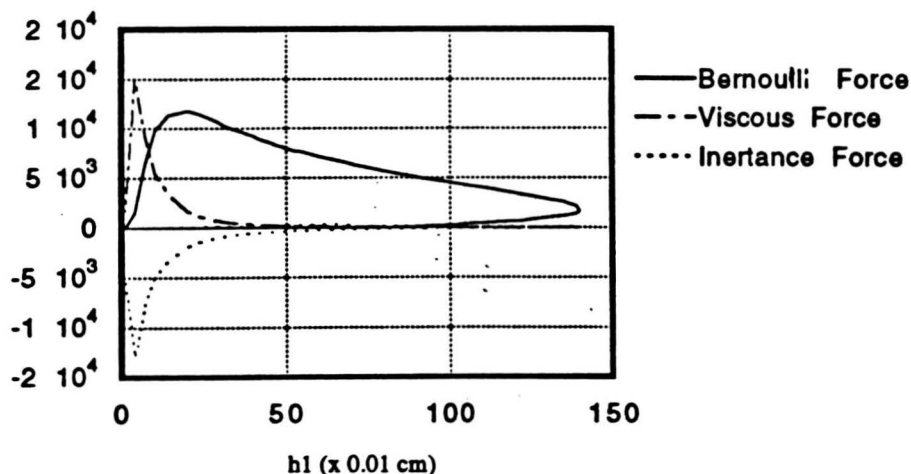
Thus, the driving forces F1 and F2 can be expressed as :

$$F_1 = \frac{P1 + P12}{2} \cdot d1 \cdot Lg = \left[ P_s - \frac{1}{2} \rho \cdot K \left( \frac{1}{h_1 \cdot Lg} \right)^2 U_g^2 - \frac{1}{2} \left( \frac{12 \cdot \mu \cdot d1}{h_1^3 \cdot Lg} \cdot U_g + \frac{\rho \cdot d1}{Lg \cdot h_1} \frac{dU_g}{dt} \right) \right] \cdot d1 \cdot Lg \quad (12)$$

$$F_2 = \frac{P21 + P2}{2} \cdot d2 \cdot Lg = F_1 \cdot \left[ \frac{1}{2} \left( \frac{12 \cdot \mu}{Lg} \left( \frac{d1}{h_1^3} + \frac{d2}{h_2^3} \right) U_g + \frac{\rho}{Lg} \left( \frac{d1}{h_1} + \frac{d2}{h_2} \right) \frac{dU_g}{dt} \right) + \frac{\rho}{2} \left( \left( \frac{1}{h_2 \cdot Lg} \right)^2 - \left( \frac{1}{h_1 \cdot Lg} \right)^2 \right) U_g^2 \right] \cdot d2 \cdot Lg \quad (13)$$

This set of equations, although formulated in a different way, is similar to the original one proposed by Ishizaka and Flanagan [2]. Anticipating our results, it is interesting to look at the respective order of magnitude of each term. Figure 21 displays the respective magnitude of each terms in equation (12) during a cycle of vibration.

Amplitude (dynes)



**Figure 21:** Magnitude of forces during a vibration phase.



The total pressure drop through the glottis is :

$$\Delta P = P_s - P_l = \frac{1}{2} \rho \cdot K \left( \frac{1}{h_1 \cdot L_g} \right)^2 U_g^2 + \frac{12 \cdot \mu \cdot d_1}{h_1^3 \cdot L_g} \cdot U_g + \frac{\rho \cdot d_1}{L_g \cdot h_1} \frac{dU_g}{dt} + \frac{\rho}{2} \left( \left( \frac{1}{h_2 \cdot L_g} \right)^2 - \left( \frac{1}{h_1 \cdot L_g} \right)^2 \right) U_g^2 + \frac{12 \cdot \mu \cdot d_2}{h_2^3 \cdot L_g} U_g + \frac{\rho \cdot d_2}{L_g \cdot h_2} \cdot \frac{dU_g}{dt} - \frac{\rho}{2} \left( \frac{U_g}{L_g \cdot h_2} \right) \cdot 2 \cdot \frac{L_g \cdot h_2}{A_l} \left( 1 - \frac{L_g \cdot h_2}{A_l} \right) \quad (14)$$

Using a two-mass model of the vocal folds, we are able to describe the phonation process using three coupled equations :

- The first two are mechanical equations (2), describing the motion of each mass,
- The third one (14) describes the flow behaviour through the glottis.

Since three unknown ( $y_1(t)$ ,  $y_2(t)$  and  $U_g(t)$  for instance) are enough to fully describe the whole process, the problem can be theoretically solved.

## II.3. Results.

Results derived from the two-mass model have already been extensively studied and discussed [38], [3]. It is not our purpose to reproduce all these works; we only present in the following section detailed results obtained for one typical configuration.

### II.3.1. Choice of the parameters.

#### - Geometrical characteristics.

This set of parameters is directly derived from the physiological observations presented in section I. The effective length of the vocal cords is chosen to value 1.4 cm while the overall stiffness is chosen as  $d_1 + d_2 = 0.3$  cm. Following the remarks of Ishizaka et al. [2], each vocal cord is divided into two unequal parts  $d_1 = 0.25$  cm and  $d_2 = 0.05$  cm.

Both Larynx and trachea are assumed to have an area of 2.5 cm<sup>2</sup>,

#### - Mechanical parameters.

This set of parameters is almost impossible to derive accurately from biological studies. Indeed, the vocal folds are a multi-layered non uniform structure. Thus it is very difficult to derive a global characteristic for the purpose of the two-mass model. Since evaluations of stiffnesses and dampings are quite uncertain, there is a wide range of possible values for the model. Guerin et al. [38] made a systematic study about the influence of these parameters. They showed, for instance, that efficient simulations were possible using only one global mass-tension parameter. In accordance with to some results they presented, we choose the following values, corresponding to a fundamental frequency of about 120 Hz :

Spring number 1 (first mass):

Stiffness  $k_1 = 45 \text{ kDynes.cm}^{-2}$ ,

Non linear coefficient  $\eta_{k1} = 100$ .

Damping  $r_1 = 17.5 \text{ Dynes.cm}^{-1}.\text{s}$

Mass  $M_1 = 0.17 \text{ g}$ .

Spring number 2 (second mass):

Stiffness  $k_2 = 8 \text{ kDynes.cm}^{-2}$ ,

Non-linear coefficient  $\eta_{k2} = 100$ .

Damping  $r_2 = 18 \text{ Dynes.cm}^{-1}.\text{s}$ ,

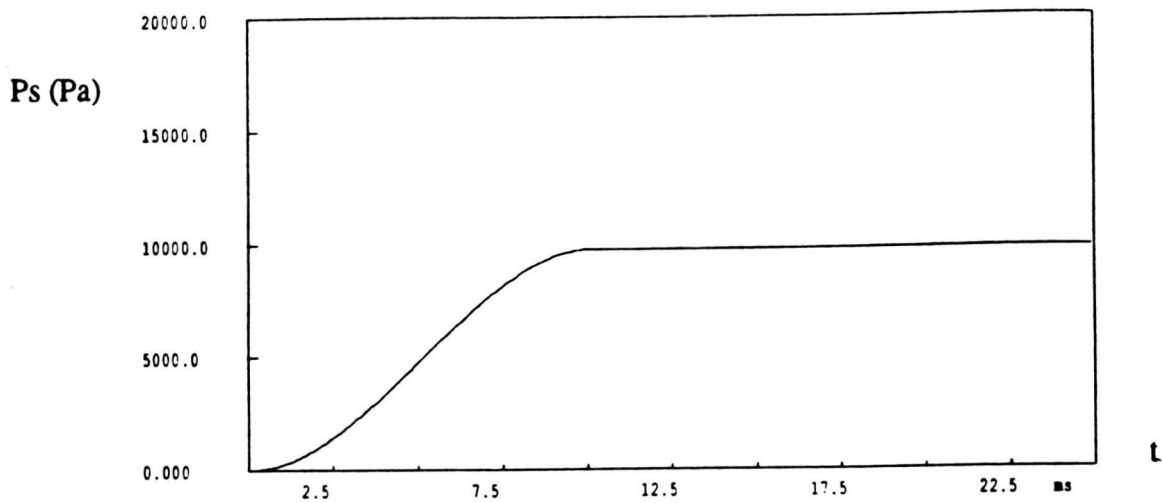
Mass  $M_2 = 0.03 \text{ g}$ .

Spring number 3 (coupling between mass 1 and mass 2):

Stiffness  $k_{12} = k_{21} = 25 \text{ kDynes.cm}^{-2}$ ,

### - Subglottic pressure

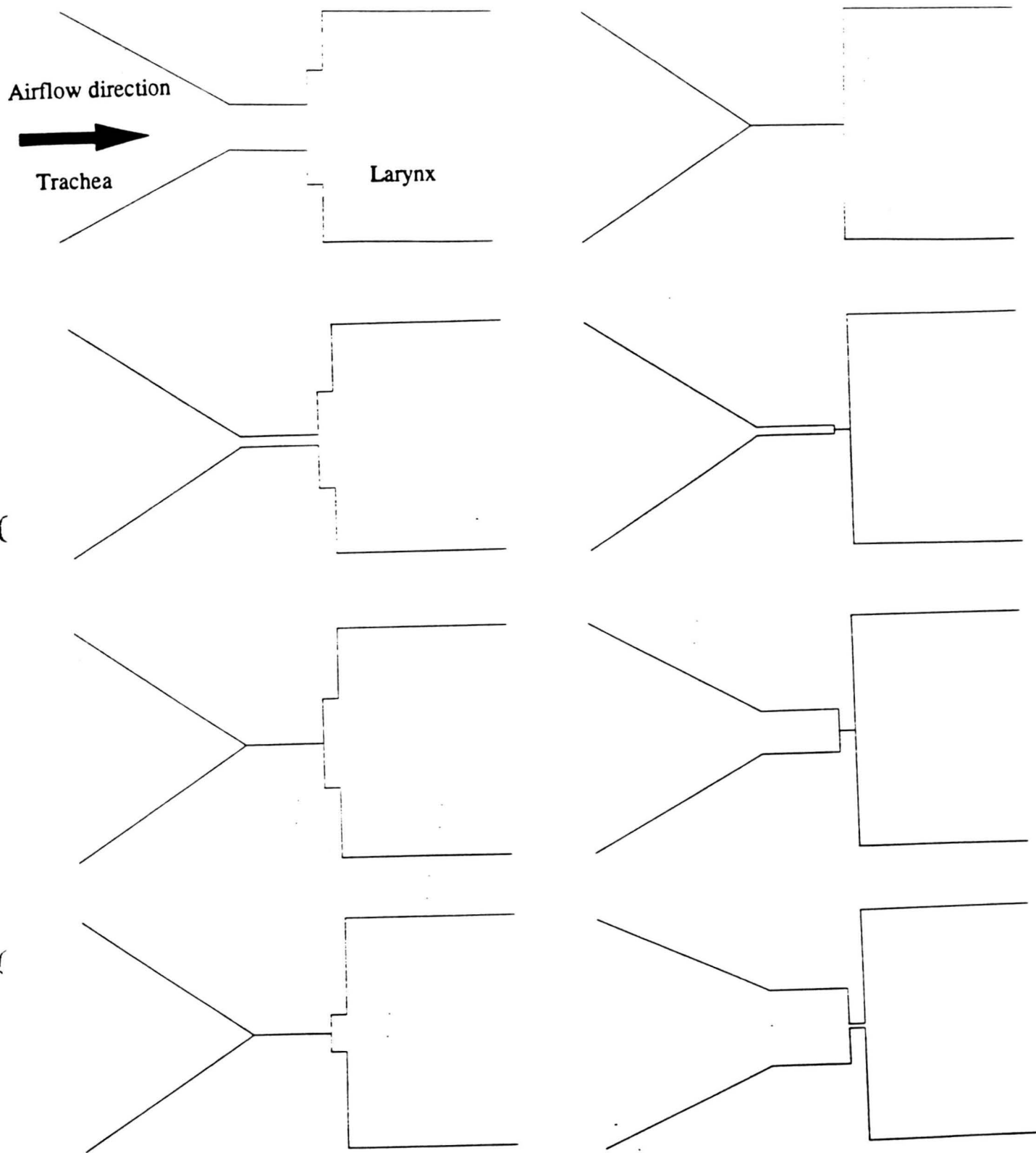
For simplicity, we modeled the subglottic pressure as a smoothly increasing function of time up to 10 ms whereas the lung pressure is maintained constant. Figure 22 presents the evolution of  $P_s$  as a function of time.



**Figure 22** : Model for subglottic pressure as a function of time.

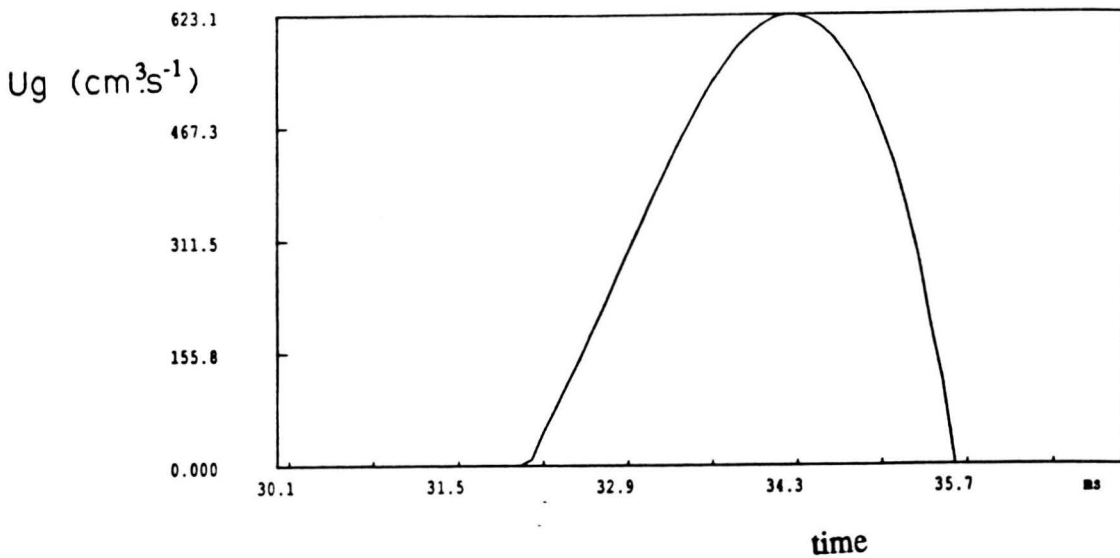
### II.3.2. Examples of results.

Figure 23 presents the sketch of the two-mass model during one cycle of oscillation. Comparing with observations upon real vibrations of the vocal folds (see figure 12 of section I.3.2), the two-mass model reproduces well the observed phase difference between its lower and upper part.



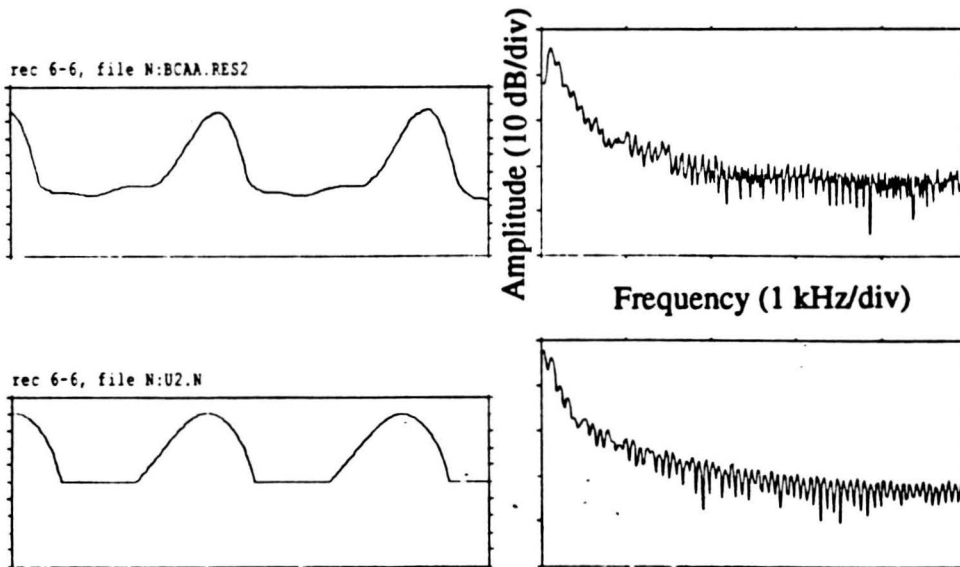
**Figure 23** : Sketch of the two mass model vibration during one cycle of vibration.

The volume velocity waveform  $U_g(t)$  is represented on figure 24.



**Figure 24** : Volume velocity  $U_g(t)$  computed with the two mass model.

These results are of course consistent with other simulations using similar models and presented in the literature [2], [38], [39]. More important, they also present similarities with speech signals derived from inverse filtering [1] as shown figure 25.



**Figure 25** : Comparison between synthesized glottal pulses (lower curve) and glottal pulses derived from inverse filtering (upper curve).

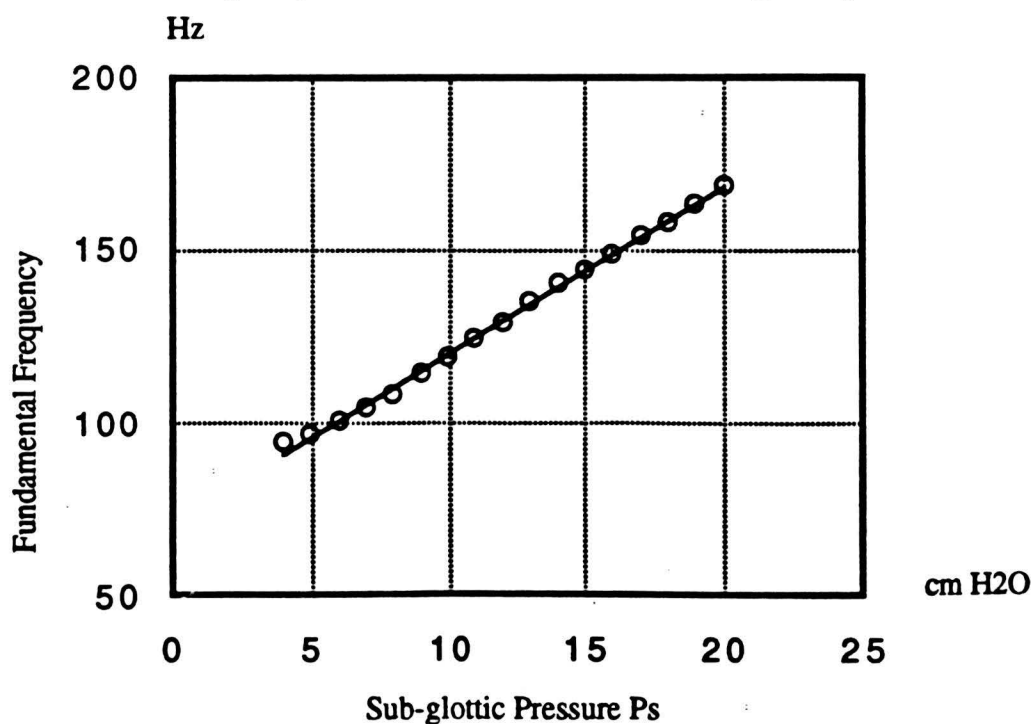
One should take the greatest care with such a comparison. First because attempting to fit reality with such simple model seems quite presumptuous, second because inverse filtering techniques also induce errors in the estimation of glottal pulses. However, qualitative comparison reveals a very similar spectrum although the glottal pulses themselves are quite different especially for the closure period.

Moreover, the following parameters were calculated upon the volume velocity waveform :

- Fundamental frequency : 125 Hz,
- Speed quotient  $Sq = 1.5$ ,
- Open quotient  $Oq = 0.6$ ,
- Peakness  $P = 0.8$ ,
- Closed-Open ratio  $C = 0.1$ .

All these values lie within the range derived from real speech analysis (see table 1).

The strong correlation between sub-glottic pressure value and fundamental frequency observed for real speech is also well described using the two-mass model. In figure 26 we present the fundamental frequency calculated as a function of the sub-glottic pressure  $P_s$ .

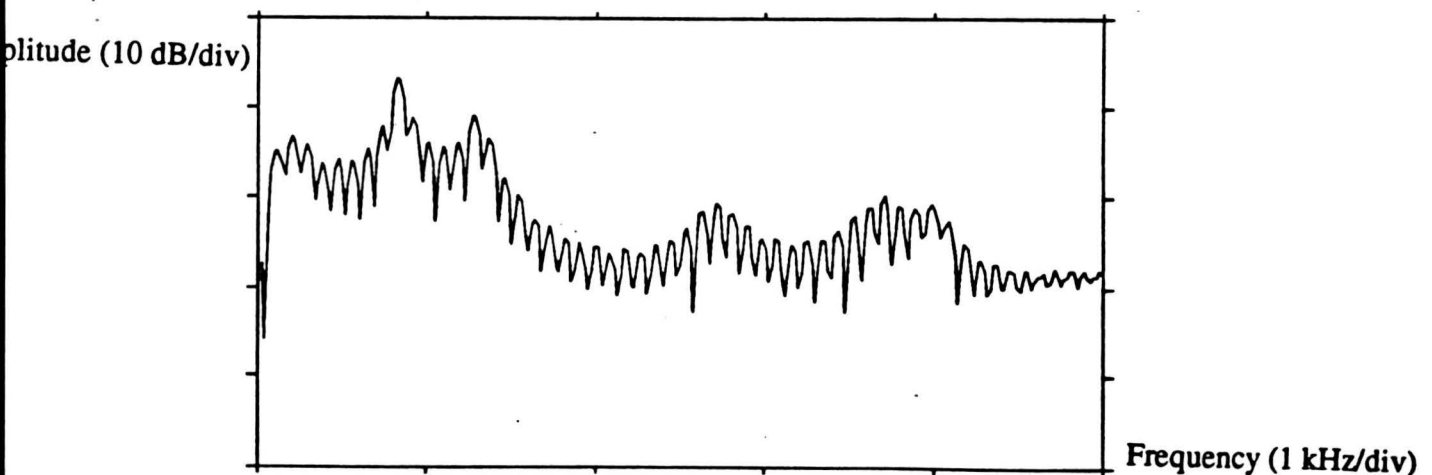


**Figure 26** : Calculated fundamental frequency as a function of the sub-glottic pressure  $P_s$ .

This result is very close to the dependency presented figure 6 of section I.1.

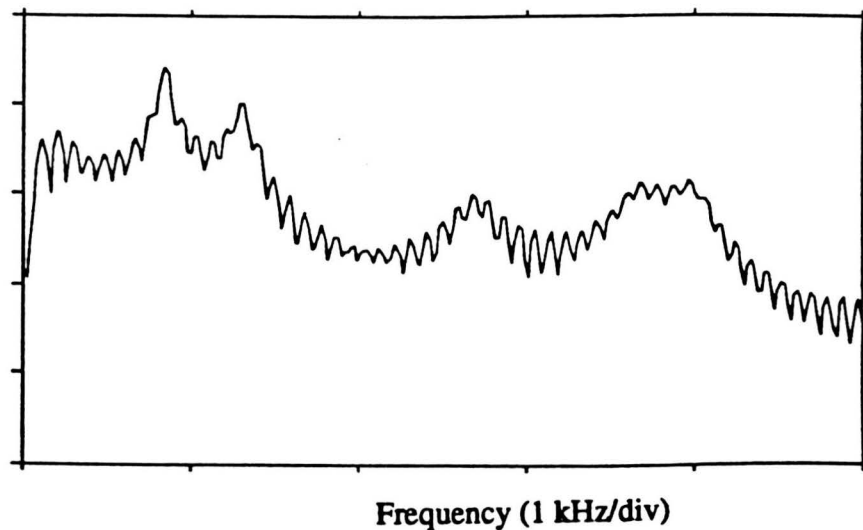
### II.3.3. Speech production, a model for the vocal tract.

Even though it was not our primary goal, it seems interesting to try to derive synthetic speech from the results of the two-mass model. To do so, one has to use a model for the vocal tract. Several different approaches are possible, each having their own advantages and drawbacks. Most of the vocal tract descriptions are based on a decomposition of the vocal tract into a number of tubes, the area of each tube depending upon the sound to be produced. An other approach, less physiologically oriented, is to modelize the vocal tract as a whole via a filter (possibly a time varying filter). A very simple and efficient description of the vocal tract has been proposed and developed by Mrayati and Guerin [40], using a simple filter whose characteristics are determined by the formant frequencies of the sound to be produced. An even simpler approach is to use the filter characteristics derived from a LPC analysis of real speech. A detailed presentation of this method can be found in reference [41]. Figure 27 presents the output signal obtained for the sustained sound /a/ using such method.



**Figure 27** : Output spectrum of a /a/ simulated by the two mass model coupled with a LPC filter for the vocal tract.

As can be seen, and heard, by comparison with real speech signal (figure 28) the results obtained describe reasonably well the main structure of voiced sounds.



**Figure 28** : Pressure output and spectrum of a /a/ for real speech.

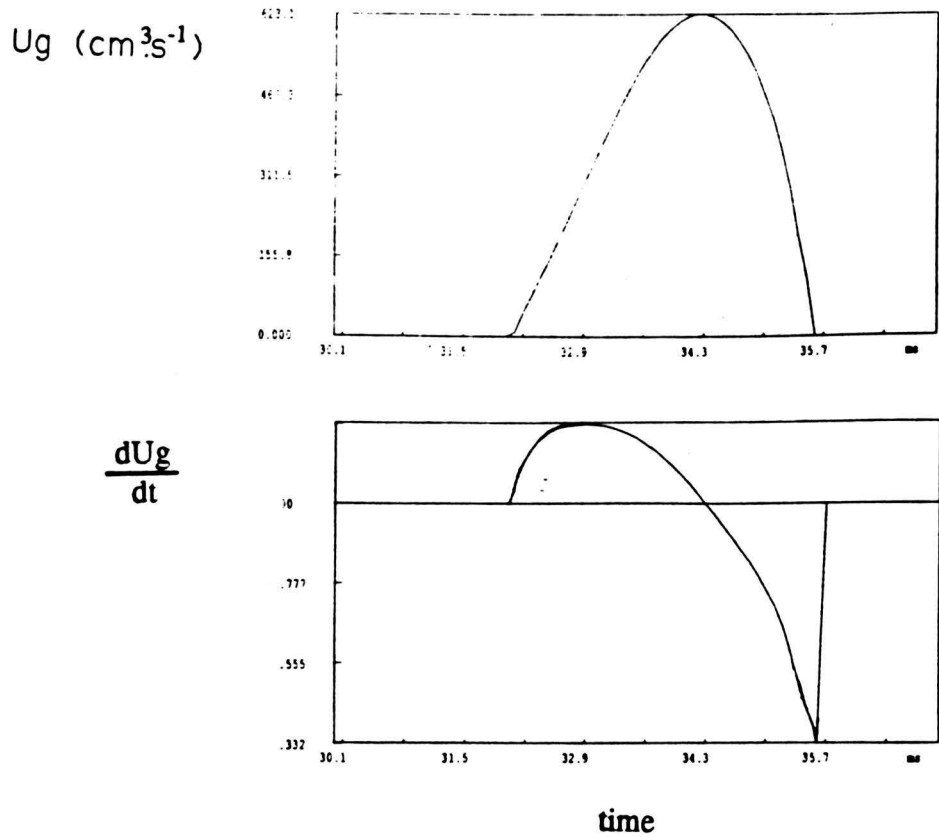
## II.4. Conclusion.

As we showed by using few examples, the two-mass model derived can already produce a glottal signal close to reality and thus reasonable synthetic speech. In addition, the simulation results fit well to some physiological observations such as :

- The phase difference between lower and upper lip of the vocal cords,
- The correlation between the fundamental frequency and the subglottic pressure,
- The ability to produce sounds for a realistic range of subglottic pressures (from 3 cm H<sub>2</sub>O).

The major problems that occur concern mainly the closing and the closed portion of the vibratory movement. First, the duration of closure we found are somewhat higher that what has been observed. Second, but not least, in the model, the volume velocity waveform exhibits a very steep decrease at closure. As a consequence, the time derivative of the volume velocity will present a sharp discontinuity as shown on figure 29. This discontinuity is not realistic and will generate a considerable amount of higher harmonics.





**Figure 29** : Volume velocity waveform and its corresponding time derivative.

Since almost all models for the vocal tract excitation are based on the time derivative of the volume velocity, this discontinuity can be acoustically important and be responsible for many synthesis imperfections.

### **III Modifications and discussion.**

Although the two mass model described here is able to produce some realistic signals, a more detailed study showed that many physiological or physical events were not correctly described and could be held responsible for many limitations. In the following, we are going to focus on two different problems. The first one concerns the biomechanical description of the closure of the glottis while the other one is more related to the theoretical description of the flow behaviour within the glottis.

#### **III.1. The collision problem.**

##### **III.1.1. Problems.**

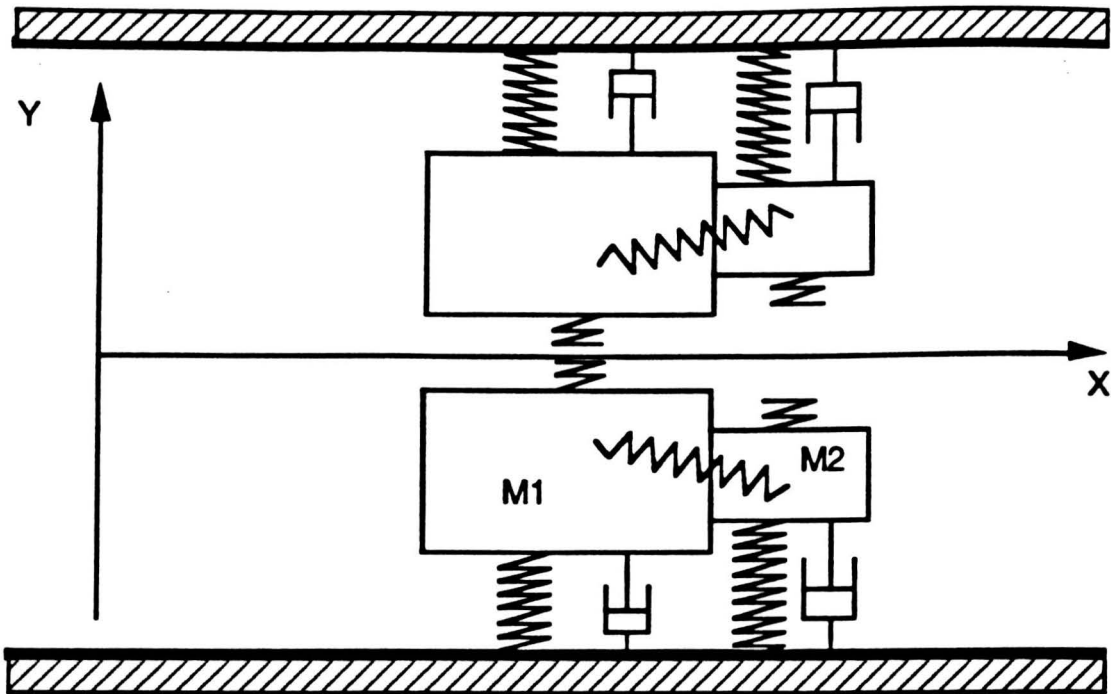
A very acute problem occurs when the masses are colliding in the model (that is when the glottis closes). Without any refinements, the simple two mass model described in the preceding section leads to unrealistic times of closure and thus to unacceptable sounds.

In the following we are going to present different methods to describe more realistic behaviour of the glottis closure. In first approximation, we can describe the collision forces as a restoring spring force. Furthermore, additional air displacement can be added leading to a less steep decrease of the volume velocity.

##### **III.1.2. Spring models.**

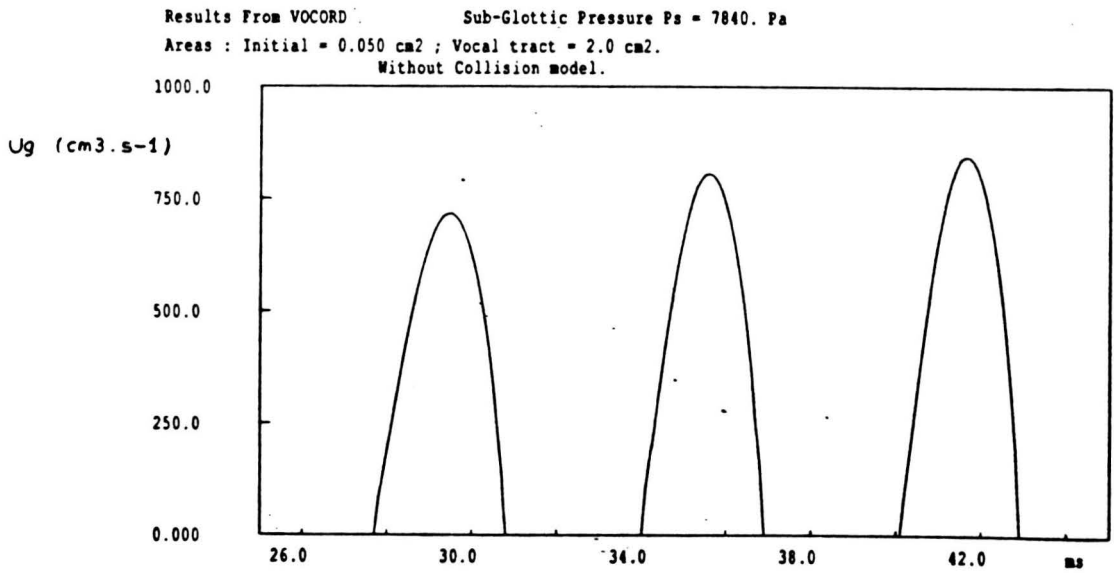
###### **III.1.2.1. Damped-spring model.**

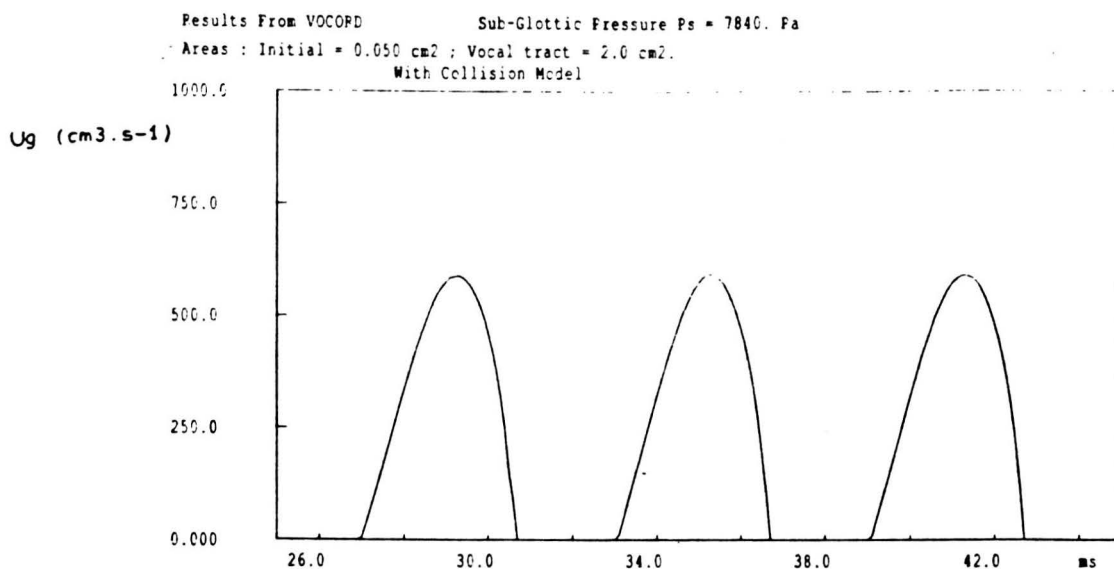
This model originally proposed by Ishizaka [2] is generally well accepted although it is a very crude approximation of reality. The restoring force is represented by an equivalent damped-spring system as shown on figure 30.



**Figure 30** : modelization of collision using a damped-spring system.

The characteristics of the additional springs are assumed to be similar as the restoring springs (i.e. non-linear). Figure 31 proposes a comparison of glottal pulses computed with and without collision model.





**Figure 31** : Comparison between three successive volume velocity waveforms with collision model (upper curve) and without collision model (lower curve).

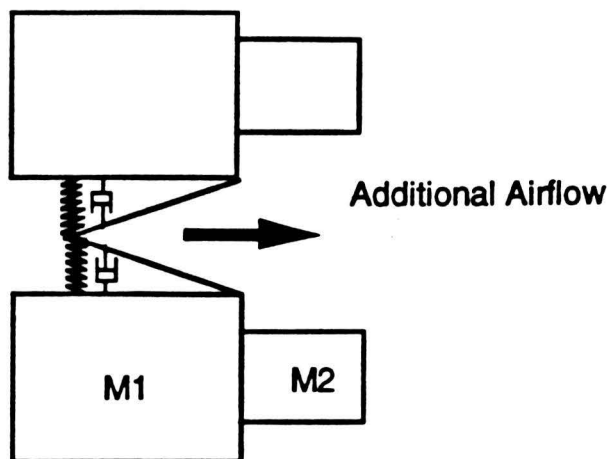
As can be seen, this model has the advantage to provide more realistic times of closure and thus speech results. However, the very steep closure still induces a discontinuity for the time derivative of the volume velocity.

### III.1.2.2. Introduction of additional displacement

Following the simple representation of the preceding section, we tried to include a new phenomenon in order to derive more realistic results of the closure of the vocal folds. In fact this problem is closely related to the shape the model has. In the original two-mass model the glottis is assumed to have a rectangular shape. In particular, as soon as the first channel is closed there is no more airflow allowed to pass through the glottis. In reality, since the glottis at closure forms roughly a divergent channel, there is still a certain amount of air exiting the glottis until the vocal cords are totally closed.

To take into account an additional air flow at closure, we propose here a very simple modification of the two-mass model. When the glottis starts to close from the anterior part

(first mass), we do no longer assume a rectangular shape of the vocal cords, the shape is considered to be divergent as shown on figure 32.



**Figure 32** : Introduction of additional air flow using a divergent shape of the glottis.

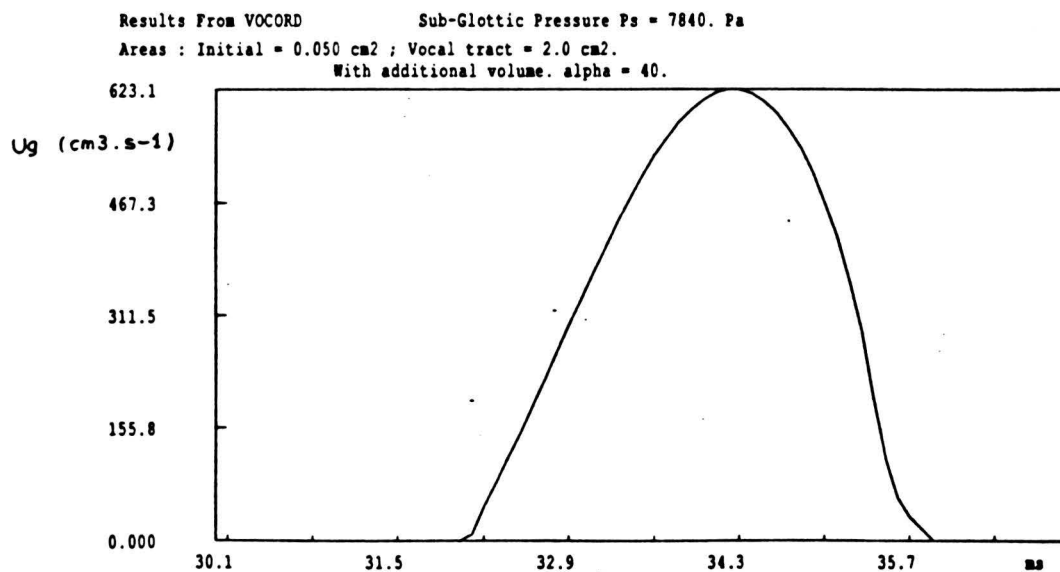
The additional volume velocity can be expressed as :

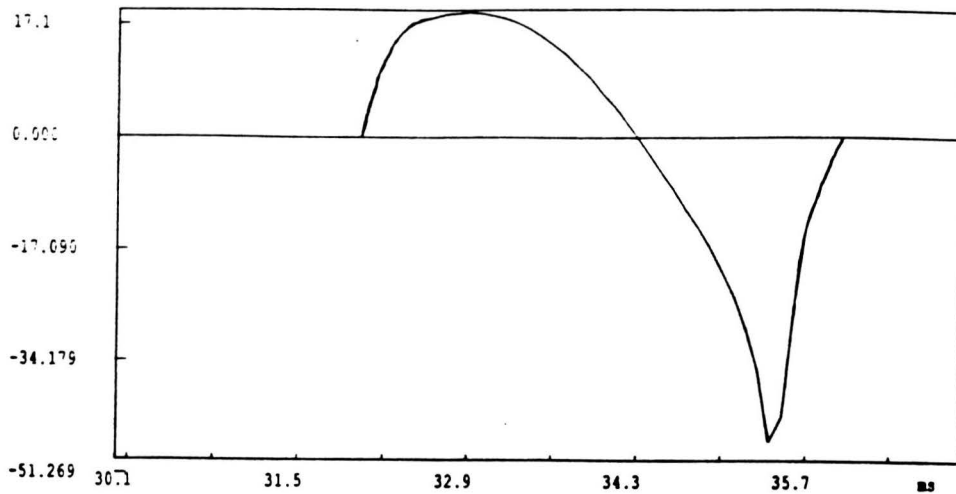
$$U_g = - \frac{dV}{dt}$$

Where  $V = \frac{1}{2} \cdot d(t) \cdot h_1(t) \cdot L_g = \frac{h_1^2(t)}{4 \cdot \tan(\alpha)} \cdot L_g$ .  $L_g$  is the volume of air between the masses,

$\alpha$  is the semi-angle of divergence.

Figure 33 is an example of result, for  $\alpha = 40^\circ$ , showing the effect of additional air-flow on both  $U_g(t)$  and its time derivative.





**Figure 33** : Computed  $U_g(t)$  with additional air flow at closure. Lower curve show the time derivative of  $U_g(t)$ .

As one can see from this figure the difference for  $U_g(t)$  is very small and has only a slight effect upon the fundamental frequency. However, there is a fair reduction of the sharp discontinuity of the time derivative which was responsible of the generation of higher harmonics.

These simple solutions lead to much more acceptable speech sounds. However, the main problem concerning the behaviour of human tissues during closure (due to the non linear contact area) and especially the model for the forces involved are still not solved. The simple model of a spring, even with non linear characteristics, is certainly a very rough approximation of reality. As a consequence, the time of closure might not be correctly predicted as well as the behaviour of the glottis just at the opening time. A more detailed study, involving a biomechanical description of the tissues, should be performed in order to solve this problem.

### III.2. A modified description of the flow within the two-mass model.

In the following we are going to present a detailed description of the flow within the channels formed by the two-mass model. We intend to introduce new effects that were not originally taken into account.

We assume the flow to be steady, incompressible and between two parallel plates.

#### III.2.1. Flow at the entrance of the glottis (first channel).

In the original model [2], the authors considered the entrance of the glottis as a sudden contraction so abrupt that a vena contracta took place. This assumption was supported by the empirical results of van den Berg et al. [35]. In fact, because the entrance of the glottis is not sharp edged, we don't expect a vena contracta to occur [42] (see appendix A). However, possible, but limited, losses can be induced by flow separation upstream of the glottis.

The pressure drop due to the entrance of the glottis is thus simply calculated from Bernoulli's formula :

$$\Delta P_{\text{inlet}} = \frac{\rho}{2} \frac{U_g^2}{A_1^2} \left( 1 + K \left( 1 - \frac{A_1}{A_t} \right)^2 \right) - \frac{\rho}{2} \frac{U_g^2}{A_t^2} \quad (14)$$

where  $A_1 = L_g \cdot h_1$  is the cross sectional area of the first channel of the glottis and  $A_t$  is the area of the trachea

Since  $A_1 \ll A_t$  this expression can be approximated by :

$$\Delta P_{\text{inlet}} = \frac{\rho}{2} \frac{U_g^2}{A_1^2} (1+K) \quad (15)$$

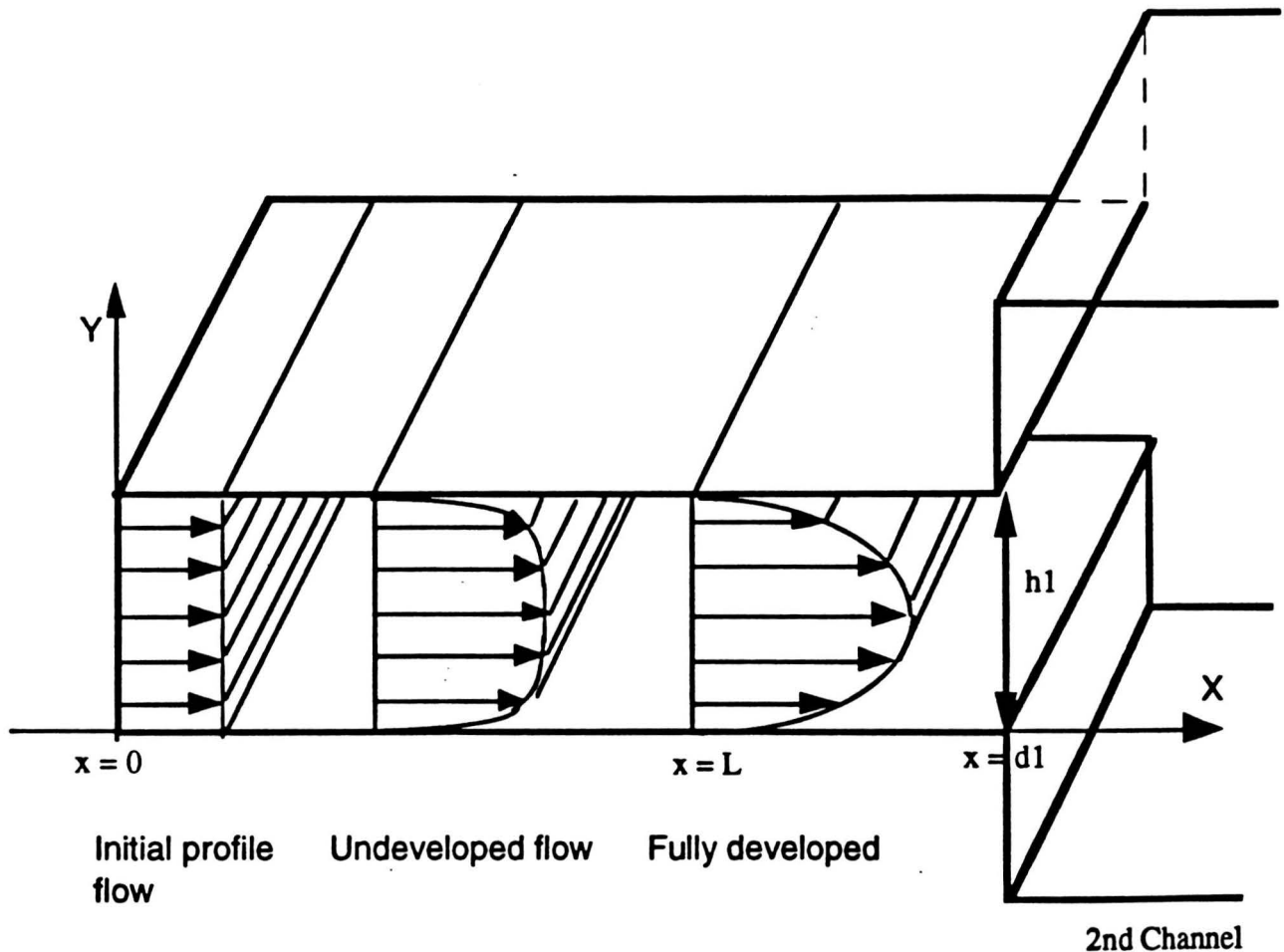
The coefficient  $K$  expresses the losses due to the contraction of the glottis. For rounded edges, as it is the case for the vocal cords, the values of  $K$  is experienced to be very small [42] and can even be neglected. In this case the pressure drop equals :

$$\Delta P_{\text{inlet}} = \frac{\rho}{2} \frac{U_g^2}{A_1^2} \quad (16)$$

### III.2.2. Flow through the first channel of the glottis.

#### Intermediate flow.

We assume the flow to be uniform at the entrance of the glottis. Because of the viscosity effects, the velocity will evolve from a uniform profile to a fully developed parabolic profile (Poiseuille flow) as shown on figure 34.



**Figure 34** : Evolution of the flow from uniform profile to parabolic profile.

This transition can be described in terms of boundary layer flow as follows [2],[42],[43]:



In the intermediate region, the flow can be divided in two regions:

- A frictionless region
- A boundary layer in which the effects of viscosity are predominant.

Let  $\delta = \delta(x)$  be the thickness of the boundary layer.

Outside the boundary layer, we assume the velocity to be uniform. Let  $u_e$  be the main flow velocity.

Within the boundary layer the momentum conservation equation can be written as (see appendix A):

$$u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -\frac{1}{\rho} \frac{\partial P}{\partial x} + \nu \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \quad (17)$$

Using a dimensional analysis, it can be shown that the term  $\frac{\partial^2 u}{\partial x^2}$  can be neglected in

comparison with  $\frac{\partial^2 u}{\partial y^2}$ . Equation (17) reduces to the form :

$$u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -\frac{1}{\rho} \frac{\partial P}{\partial x} + \nu \frac{\partial^2 u}{\partial y^2} \quad (18)$$

Integration of this equation over the boundary layer can be done and leads to the von Karman equation :

$$\frac{d\theta}{dx} + (\delta_* + 2\theta) \frac{1}{u_e} \frac{du_e}{dx} = \frac{\tau_w}{\rho u_e^2} \quad (19)$$

where :

$\tau_w = \mu \left( \frac{\partial u}{\partial y} \right)_{y=0}$  is the viscous shear stress at the wall,

$\delta_* = \int_0^\delta \left( 1 - \frac{u}{u_e} \right) dy$  is the thickness displacement,

$\theta = \int_0^\delta \frac{u}{u_e} \left( 1 - \frac{u}{u_e} \right) dy$  is the momentum thickness.

The analysis of the boundary layer can be done using the approximate method of Karman-Polhausen and neglecting the variation in the shape of the boundary layer [44]. Let us assume that the velocity profile follows a fourth order polynomial law of the form :

$$u = a_0 + a_1 y + a_2 y^2 + a_3 y^3 + a_4 y^4$$

with respect to the boundary conditions :

$$\text{for } y = 0 \quad u = 0$$

$$\text{for } y = \delta \quad u = u_e, \quad \frac{\partial u}{\partial y} = 0, \quad \frac{\partial^2 u}{\partial y^2} = 0$$

the constants  $a_0$  and  $a_1$  can be determined and the velocity profile becomes :

$$u = \left( 2 \frac{y}{\delta} - 2 \left( \frac{y}{\delta} \right)^3 + \left( \frac{y}{\delta} \right)^4 \right) u_e \quad (20)$$

Using this approximation for the velocity together with the von Karman equation (18) we find:

$$\frac{37}{315} u_e \delta \frac{d\delta}{dx} + \frac{337}{630} \delta^2 \frac{du_e}{dx} = 2\nu \quad (21)$$

$$\text{since } \theta = \frac{37}{315} \delta, \quad \delta_* = \frac{3}{10} \delta \text{ and } \tau_w = 2\mu \frac{u_e}{\delta}$$

Introducing the volume velocity  $U_g = Lg \left( h_1 - \frac{3}{5} \delta \right) u_e$ , equation (20) takes the form :

$$\frac{U_g}{6300 \cdot Lg \cdot v} \left( \frac{370 \cdot h_1 \cdot \delta + 789 \cdot \delta^2}{\left( h_1 - \frac{3}{5} \cdot \delta \right)^2} \right) \frac{d\delta}{dx} = 1 \quad (22)$$

Integration of this equation can be easily done and leads to the relationship :

$$\frac{U_g}{756 Lg v} \left( 263 \cdot \frac{\delta}{h_1} - \frac{1685}{3} \cdot \frac{\frac{3}{5} \cdot \delta}{1 - \frac{3}{5} \cdot \frac{\delta}{h_1}} + 1000 \cdot \ln \left( 1 - \frac{3}{5} \cdot \frac{\delta}{h_1} \right) \right) = \frac{x}{h_1} \quad (23)$$

Let  $L$  be the distance needed for the flow to become fully developed. Two cases are possible:

1) In the first one, the length of the channel,  $d_1$ , is not long enough for a fully developed flow to establish. Namely this will occur if  $d_1 < L$  or equivalently if  $\delta(d_1) \leq \delta(L)$ . In this case, the flow doesn't have time to become parabolic. The pressure drop is thus different from a fully developed Poiseuille flow and can be expressed as :

$$\Delta P_{undeveloped} = P(0) - P_{12} = \frac{\rho}{2} \left( \frac{U_g}{\left( \left( 1 - \frac{3}{5} \cdot \frac{\delta(d_1)}{h_1} \right) h_1 Lg \right)} \right)^2 \quad (24)$$

$\delta(d_1)$  can be calculated using relation (23) for  $x = d_1$  :

$$\frac{U_g}{756 Lg v} \left( 263 \cdot \frac{\delta(d_1)}{h_1} - \frac{1685}{3} \cdot \frac{\frac{3}{5} \cdot \delta(d_1)}{1 - \frac{3}{5} \cdot \frac{\delta(d_1)}{h_1}} + 1000 \cdot \ln \left( 1 - \frac{3}{5} \cdot \frac{\delta(d_1)}{h_1} \right) \right) = \frac{d_1}{h_1} \quad (25)$$

2) In the second case, the length of the channel  $d1$  is high enough to allow a fully developed Poiseuille flow. The pressure drop due to the development of the flow can be expressed as:

$$\Delta P_D = P(0) - P(L) = \frac{\rho}{2} \frac{U_g^2}{h1^2 \cdot Lg \cdot \left(1 - \frac{3}{5} \cdot \frac{\delta(L)}{h1}\right)^2} \quad (26)$$

while the pressure loss in the last part of the channel ( $L \leq x \leq d1$ ) is simply described by a Poiseuille's law :

$$\Delta P_P = P(L) - P(d1) = \frac{12 \mu (d1-L)}{Lg \cdot h1^3} U_g \quad (27)$$

Thus, the total pressure drop over the entire region  $0 \leq x \leq d1$  is ((26) + (27)):

$$\Delta P_{\text{developed}} = \Delta P_D + \Delta P_P = \frac{\rho}{2} \frac{U_g^2}{h1^2 \cdot Lg \cdot \left(1 - \frac{3}{5} \cdot \frac{\delta(L)}{h1}\right)^2} + \frac{12 \mu (d1-L)}{Lg \cdot h1^3} U_g \quad (28)$$

**Remark** : If we neglect the losses due to the development of the flow, that is  $L = 0$ , we find the pressure drop to be :

$$\Delta P_{\text{developed}} = P(0) - P(d1) = \frac{12 \mu d1}{Lg h1^3} U_g \quad (29)$$

which is the classical formula (8) used for the two-mass model.

### III.2.3. Flow in the second channel.

Depending on the shape of the glottis, two situations should be considered :

1) Flow in a convergent glottis (opening glottis).

Losses can be induced by the contraction of the flow. Following the description for the flow at the entrance of the first channel we can write the pressure drop due to the contraction between section 1 and 2 in the form :

$$\Delta P_{12c} = P_-(d1) - P_+(d1) = \frac{\rho}{2} \frac{U_g^2}{A_2^2} \left( 1 + K_c \left( 1 - \frac{A_1}{A_2} \right)^2 \right) - \frac{\rho}{2} \frac{U_g^2}{A_1^2} \quad (30)$$

Where  $P_-(d1)$  and  $P_+(d1)$  are respectively the pressures just upstream and downstream at  $x = d1$ .  $A_2 = h_2 \cdot d_2$  is the area of the second channel.  $K_c$  is the pressure loss coefficient due to the contraction.

Remark : Neglecting the possibility of a pressure loss due to the contraction ( $K_c = 0$ ), equation (30) becomes :

$$\Delta P_{12} = \frac{\rho}{2} \left( \frac{1}{A_2^2} - \frac{1}{A_1^2} \right) U_g^2 \quad (31)$$

which is formula (9) of the classical model.

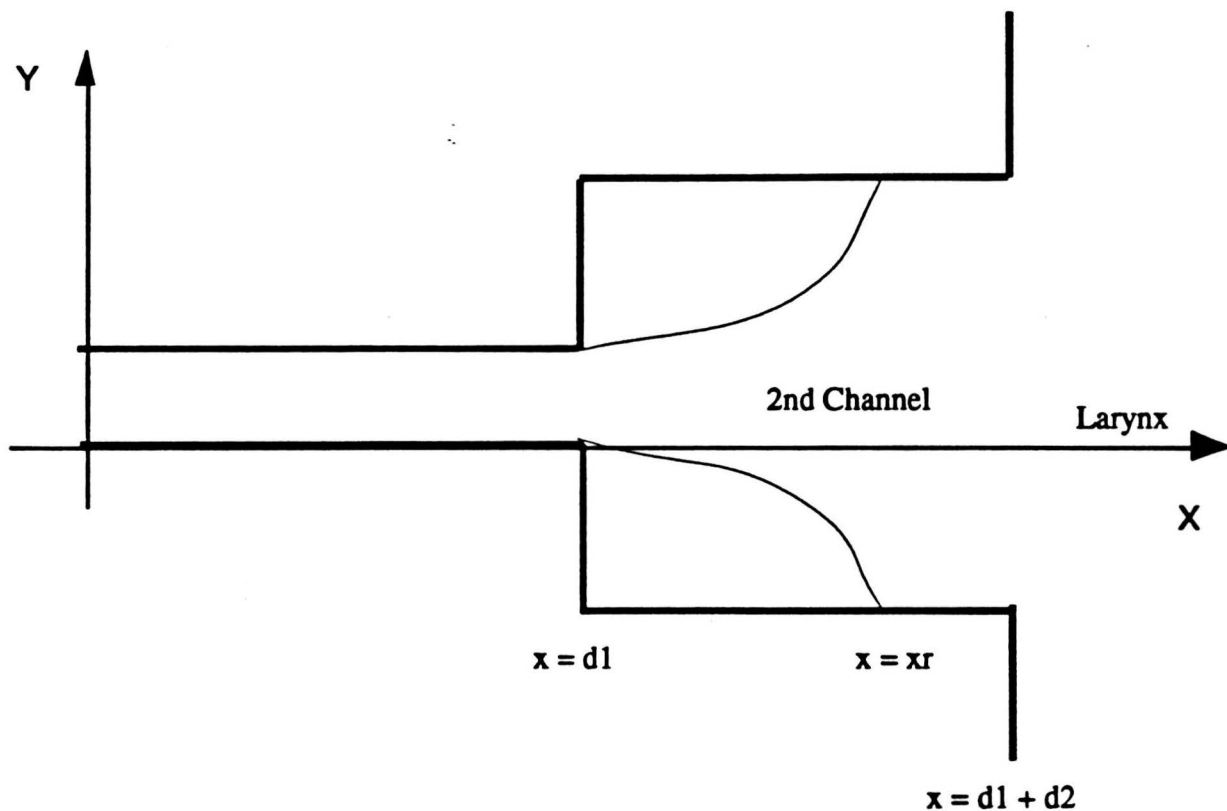
The flow within the second channel can be formally described in the same way as for the first channel noticing however that since  $d_2$  is even smaller than  $d_1$ , except for very small apertures, the flow is even more likely to be undeveloped at the outlet of the second channel.

2) Flow in a divergent glottis (closing phase).

For an abrupt discontinuity ( $h_1 \ll h_2$ ) the flow might separate from the walls forming a jet within the second channel. In this case, there will be no pressure loss due to the discontinuity:

$$\Delta P_{12d} = 0$$

Let  $x = x_r$  be the eventual point of reattachment of the jet (figure 35).



**Figure 35** : Flow separation in a divergent two-mass model.

We considered two situations :

- In the first one, the channel is not long enough ( $x_r > d_1 + d_2$ ) so that the flow is separated over the entire length of the second channel. No pressure drop is associated with this situation.

- In the second case ( $x_r < d_1 + d_2$ ), the flow reattaches to the second channel. The pressure drop due to the flow separation can be written as (see appendix A) :

$$\Delta P_{\text{separation}} = P_+(d_1) - P(x_r) = \frac{\rho}{2} \frac{U_g^2}{A_1^2} \cdot 2 \cdot \frac{A_1}{A_2} \left( \frac{A_1}{A_2} - 1 \right) \quad (32)$$

where  $A_2$  is the area of the second channel.

### III.2.4. Flow in the larynx.

Because the larynx constitutes a sudden expansion, we consider flow separation at the end of the second channel. Ishizaka et al. assume the flow to reattach further downstream and thus expect a pressure recovery. In fact, experiments show that there is almost no pressure recovery since the energy of the jet is dissipated by turbulence.

### Summary.

Equations (14) to (31) describe in some detail the flow behaviour at the entrance, through and at the end of the glottis. Table 2 presents a comparison between classical theory and the modified one.

|                |   | Pressure Difference   |                |
|----------------|---|---|----------------|
|                |   | Classical model   | Modified Model |
| Inlet          | Vena contracta effect<br>$\frac{\rho}{2} \frac{U_g^2}{A_1^2} (1+K)$ with $K = 0.375$ .  | No vena contracta<br>$\frac{\rho}{2} \frac{U_g^2}{A_1^2} (1+K)$ with $K \ll 1$ .  |                |
| First Channel  | Fully developed flow:<br>$\frac{12 \cdot \mu \cdot d_1}{h_1^3 \cdot L_g} \cdot U_g$   | - $d_1 < L$ Undeveloped flow :<br>$\frac{\rho}{2} \left( \frac{U_g}{\left(1 - \frac{\delta(d_1)}{h_1}\right) h_1 L_g} \right)^2$<br>- $d_1 > L$ Development and Fully developed flow :<br>$\frac{\rho}{2} \frac{U_g^2}{h_1^2 \cdot L_g \cdot \left(1 - \frac{\delta(L)}{h_1}\right)^2} + \frac{12 \mu (d_1-L)}{L_g \cdot h_1^3} U_g$  |                |
| Transition     | Bernoulli formula :<br>$\frac{\rho}{2} \left( \left(\frac{1}{A_2}\right)^2 - \left(\frac{1}{A_1}\right)^2 \right) \cdot U_g^2$                    | - Converging glottis :<br>$\frac{\rho}{2} \frac{U_g^2}{A_2^2} \left(1 + K_d \left(1 - \frac{A_1^2}{A_2^2}\right)\right) - \frac{\rho}{2} \frac{U_g^2}{A_1^2}$<br>- Diverging glottis :<br>- $h_1 < h_2$ . Flow separation :<br>0<br>- $h_1 = O(h_2)$ no separation :<br>$\frac{\rho}{2} \left( \left(\frac{1}{A_2}\right)^2 - \left(\frac{1}{A_1}\right)^2 \right) \cdot U_g^2$ |                |
| Second Channel | Fully developed flow (see First channel)  | - Separated flow :<br>or<br>- Undeveloped flow (see First channel)<br>or<br>- Development + Fully developed flow (unlikely)   |                |
| Outlet         | Flow separation and reattachment :<br>$\frac{\rho}{2} \left(\frac{U_g}{A_2}\right)^2 \cdot 2 \cdot \frac{A_2}{A_L} (A_L - 1)$ (pressure recovery) | Flow separation and no pressure recovery :<br>0   |                |

**Table 2** : Comparison of pressure differences predicted by classical theory and the modified one.



Compared to the classical formulation [2] presented in section II, the main differences are :

- The head loss at the entrance is neglected (no vena contracta effect),
- Introduction of flow separation in various parts of the glottis,
- Description of flow development within the first and the second channel.

Introduction of the first and last modifications in the model is easy due to the analytical formulation (equations (16) and (24)). We now need some information about the possible importance of flow separation.

### **III.3. Introduction of flow separation in the two-mass model.**

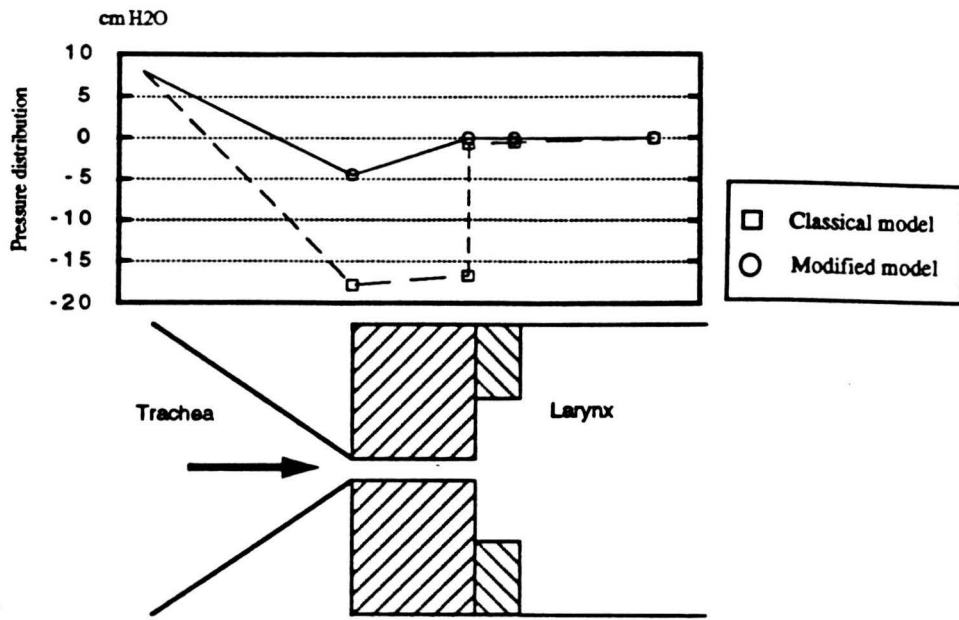
In order to evaluate the possible effects and the relative importance of flow separation within the glottis, we made a modified version of our two-mass model introducing the possibility of flow separation within the second channel. Our purpose is not, for the moment, to produce more realistic sounds but just to have an idea of the possible importance of such effect. Condition for flow separation can be expressed using Reynolds number or critical ratios  $h_2/h_1$ .

Because of the lack of knowledge about reattachment positions, we only consider the cases of fully separated flow through the second channel. In other words, we assume that when the flow separates from the wall, it never reattaches.

In the following, we will denote as "classical model" the original 2-mass model as described in section II. The "modified model" will refer to our version of the 2-mass model in which effects of separation are included.

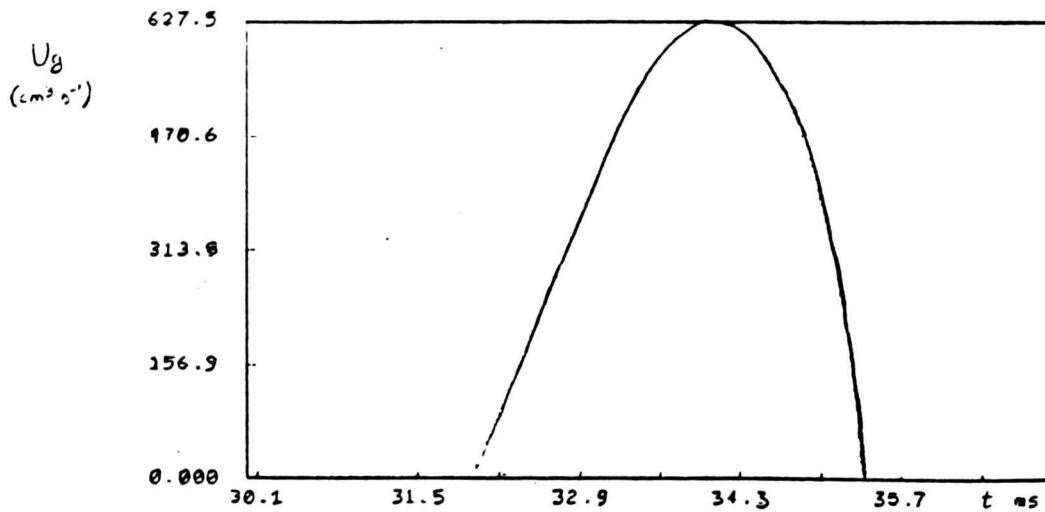
#### **III.3.1. Discussion about the relative importance of separation process.**

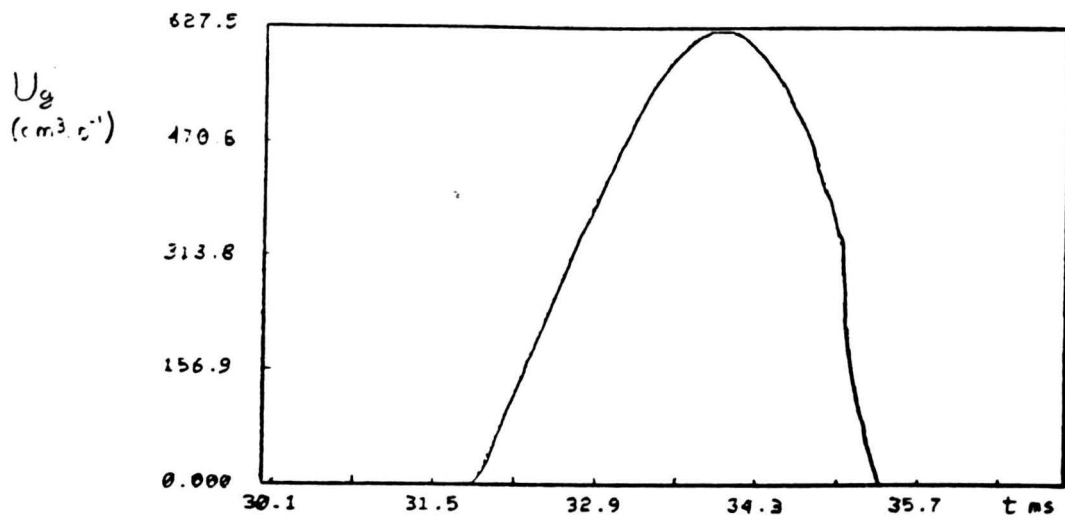
Figure 36 is an example of comparison between the pressure distribution within the glottis calculated with the original program and with the modified one. The condition for occurrence of separation is  $h_2/h_1 > 3$  and  $Re = u \cdot h_1 / \nu > 300$ .



**Figure 36** : Distribution of pressure within the glottis. Comparison between classical model and modified version.

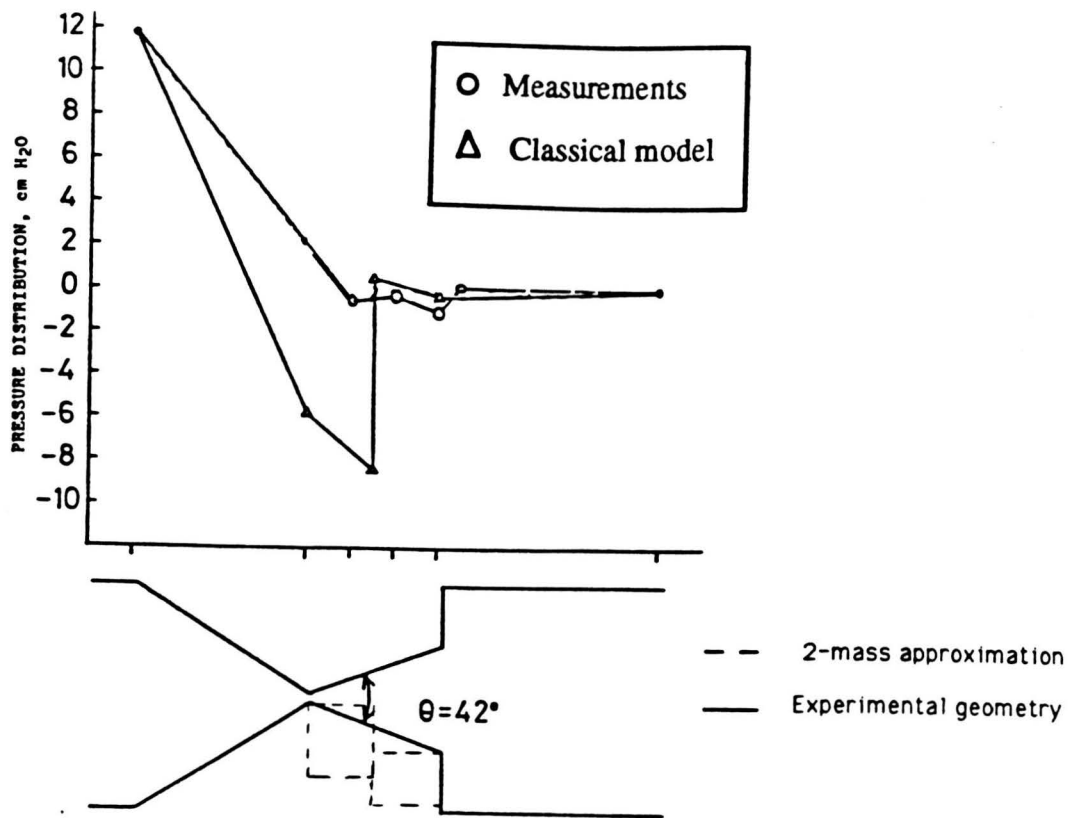
The differences are thus quantitatively important. Moreover, Figure 37 presents a comparison of the volume velocity waveforms calculated with and without separation process.





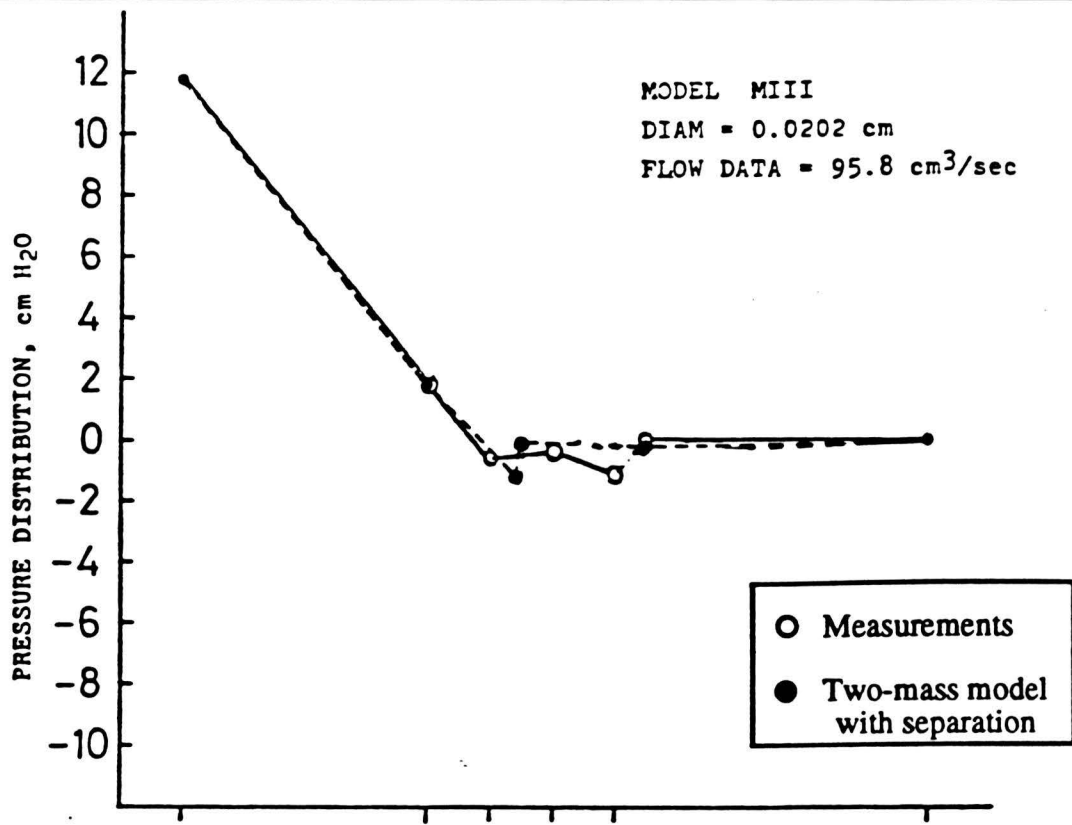
**Figure 37** : Comparison of volume velocity waveforms computed with classical model and with modified version.

Since the differences were found to be effective on both the pressure distribution and the volume velocity, we can expect flow separation to be an important phenomenon. This conclusion is also supported by recent experiments on models of the vocal folds [4]. The results showed that the flow prediction of the two mass model could be quite inaccurate especially for closing time (for a diverging glottis). Figure 38 presents results from measurements derived by Sherer compared to the results simulated with the two mass model.



**Figure 38** : Comparison of pressure distribution computed using the two-mass model and measured on a model (adapted from [4]).

Ishizaka [45], argued that such differences could be explained, and empirically taken into account, by considering additional losses due to the abrupt discontinuity formed by the two masses at closure. In such a case he introduced flow separation effects as presented in the preceding section and assumed reattachment in all cases. Ishizaka introduced empirical factors to take into account this phenomenon and fitted well some results of Sherer (figure 39).

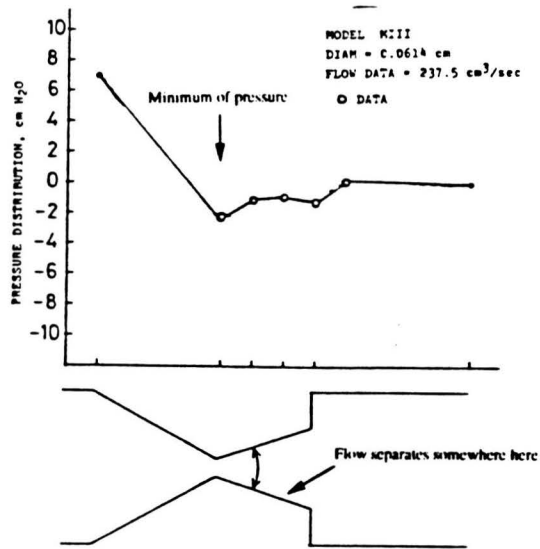


**Figure 39** : Comparison of measured pressure distribution and prediction of modified two-mass model including separation-reattachment in the second channel.(from [44])

These results also tend to prove the importance of flow separation, at least from a fluid dynamical point of view.

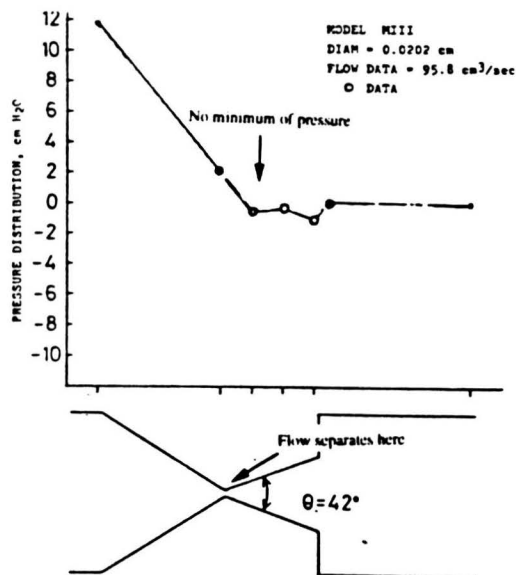
### III.3.2. Unsteady separation.

The results of Scherer et al. also evidence the unsteadiness of the flow separation position. For wide apertures, the separation clearly occurs at the minimum constriction point, that is at the entrance of the model (figure 40).



**Figure 40** : Measured pressure distribution related to the separation position (from [4]).

As the aperture is made narrower (the 0.02 cm case), due to the high viscosity effects, the separation occurs further downstream as shown on figure 41.



**Figure 41** : Interpretation of one result from Scherer from the point of view of flow separation.

Any attempt to describe this effect using the two-mass model is likely to be poor. Indeed, the two-mass model uses a step-wise approximation in which we assume an abrupt discontinuity to be able to describe the flow modifications due to the varying shape of the glottis. As we showed, flow separation can be introduced as a steady process (ocuring at a given position). It can be easily understood that in order to introduce more details such as unsteady flow separation, the step-wise approximation is no longer adequate.

Since the separation position determines the Bernoulli force acting on the vocal folds, we expect unsteady flow separation to be important for phonation.

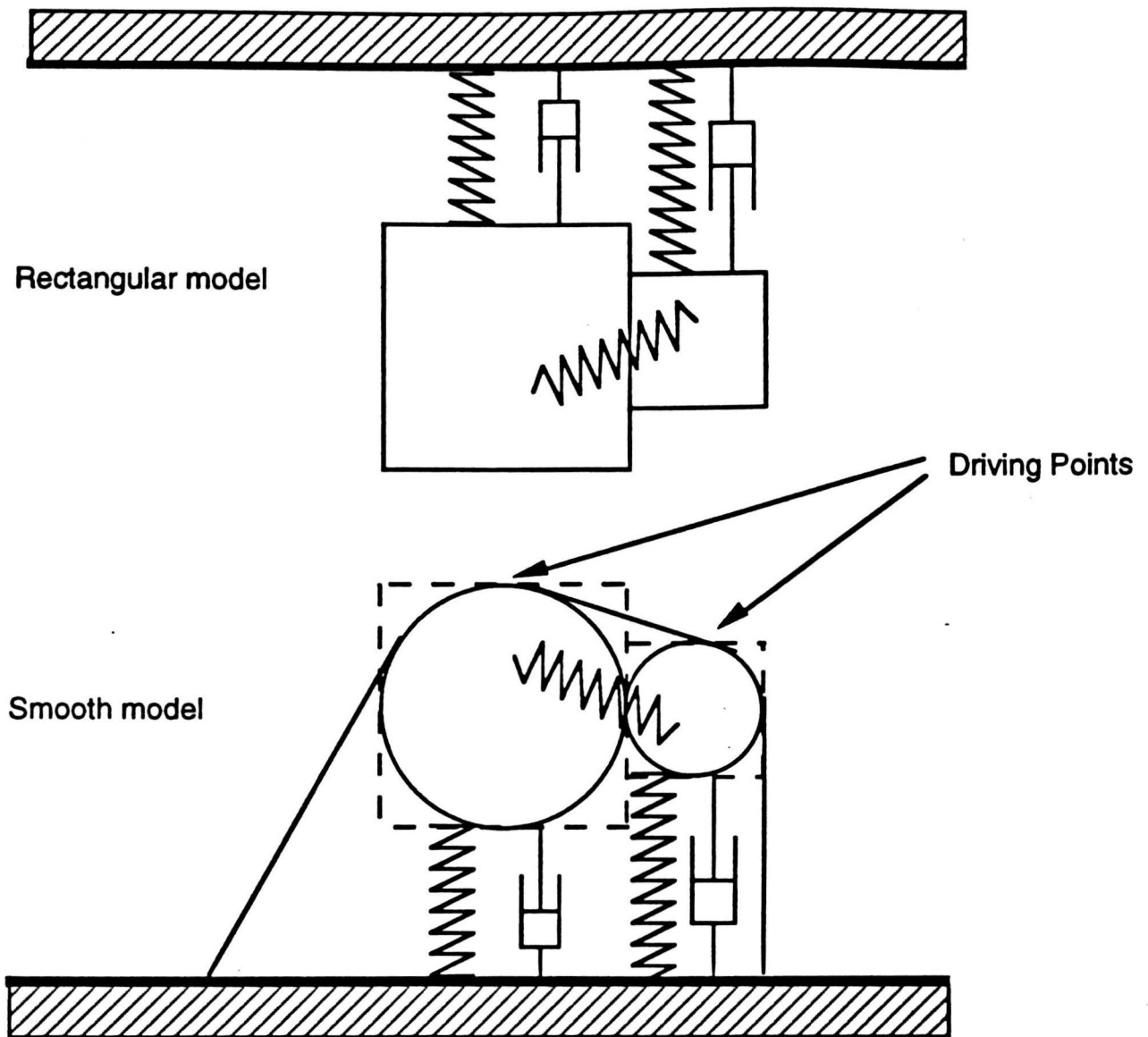
Our longer term research is thus to modify the two-mass model in order to take into account this effect. A possible approach, suggested by Auregan [47], is to consider a model still mechanically governed by the two-mass approximation but having a smooth shape instead of a rectangular one.

Several solutions are possible to derive a smooth shape for the model. The most simple one is to consider the original position of the masses as two parameters controlling a fitting curve.

Different fits were tried. Among them :

- Splines interpolation,
- Polynomial, exponential fitting,
- Analytic solution.

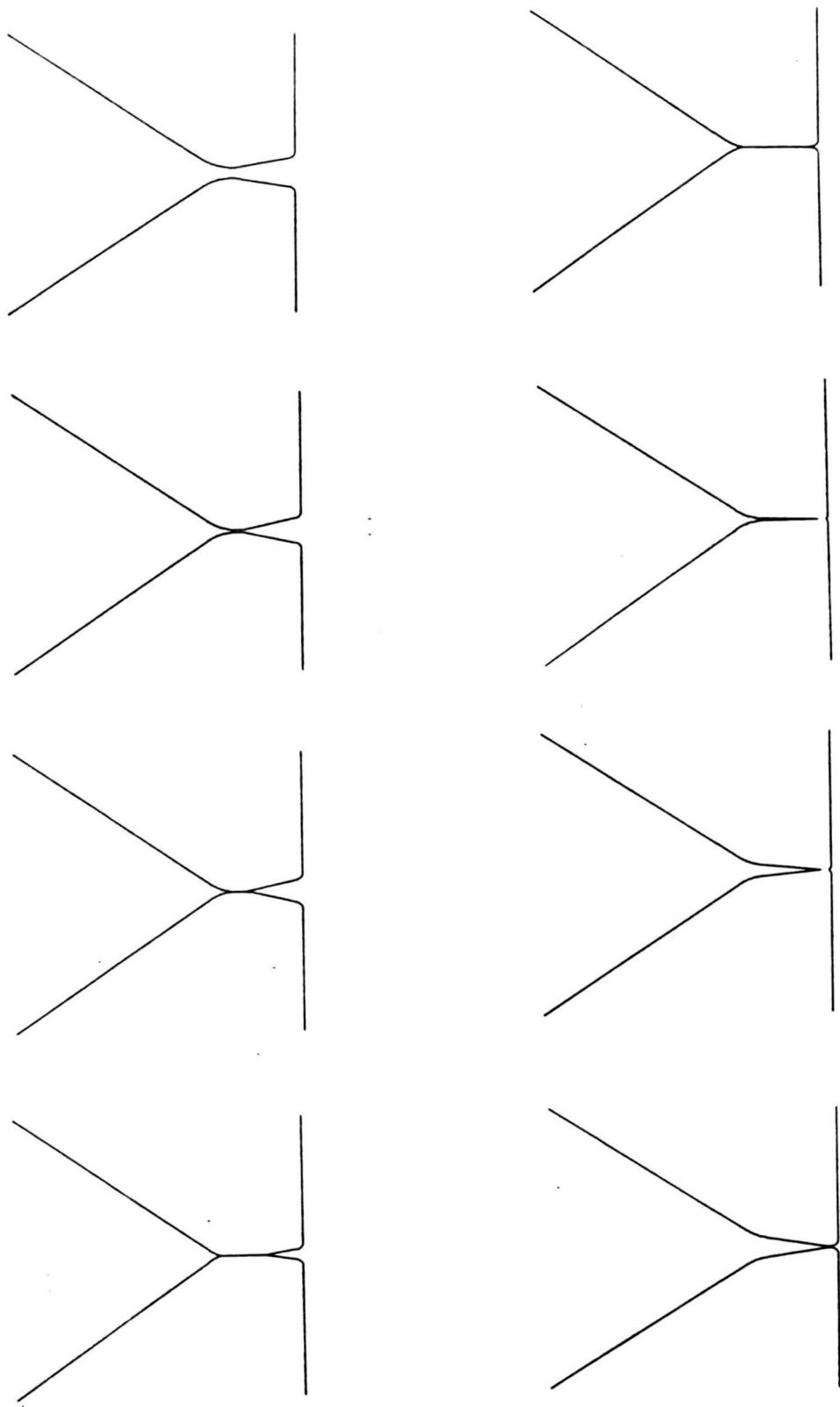
Both splines and approximate fittings were found unacceptable (i.e : led to unrealistic shapes) because of the very few constraints. Thus, an analytic solution seems more wise. Figure 42 present an example of the model.



**Figure 42** : Geometrical description of the vocal folds from the two-mass model.

In this model, the two mass are now represented as two cylinders. Connexions between cylinders and between cylinders and larynx or trachea are approximated by tangent surfaces. The analytical formulation can be found in appendix B. Figure 43 presents a sketch of one vibration cycle obtained with such model. In all cases the patterns are comparable with the real shape of the glottis.





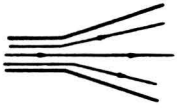
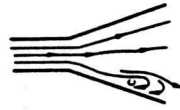



**Figure 43:** Sketch of vibration of a two-mass model with smooth shape.

### III.3.3. Conclusions and further research.

We showed in this section that flow separation within the glottis is an important effect and thus should be taken into account. This conclusion is supported by our simulation results as well as by experimental results of other researchers. Because separation appears not to be a steady process, the two-mass approximation using a rectangular shape is inefficient. Thus it seems necessary to describe more accurately the real shape of the glottis.

A smooth model will also present other advantages for instance :

- Possibility of a more precise description of the collision of the vocal folds (introduction of additional flow at closure, biomechanical description of the collision process...).
- Possibility for describing more complex phenomena. For instance, in a diverging channel, such as the glottis at closure, flow separation is not the only alternative to a Bernoulli flow as shown in table 3.

| Description  | Characteristics  |
|--|--|
| <p>1. Unstalled</p>               | <p>Flow follows diffuser contours.<br/>Flow is steady.</p>   |
| <p>2. Appreciable Stall</p>       | <p>Flow generally follows diffuser contours. Boundary layers thicken. Small regions of separation and erratic flow are generally first seen in corners and they occupy no more than 1/5 of diffuser wall. There is little or no reverse flow.</p>  |
| <p>3. Large Transitory Stall</p>  | <p>Flow is erratic with gross oscillation of pressure and overall flow pattern. Stalled regions with reverse flow form and then wash out.</p> <p><math>N/W_1 &lt; 4</math>: stalls occur on 1 diverging wall.</p> <p><math>4 &lt; N/W_1 &lt; 12</math>: stalls occur on both diverging walls (shown).</p> <p><math>N/W_1 &gt; 16</math>: stalls occur on parallel walls.</p> |
| <p>4. Fully Developed Stall</p>   | <p>Flow separates near throat and forms a large, stable, fixed eddy along one diverging wall while the flow follows second diverging wall. Near-steady flow with reverse flow in eddy. Eddy can be moved from one wall to the other wall only by large disturbances.</p>   |
| <p>5. Jet Flow</p>                | <p>Incoming flow separates from both diffuser walls near throat and proceeds as a jet down diffuser. Large fixed eddies form on diverging walls. Flow is steady with substantial regions of reverse flow. Diffuser pressure recovery is very poor.</p>   |

**Table 3** : Description of flow through diffusers. (from [48])

This indicates, if necessary, the need for experiments combined with theoretical research. Thus, our future work will start with experiments, such as flow visualization and dynamic flow measurements.

This set of experiments is first intended to provide insight about the occurrence and the importance of fluid dynamical effects. For instance, the literature often refer to a vena contracta effect at the inlet of the glottis. As the matter of fact, in the original two-mass model a considerable part of the losses (and thus of the forces acting on the vocal folds) is assumed to be caused by such effect. From a fluid dynamical point of view, experience showed that, for non-abrupt smooth rounded entrance, there are almost no losses at all.

Flow visualization, applied smooth models, should thus provide an answer to this important contradiction. Furthermore, flow separation occurrence could be qualitatively studied and lead us to an idea about the influence of characteristics of the flow upon separation and separation position.

## Conclusion

The two-mass model seems a reasonable first approximation of the flow through the glottis. The computer program we derived is indeed able to produce acceptable voiced sounds. However, we tried to show that several effects, susceptible to improve the model, were not taken into account. Introduction of these important effects is quite limited by the shape of the glottis assumed in the model. For instance, flow separation can only occur at a given position. An other example was shown when studying the collision of the masses. We demonstrated that a better description required a modification of the shape. As a starting point, we proposed a simple, but realistic, modified description of the vocal folds.

Our further research is intended to provide theoretical and experimental insight into flow details. Stability of the flow seems crucial and will be first studied. Furthermore, the separation process as well as the various losses along the glottis will, of course be studied. They will be measured and observed in models. According to these results, our goal is to develop and validate a new numerical model mechanically equivalent to the two-mass model, but using a smooth shape for the fluid dynamical description.

## APPENDIX A : Fluid dynamical equations

### A.1. Basic equations [44]

Let us first recall the dynamic equations for a newtonian incompressible fluid:

Mass conservation :

$$\vec{\nabla} \cdot \vec{V} = 0 \quad (\text{A.1})$$

Momentum conservation (Navier-Stokes equation) :

$$\frac{\partial \vec{V}}{\partial t} + \vec{V} \cdot \vec{\nabla} \vec{V} = -\frac{1}{\rho} \vec{\nabla} P + \nu \nabla^2 \vec{V} \quad (\text{A.2})$$

where  $\vec{V}$  is the fluid velocity vector,  
 $\nu$  is the kinematic viscosity,  
 $\rho$  is the fluid density,  
 $P$  is the static pressure.

Let  $u(x,y,z,t)$ ,  $v(x,y,z,t)$  and  $w(x,y,z,t)$  be the components of vector  $\vec{V}$ . Equations (A.1) and (A.2) can be rewritten in cartesian coordinates as :

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad (\text{A3})$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial x} + \nu \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) \quad (\text{A4})$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial y} + \nu \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} \right) \quad (\text{A5})$$

$$\frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} = -\frac{1}{\rho} \frac{\partial p}{\partial z} + \nu \left( \frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2} \right) \quad (\text{A6})$$

## A.2. Bernouilli theorem:

Let us now consider the special case of a frictionless fluid ( $\nu = 0$ ). The equation (A2) can be rewritten as :

$$\frac{\partial \vec{V}}{\partial t} + (\vec{V} \cdot \nabla) \vec{V} = -\frac{1}{\rho} \nabla P \quad (\text{A7})$$

Using the vectorial relationship :  $\vec{V} \cdot \nabla \vec{V} = \nabla \frac{V^2}{2} + (\nabla \times \vec{V}) \times \vec{V}$ , where  $\times$  is the vectorial product and  $V$  the modulus of vector  $\vec{V}$ , equation (A7) becomes :

$$\frac{\partial \vec{V}}{\partial t} + \nabla \left( \frac{V^2}{2} + \frac{P}{\rho} \right) + (\nabla \times \vec{V}) \times \vec{V} = 0 \quad (\text{A8})$$

If the flow is irrotational ( $\nabla \times \vec{V} = 0$ ), we can define a potential  $\phi$  so that  $\vec{V} = \nabla \phi$  and equation (A8) can be rewritten as :

$$\nabla \left( \frac{\partial \phi}{\partial t} + \frac{V^2}{2} + \frac{P}{\rho} \right) = \vec{0} \quad (\text{A9})$$

In the special case where the unsteady term  $\frac{\partial \phi}{\partial t}$  can be neglected (quasi-steady flow), we obtain the well known form of the Bernouilli theorem for a steady incompressible potential flow :

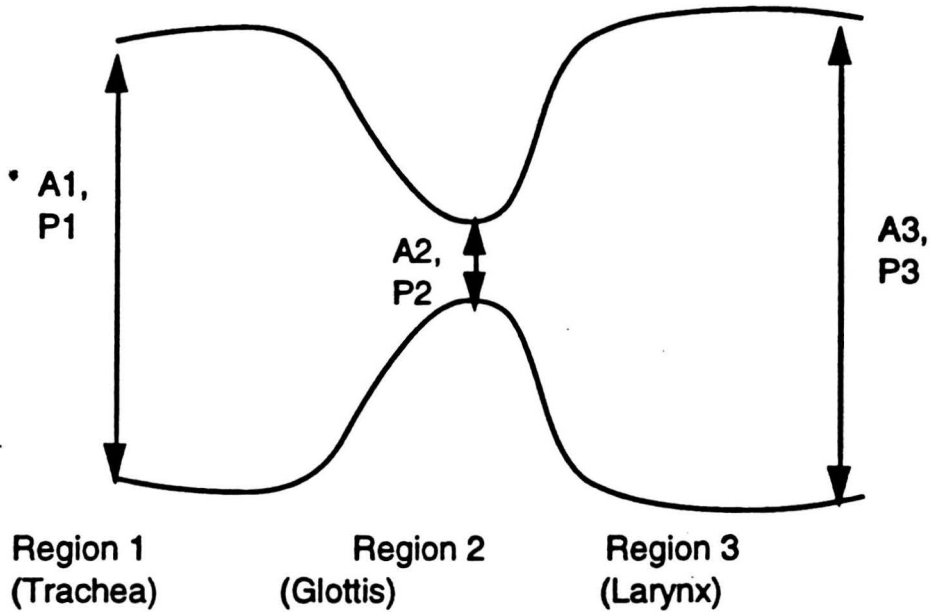
$$\nabla \left( \frac{V^2}{2} + \frac{P}{\rho} \right) = \vec{0} \quad (\text{A10})$$

or equivalently :

$$\frac{V^2}{2} + \frac{P}{\rho} = \text{constant} \quad (\text{A11})$$

Equation (A11) indicates that if the velocity increases, pressure decreases and *vice versa*.

As an example of application of theorem (A11), let us consider the flow in the varying channel (similar to the glottis) shown figure A1.



**Figure A1** : Glottal-like channel.

Because of the continuity of the flux, we have the relationship :  
 $A_2.V_2 = A_1.V_1 = A_3.V_3 = U_g$  (A12)

Where :

$A_1$ ,  $A_2$  and  $A_3$  are the respective cross sectional areas of region 1, 2 and 3,  
 $V_1$ ,  $V_2$  and  $V_3$  the fluid velocity in section 1, 2 and 3.

Applying Bernouilli theorem between section 1 and 2 and between section 2 and 3, we can see that a minimum of pressure (or a maximum of velocity) occurs at the contraction 2.

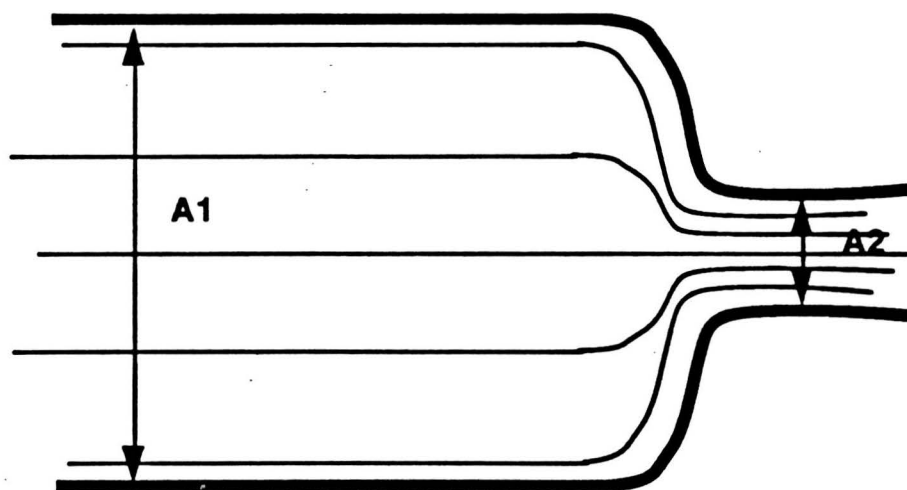
Let us now consider the pressure variation between the whole glottis, that is between section 1 and 3. The Bernouilli theorem (A11) together with relation (A12) leads to the following relationship :

$$P_3 - P_1 = \frac{\rho}{2} (V_1^2 - V_3^2) = \frac{\rho}{2} U_g^2 \left( \frac{1}{A_1^2} - \frac{1}{A_3^2} \right) \quad (\text{A13})$$

In the special case where  $A_1 = A_3$ , which is roughly the case for the glottis, we find a zero pressure difference and hence no variations for the velocity whatever the shape of the contraction between section 1 and 3 is. This constitutes a special case of the well known D'Alembert's paradox and indicates that more details about the flow, especially viscosity effects, must be taken into account in order to describe the volume flow control by a change of geometry of the glottis.

### A.3. Pressure loss due to a sudden contraction.

We examine here the effects of a sudden contraction for the flow as encountered at the entrance of the glottis. The pressure loss due to this abrupt contraction can be evaluated in first approximation using Bernouilli theorem, assuming the flow to follow the walls (figure A2).



**Figure A2** : Bernouilli flow at the entrance of the glottis.



Applying Bernoulli theorem, the pressure drop can be written as :

$$P_1 - P_2 = \frac{\rho}{2} \cdot U_g^2 \cdot \left( \frac{1}{A_2^2} - \frac{1}{A_1^2} \right) \quad (A14)$$

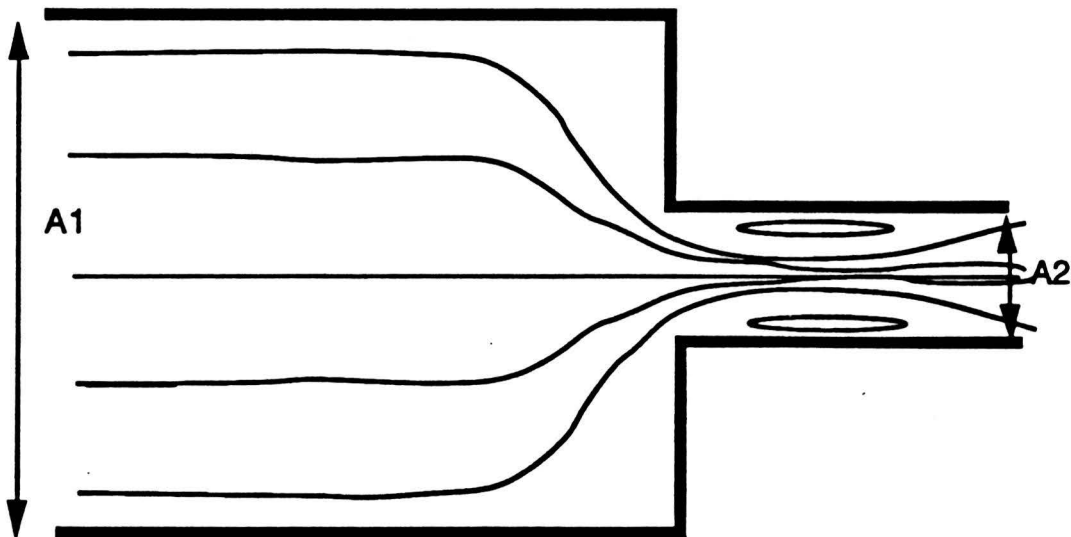
In reality, the actual pressure difference does not equal the theoretical Bernoulli law. Losses will occur to viscosity effects downstream and mainly upstream of the contraction. Taking in account these losses, we rewrite equation (A14) in the modified form :

$$P_1 - P_2 = \frac{\rho}{2} \cdot U_g^2 \cdot \left( \frac{1}{A_2^2} - \frac{1}{A_1^2} \right) + \frac{\rho}{2} \cdot K \cdot \frac{U_g^2}{A_2^2} \quad (A15)$$

where  $K$  is the pressure loss coefficient.

In the special case of sharp edged abrupt contraction (figure A3), empirical results gave the following values for  $K$  as a function of  $A_2/A_1$  [49] in the case of circular pipes:

|           |      |      |      |      |      |      |      |      |      |   |
|-----------|------|------|------|------|------|------|------|------|------|---|
| $A_2/A_1$ | 0.1  | 0.2  | 0.3  | 0.4  | 0.5  | 0.6  | 0.7  | 0.8  | 0.9  | 1 |
| $K$       | 0.37 | 0.35 | 0.32 | 0.27 | 0.22 | 0.17 | 0.10 | 0.06 | 0.02 | 0 |



**Figure A3** : Abrupt contraction in a channel

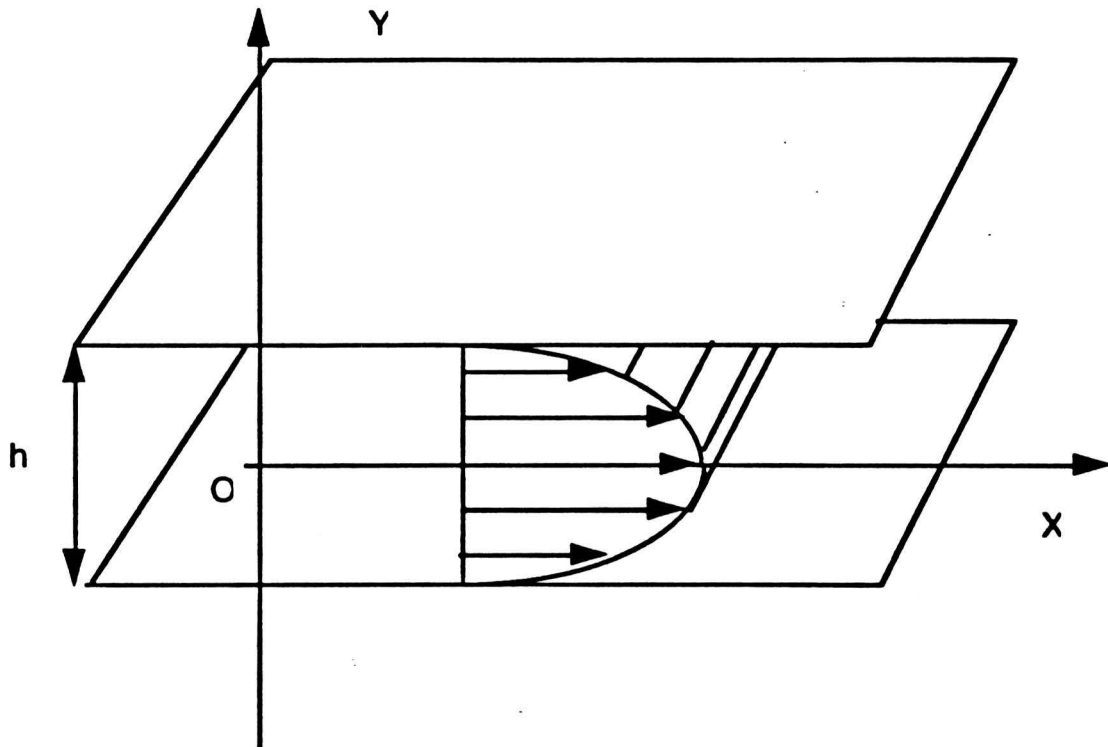
As can be seen the losses can be important. They are due mainly by a flow separation at the entrance of the contraction (vena contracta). It is interesting to note that using  $A_1 = 1$  cm,

$A_2 = 0.1$  cm (typical values for the trachea and the glottis) we find a value of  $K = 0.37$  which is comparable to the value 0.375 reported by van den Berg but for a rectangular shape [35].

However, the real shape of the vocal folds is not sharp, as well as the contraction is not abrupt. In this case, experiments showed very small values of  $K$  (order of 0.04 for a circular shape). More precisely, the losses depend upon the length of the contraction, the radius of curvature of the entrance as well as the areas  $A_1$  and  $A_2$ . According to typical values for the vocal folds empirical tables [48] predict almost no pressure losses ( $K \ll 1$ ).

#### A.4. Parallel flow through a bidimensional channel.

Let us now consider the steady incompressible flow in a bidimensional channel, as in our approximation of the glottis (figure A4).



**Figure A4** : Parallel flow in a bidimensional channel.

Let  $h$  be the height of the channel.

Since we assume the flow to be parallel to the walls, the components of the velocity  $v$  and  $w$  are zero everywhere. The mass conservation equation reduces to :

$$\frac{\partial u}{\partial x} = 0 \quad (\text{A16})$$

expressing that the component  $u$  does not depend upon  $x$ . The Navier Stockes equations take the form :

$$-\frac{1}{\rho} \frac{\partial P}{\partial x} + \nu \left( \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} \right) = 0 \quad (\text{A17})$$

$$\frac{1}{\rho} \frac{\partial P}{\partial y} = 0 \quad (\text{A18})$$

$$\frac{1}{\rho} \frac{\partial P}{\partial z} = 0 \quad (\text{A19})$$

From equations (A18) and (A19) we deduce that  $P$  is a function of  $x$  only. Furthermore, by taking the  $x$ -derivative of (A17) and using (A16), we can show that  $P$  is a linear function of  $x$  :

$$\left( \frac{d^2 P}{dx^2} = 0 \right)$$

Very generally, we considering the flow trough a channel one has to distinguish between two different regions :

- The first one is called the frictionless region in which the flow can be describe by potential flow equations.

- The second one, in the vicinity of the walls, is called the boundary layer. In this region the effects of viscosity are predominant and the velocity is considerably lower than in the frictionless region. The thickness of the boundary layer is an increasing function of the  $x$  coordinate.

Thus, starting from an initial profile (uniform), as the thickness of the boundary layer increases, the velocity distribution develops to a parabolic shape (Poiseuille's law). At a sufficiently large distance from the entrance region the entire flow is governed by viscosity effects and the velocity profile becomes parabolic (Fully developed Poiseuille flow) and do not depend upon the  $x$  coordinate any longer. For such a flow, the velocity  $u$  only depends upon  $y$ , the Navier Stokes equations (A17), (A18), (A19) will reduce to the simple form:

$$\mu \cdot \frac{\partial^2 u}{\partial y^2} = \frac{dP}{dx} \quad (\text{A20})$$

where  $\mu = \rho \nu$  is the dynamic viscosity.

With the boundary conditions :

$$u = 0 \text{ for } y = \pm \frac{h}{2} \quad (\text{A21})$$

Since  $u=u(y)$  and  $P=P(x)$ , each member of the equation has to be constant. The solution of this equation, with respect to the boundary conditions (A21) is :

$$u(y) = \frac{1}{2\mu} \frac{dP}{dx} \left( y^2 - \frac{h^2}{4} \right) \quad (\text{A22})$$

The volume velocity  $U_g$  can be derived by integrating  $u(y)$  over  $y$  :

$$U_g = - \frac{h^3 \cdot L_g}{12 \cdot \mu} \frac{dp}{dx} \quad (\text{A23})$$

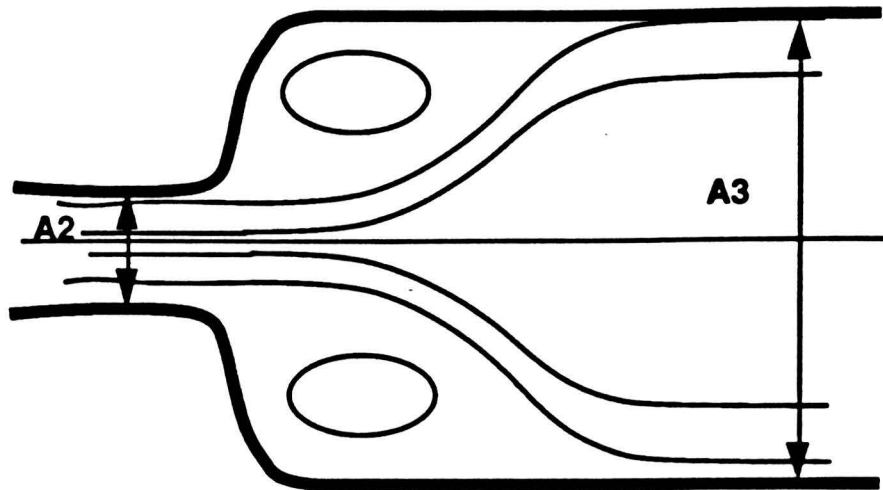
The pressure drop over a distance  $d1$  in the  $x$  direction will be :

$$P_2 - P_1 = - \frac{12 \cdot \mu \cdot d1}{h^3 \cdot L_g} U_g \quad (\text{A24})$$

This equation indicates that the pressure loss due to viscosity effects is proportional to the volume velocity and to the third power of the channel's aperture.

### A.5. Pressure loss due to a sudden expansion.

We now turn to the case of a sudden expansion as occurs at the transition from the glottis to the larynx (figure A5).



**Figure A5** : Sudden expansion of the larynx.

Because the flow at the end of the glottis will separate from the walls, energy will be dissipated due to formation of vortices. The Bernoulli law is thus not valid.

If we assume the pressure in the region A3 - A2 to be P2 (this was verified experimentally), the conservation of the momentum provides, neglecting wall friction, the following relationship :

$$\rho \cdot A_3 \cdot V_3^2 - \rho \cdot A_2 \cdot V_2^2 + P_3 \cdot A_3 - P_2 \cdot A_2 = P_2 \cdot (A_3 - A_2) \quad (A25)$$

and thus, using mass conservation :  $A_2 \cdot V_2 = A_3 \cdot V_3$ , we find :

$$P_2 - P_3 = \frac{\rho}{2} \frac{U_g^2}{A_2^2} \cdot 2 \cdot \frac{A_2}{A_3} \cdot \left( \frac{A_2}{A_3} - 1 \right) \quad (A26)$$

A pressure recovery coefficient can be defined in a similar way as (A24) and values here :

$$K = \left( \frac{A_2}{A_3} - 1 \right)^2 \quad (A27)$$

Note that for the case of the larynx, since  $A_3 \gg A_2$ , the pressure recovery coefficient is very small.

## APPENDIX B : Calculation of new vocal cords shape.

### B.1. Notations :

Let  $f$  be any scalar quantity, we define the quantity  $\Delta_{ij}$  as :

$$\Delta_{ij}f = f_j - f_i \quad (\text{B.1})$$

### B.2. Preliminary results :

Let  $C1(O1, R1)$  be the circle centered on  $O1(x_{o1}, y_{o1})$  with a radius of  $R1$ . It can be easily shown that any line tangent to this circle has the following equation :

$$y = mx + y_{o1} - mx_{o1} + R1\sqrt{m^2 + 1} \quad (\text{B.2})$$

where the slope  $m$  can be any real value.

Let  $C2(O2, R2)$  be the circle centered on  $O2(x_{o2}, y_{o2})$  with a radius of  $R2$ . Any line tangent both to  $C1$  and  $C2$  has to verify the following system :

$$\begin{aligned} y &= mx + y_{o1} - mx_{o1} + R1\sqrt{m^2 + 1} \\ y &= mx + y_{o2} - mx_{o2} + R2\sqrt{m^2 + 1} \end{aligned} \quad (\text{B.3})$$

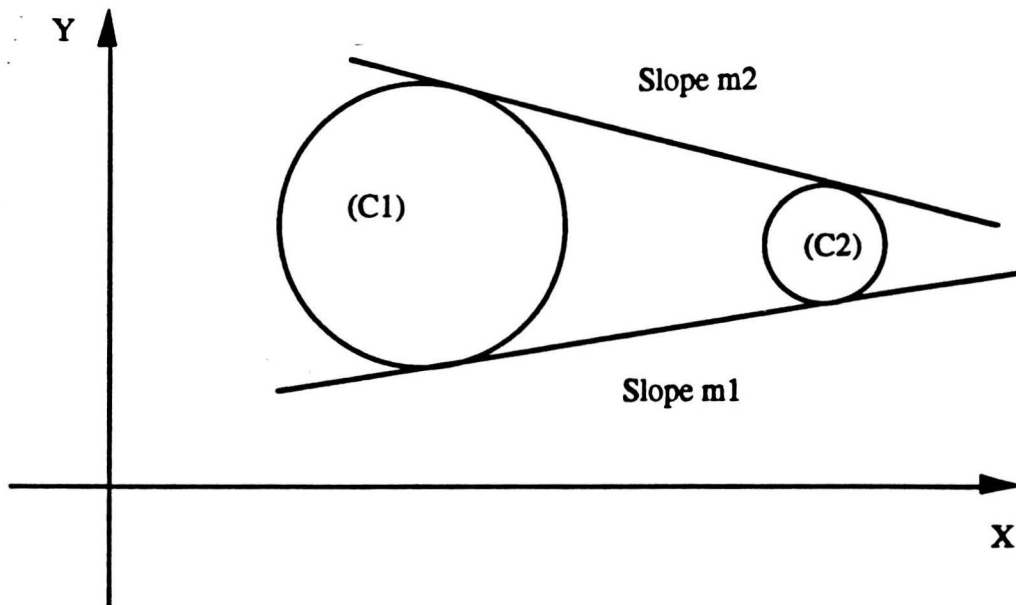
Thus,  $m$  has to verify :

$$m^2(\Delta_{12}^2 R - \Delta_{12}^2 x) + 2m\Delta_{12}x \cdot \Delta_{12}y + \Delta_{12}^2 R - \Delta_{12}^2 y = 0 \quad (\text{B.4})$$

if  $O1 \neq O2$  it can be shown that this equation has two solutions :

$$\begin{aligned} m_1 &= \frac{-\Delta_{12}x \cdot \Delta_{12}y + \Delta_{12}R \sqrt{\Delta_{12}^2 x + \Delta_{12}^2 y - \Delta_{12}^2 R}}{\Delta_{12}^2 R - \Delta_{12}^2 x} \\ m_2 &= \frac{-\Delta_{12}x \cdot \Delta_{12}y - \Delta_{12}R \sqrt{\Delta_{12}^2 x + \Delta_{12}^2 y - \Delta_{12}^2 R}}{\Delta_{12}^2 R - \Delta_{12}^2 x} \end{aligned} \quad (\text{B.5})$$

corresponding to the two situations depicted in figure B1.



**Figure B1** : Interpretation of the two solution of equation (B.4)

In our case, we will only consider the line of lower slope namely :

$$m2 = \frac{-\Delta_{12}x \cdot \Delta_{12}y - \Delta_{12}R \sqrt{\Delta_{12}^2x + \Delta_{12}^2y - \Delta_{12}^2R}}{\Delta_{12}^2R - \Delta_{12}^2x} \quad (B.6)$$

Special case :

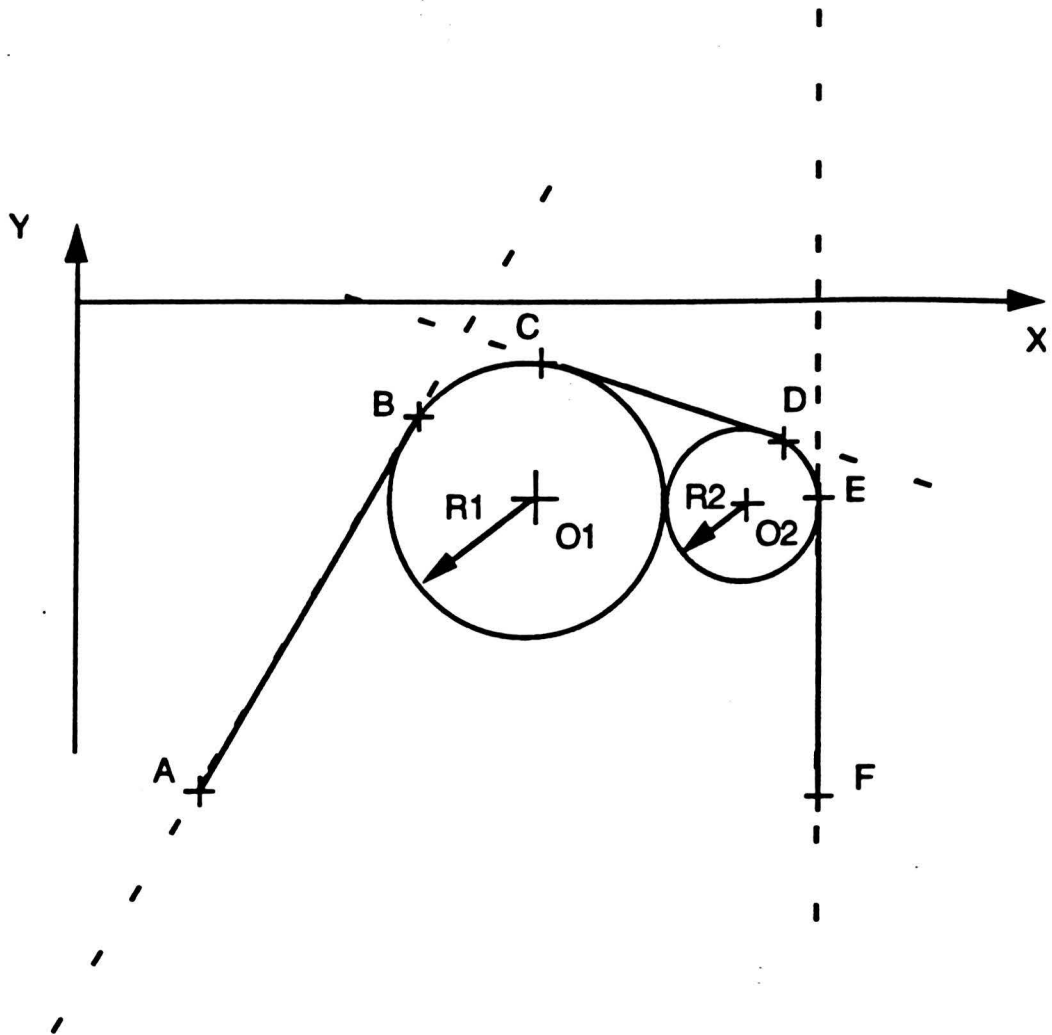
When  $R2 = 0$  the problem leads to a line tangent to  $C2$  and including  $O2$ . Noting that, in this case  $\Delta_{12}R = -R_1$  the equation is found to be :

$$m2 = \frac{-\Delta_{12}x \cdot \Delta_{12}y + R_1^2 \sqrt{\Delta_{12}^2x + \Delta_{12}^2y - R_1^2}}{R_1^2 - \Delta_{12}^2x} \quad (B.7)$$

### B.3. Description of a smooth model of the vocal cords.

The general aspect is recalled in figure B2.





**Figure B2** : Smooth model for the vocal folds.

The model can be divided into five parts :

1) Line from  $A(x_A, y_A)$  to  $B(x_B, y_B)$  and tangent to C1 :

The equation of this line is (using (B.2)) :

$$y = m_{AB}x + y_{o1} - m_{AB}x_{o1} + R1\sqrt{m_{AB}^2 + 1}$$

$m_{AB}$  is given by (B.7) :

$$m_{AB} = \frac{-\Delta_{ABx} \cdot \Delta_{ABY} - R1 \sqrt{\Delta_{ABx}^2 + \Delta_{ABY}^2 - R1}}{R1 - \Delta_{ABx}^2}$$

2) Part of a circle from B(x<sub>B</sub>, y<sub>B</sub>) to C(x<sub>C</sub>, y<sub>C</sub>):

Equation of the circle :

$$(x - x_{01})^2 + (y - y_{01})^2 = R1^2$$

with x<sub>B</sub> < x < x<sub>C</sub>.

3) Line tangent to both C1 and C2 :

The general equation is :

$$y = m_{CD}x + y_{01} - m_{CD}x_{01} + R1 \sqrt{m_{CD}^2 + 1}$$

From equation (B.6) the slope of the line is found to be :

$$m_{CD} = \frac{-\Delta_{CDx} \cdot \Delta_{CDY} - \Delta_{12}R \sqrt{\Delta_{CDx}^2 + \Delta_{CDY}^2 - \Delta_{12}^2R}}{\Delta_{12}^2R - \Delta_{CDx}^2}$$

4) Part of a circle from D(x<sub>D</sub>, y<sub>D</sub>) to E(x<sub>E</sub>, y<sub>E</sub>):

$$(x - x_{02})^2 + (y - y_{02})^2 = R2^2$$

with x<sub>D</sub> < x < x<sub>E</sub>.

5) Line from E(x<sub>E</sub>, y<sub>E</sub>) to F(x<sub>F</sub>, y<sub>F</sub>) tangent to C2 :

Equation form :

$$y = m_{EF}x + y_{02} - m_{EF}x_{02} + R2 \sqrt{m_{EF}^2 + 1}$$

With slope given by equation (B.7) :

$$m_{EF} = \frac{-\Delta_{EFX} \cdot \Delta_{EFY} - R^2 \sqrt{\Delta_{EFX}^2 + \Delta_{EFY}^2} - R^2}{R^2 - \Delta_{EFX}^2}$$

## References.

- [1] EGGEN, J.H. (1992)  
On the Quality of Synthetic Speech. Evaluation and improvements.  
Doctoral Thesis, Techn. Univ. Eindhoven.
- [2] ISHIZAKA K, FLANAGAN J.L (1972)  
Synthesis of Voiced Sounds from a Two-Mass Model of the Vocal Cords.  
B.S.T.J 51 pp 1233-1268.
- [3] AL ANSARI A (1981)  
Etude du fonctionnement et simulation en temps reel d'un modele de la source vocale.  
Thesis INP-Grenoble.
- [4] SCHERER, R.C, TITZE I.R, CURTIS, J.F. (1983).  
Pressure-flow relationships in two-models of the larynx having rectangular glottal shapes.  
J. Acoust. Soc. Am. 73 pp 668-676.
- [5] TEAGER H.M, TEAGER S.M (1985)  
Active Fluid Dynamic voice Production, or there is a unicorn in the garden.  
in Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control ed. by I. R. Titze and R. s. Scherer (The Denver Center for the Performing Arts, Denver, CO),  
pp 387-394.
- [6] BORDEN G.J, HARRIS K.S (1984).  
Speech Science Primer  
Williams and Wilking, Baltimore.
- [7] LADEFOGED P. (1982).  
A course in Phonetics  
Harcourt Brace Jovanovich : New-York.

- [8] HARDCASTLE, J.W (1976)  
Physiology of speech production,  
An introduction for speech scientists.  
Acad. Press, London.
- [9] LIEBERMAN P (1968)  
Vocal cord motion in man.  
Ann. N. Y. Acad. Sci. 115, pp 28-38.
- [10] KOIKE Y. (1980)  
Sub- and supraglottal pressure variation during phonation  
in Vocal Fold Physiology, Ed. by K. N. Stevens and M. Hirano, Tokyo : Univ. of  
Tokyo, pp 181,192.
- [11] van den BERG J. W (1956).  
Direct and indirect determination of the mean subglottic pressure.  
Folia Phoniatica 8 1956, pp 1-24.
- [12] CRANEN B, BOVES L. (1985)  
Pressure measurements during speech production using miniature pressure transducers :  
Impact on models for speech production.  
J. Acoust. Soc. Am. 77 pp 1543-1551.
- [13] LADEFOGED P (1963)  
Some physiological parameters in speech.  
Language and Speech, 6, pp 109-119.
- [14] LADEFOGED P, MCKINNEY N.P (1963)  
Loudness, Sound Pressure and Subglottal Pressure in Speech.  
J. Acoust. Soc. Am. 35 pp 454-460.
- [15] McGLONE R.E, SHIP T. (1970)  
Changes in Subglottal Air Pressure Associated with Changes of Fundamental Frequency  
J. Acoust. Soc. Am. 48 p 118A.

[16] FUJIMURA O.

Of the vocal fold and its phonetic implications.

in Vocal Fold Physiology, Ed. by K. N. Stevens and M. Hirano, Tokyo : Univ. of Tokyo, pp 271-288.

[17] HIRANO M. KURITA S. NAKASHIMA T.

Growth, Development and aging of human vocal folds.

in Vocal Fold Physiology, Ed. by K. N. Stevens and M. Hirano, Tokyo : Univ. of Tokyo, pp 22-43

[18] SAWASHIMA M. HIROSE H. HONA K. YOSHIOKA H. HIBI S. R. KAWASE  
N. YAMADA M.

Stereoendoscopic Measurements of the Laryngeal Structure.

in Vocal Fold Physiology, Ed. by K. N. Stevens and M. Hirano, Tokyo : Univ. of Tokyo, pp 264-276.

[19] HOLLIEN H. (1962)

Vocal Fold Thickness and Fundamental Frequency of Phonation.

J. Speech Hearing Res., 5, pp 237-243.

[20] HOLLIEN H., CURTIS J.F (1960)

A Laminagraphic Study of Vocal Pitch.

J. Speech Hearing Res., 3, pp 361-371.

[21] SONIES B.C, SHWAKER T.H, HALL T.E, GERBER L.H, LEIGHTON S.B  
(1981)

Ultrasonic visualization of tongue motion during speech

J. Acoust. Soc. Am. 70, pp 683-686

[22] HIRANO M. (1977)

Structure and vibratory behaviour of the vocal folds

in Dynamic Aspects of Speech Production. M. Sawashima and F. S. Cooper. Univ. of Tokyo Press. Tokyo, Japan, pp 13-27.

[23] LINDQVIST J. (1970)

The voice source studied by means of inverse filtering.

QPSR Speech trans. Lab., Roy. Inst. Techn. Stockholm, 3-9.

[24] LINDQVIST J. (1965)

Studies of the voice source by means of inverse filtering techniques.

5th Int. Cong. Acoust. Liege A-35.

[25] SUNDBERG J. (1987)

The science of the singing voice.

Northern Illinois University Press., DeKalb.

[26] KLATT D.H, KLATT L.C (1990)

Analysis, synthesis and perception of voice quality variations among female and male talkers.

J. Acoust. Soc. Am. 87, pp 820-857.

[27] FLANAGAN J.L, LANDGRAF L.L. (1967)

Self-oscillating Source for Vocal-Tract Synthesizers.

Proc. IEEE-AFCRL. Symposium on Speech Commun. and Process., Boston Mass.

[28] BOE L.J, GUERIN B (1977)

Etude du fonctionnement d'un modele des cordes vocales a deux masses; determination des parametres de commande et de leurs influences respectives.

Bulletin de l'Institut de Phonetique de Grenoble, 6, pp 1-56.

[29] TITZE, I.R, TALKIN D.T (1979)

A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation.

J. Acoust. Soc. Am. 66 pp 60-74.

[30] TITZE, I.R (1973)

The human vocal cords : A mathematical Model part I.

Phonetica 28 pp 129-170.

- [31] TITZE I. R (1974)  
The Human Vocal Cords : A Mathematical Model part II.  
Phonetica 29, pp 1-21.
- [32] DESCOUT R, AULOGE J, Y, GUERIN B (1980)  
Continuous model of the vocal source.  
Proc. ICA-SSP pp 61-64.
- [33] CONRAD W. A (1969)  
Pressure-Flow Relationships in Collapsible Tubes.  
IEEE Trans. on biomedical Engineering BME 16, pp 284-295.
- [34] CONRAD W. A (1985)  
Collapsible tube model of the larynx.  
in Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control ed. by I. R. Titze and R. s. Scherer (The Denver Center for the Performing Arts, Denver, CO), pp 328-348.
- [35] BERG van den J, W, ZANTEMA J, T, DOORNENBAL Jr, P (1957)  
On the Air Resistance and the Bernoulli Effect of the Human Larynx.  
J. Acoust. Soc. Am. 29, pp 626-631.
- [36] ISHIZAKA K, KANENKO T (1968)  
On the equivalent Mechanical Constants of the vocal cords.  
J. Acoust. Soc. Japan 24 pp 312-313.
- [37] ISHIZAKA K, MATSUDAIRA M (1972)  
Fluid Mechanical Considerations of Vocal Cord Vibration.  
SCRL Monograph No 8.
- [38] GUERIN B., BOE J, L (1977)  
A two mass model for the vocal cords : determination of control parameters and their respective consequences.  
IEEE Int. Conf. on Acoust. , Speech and Signal Processing, Hartford. pp 583-586.



[39] FLANAGAN J.L, ISHIZAKA K (1978)

Computer model to characterize the air volume displaced by the vibrating vocal cords.  
J. Acoust. Soc. Am. 63, pp 1559-1565.

[40] MRAYATI M, GUERIN B (1976)

Etude de l'impedance d'entree du conduit vocal. couplage source-conduit vocal.  
Acustica 35, pp 330-340.

[41] VOGTEN L. L. M (1990)

Syllabus van het college Spraaktechnologie.  
Intitute for Perception Research Repport No 649.

[42] ZON van J (1989)

Stromingsgeïnduceerde klepinstabiliteiten  
Vakgroep Transportfysica R-1024-A. Eindhoven University of Technology.

[43] ZON van J., HIRSCHBERG A., GILBERT J., WIJNANDS A. P. J (1990)

Flow through the reed channel of a single reed music instrument.  
In Colloque de Physique, Supp. au Journal de Physique, Fasc. 2, 51, pp 821-824.

[44] SCHLICHTING H (1968)

Boundary-Layer Theory  
Mc Graw Hill, New York.

[45] ISHIZAKA K (1985)

Air resistance and intra-glottal Pressure in a Model of the Larynx.  
in Vocal Fold Physiology: Biomechanics, Acoustics and Phonatory Control ed. by I. R.  
Titze and R. s. Scherer (The Denver Center for the Performing Arts, Denver, CO),  
pp 414-424..

[46] HIRSCHBERG A. (1992)

Some Fluid Dynamical Aspects of Speech.  
To appear in Bulletin de la Communication Parlee.

[47] AUREGAN Y. (1992)

Personnal communication.

[48] BLEVINS R.D (1984)  
Applied Fluid Dynamic Handbook.  
Van Nostrand Reinhold Co. : New York.

[49] CANDEL S. (1990)  
Mecanique des Fluides  
Dunod Universite, Bordas Paris.