Chapter 4

Numerical simulation of a voice-producing element based on Navier Stokes equations and the Finite Element Method

M.P. de Vries, M.C. Hamburg, A.E.P. Veldman, H.K. Schutte, G.J. Verkerke

Submitted to JASA
Abstract

When the larynx of a patient has been removed surgically to treat laryngeal cancer, voice restoration can result in a rather poor speech quality. To improve the rehabilitation process a voice-producing element has been developed, based on the so-called lip principle, in which the lips behave like the lips of a musician while playing a brass instrument. To optimize the voice-producing element, a numerical model has been developed. In this model, the Finite Element Method (FEM) is used to describe the geometry and mechanical properties of the oscillating part. The flow through the voice-producing element is described by two-dimensional incompressible Navier-Stokes equations. The interaction between lip and airflow is modeled by placing the lip and the aerodynamics in a collective computational grid, and requiring continuity of forces and velocities.

By applying a pressure to the model, air flows through the voice-producing element. In the numerical model, the geometric properties of a single lip have been varied, until pulses comparable to glottal volume velocity waveforms are obtained. These pulses can be considered as artificial glottal waves. By varying several geometric parameters of the voice-producing element, voice quality has been optimized.

To validate this numerical model, an in vitro test with a prototype of the voice-producing element has been performed. The mean speaking fundamental frequency of the sound has been measured as a function of the flow through a prototype of the voice-producing element and compared to the results of the numerical model. Experimental and numerical results show a rather good similarity.

4.1 Introduction

In some of the patients in whom cancer of the larynx is diagnosed, the larynx has been removed surgically. The consequences of this removal are serious, implying amongst others the loss of the vocal folds. Since the first surgical removal of the larynx, over hundred years ago, several methods to replace the voice source have been developed (Mahieu 1988). Nowadays, most patients can reach reasonable voice quality by compressing air in the esophagus and releasing it, causing the mucosa at the esophageal entrance to vibrate (Blom, Singer, and Hamaker 1988). This mechanism of voice production is comparable to belching. A method of transporting air to the esophagus is via a shunt valve between the trachea and the esophagus (see Fig.1), a one-way valve made of silicone rubber that prevents leakage of food and fluids from the upper digestive tract into the airways, (Mahieu 1988).

The basic sound is (as in laryngeal speech production) acoustically converted to speech sound in the pharynx and oral and nasal cavity. However, the source sound has a low mean speaking fundamental frequency $F_0$ (60-80 Hz.), (Cornut, Laganier, and Agnes 1968; Damsté 1958; Snidecor and Curry 1959), whereas values for $F_0$ in laryngeal speech are about 110 Hz for male and 210 Hz for female.
Especially for a female laryngectomized person this low $F_0$ is disturbing. Moreover, speech is often monotonous.

To overcome these drawbacks, a voice-producing element that produces a source sound with a higher $F_0$ and frequency variation during speech is under development. The underlying principle is comparable to the oscillating lips of a musician playing a brass instrument (Adachi and Sato 1996; Sram 1989). In the voice-producing element, only one oscillating lip will be placed because a voice-producing element consisting of one lip is easier and more reproducible to manufacture (Fig. 2). In the neutral position, the lip is slightly pressed against the opposite wall. When air pressure is applied at the inlet of the voice-producing element, the initially closed lip opens and the flowing air starts the lip to oscillate as result of aerodynamic and mechanic forces that occur in the lip.

The voice-producing element will be placed in the shunt valve between the trachea and the esophagus and consists of a single lip of silicone rubber. The choice of the voice-producing element to be placed in the shunt valve leads to several geometric
requirements to the element: it has to be placed in a lumen with an inner diameter of 5 mm and a length of 5 to 11 mm.

Other requirements concern the sound that is produced by the voice-producing element; \( F_0 \), pressure, flow and Sound Pressure Level (SPL) has to fall within physiologic ranges:

- We aim at a voice-producing element having a \( F_0 \) of about 110 Hz for male and about 210 Hz for female.
- The sound pressure level at comfortable effort has to be between 70 dB and 80 dB, measured at 30 cm distance from the mouth.
- Because the lungs will be the energy source of the voice-producing element, they must be able to supply the pressure needed for sound production, ranging from 0.4 to 3.0 kPa, (Schutte 1980).
- The lungs also provide restrictions to the flow through the element. Physiological flow at which phonation occurs ranges from 0.1 to 0.3 L/s, (Schutte 1980).
- The user of the voice-producing element must be able to vary intensity and \( F_0 \) of the sound. Because no external control mechanisms are desired, both frequency and intensity must increase under influence of an increasing flow through the voice-producing element. While varying the applied aerodynamic forces \( F_0 \) must vary between about \(-10\) and \(+30\)% of the values, mentioned above; and the intensity must vary between 70 and 80 dB, (Schutte 1992).

To find the optimal geometry and material for the voice-producing element within the constraints of physiological possibilities, numerical modeling is used to avoid a trial-and-error approach. Because no numerical model on voice-producing
prostheses exists, numerical models of the vocal folds have been studied. Since the human vocal folds and the new voice-producing element have much resemblance (both sound generators produce glottal waves as a result of interaction between the airflow and the oscillating lip), adapting a numerical model of the vocal folds could be a fast way to obtain a realistic model of the voice-producing element.

In numerical models of the vocal folds, both vocal folds and flowing air are simplified considerably, as in the models of several authors, (Ishizaka and Flanagan 1972; Titze 1973; Titze 1974; Story and Titze 1995; Pelorson, Hirschberg, van Hassel, Wijnands, and Auregan 1994). These models are often referred to as lumped-parameter models, because the properties of the vocal folds are lumped together in a small number of parameters. In most commonly used lumped-parameter models, the vocal folds are described by a number of masses, connected to each other by a number of springs and dampers. In these models, the behavior of the flowing air is approximated by the one-dimensional Bernoulli equation, implying a constant pressure distribution in a cross section. The air around the vocal folds does not behave like a one-dimensional flow, because aerodynamic pressure acts at both sides of the folds. As a consequence, a more detailed description of the aerodynamics was essential for an accurate examination of the behavior of the vocal folds. New numerical models have been developed to describe the behavior of the airflow around the lip more accurately using two-dimensional Navier-Stokes equations (Alipour and Titze, 1996; De Vries, Schutte, Veldman, and Verkerke 2000). These models are a good basis for modelling the voice-producing element. Although it is possible to describe the lip by a number of lumped parameters using the numerical method described by de Vries et al. (De Vries, Schutte, and Verkerke 1999), the lip can be represented more accurately with a finite element method (FEM) model.

The aim of this study is to determine the optimum configuration of the voice-producing element using a numerical model. For different configurations of the lip the pressure at which vibration starts, the range of self-sustained oscillation, the flow needed for voice production, and the possibilities to vary F0 will be determined. From the results of the numerical simulations, an optimal configuration for the lip in the voice-producing element will be derived. To validate the numerical model, in-vitro experiments have been performed.

4.2 Materials

The numerical model developed consists of two parts: a model of the aerodynamics based on Navier Stokes equations and a model of the mechanics of the lips based on FEM. First the aerodynamics and mechanics are discussed separately, followed by the description of the interaction between the two models.
4.2.1 Aerodynamics

For the computation of the aerodynamic part of this model, incompressible two-dimensional Navier-Stokes equations are used, as described by de Vries et al., 2000. Spatial discretization is performed in a Cartesian grid. This grid can be refined at places of particular interest. The three degrees of freedom of each cell in the grid (Fig. 3) are the horizontal velocity \( u \), the vertical velocity \( v \) and the pressure \( p \).

![Schematic representation of the grid of aerodynamic cells (white) in interaction with mechanic cells (grey). Direction of velocities \( u \) and \( v \) is indicated.](image)

As boundary conditions, velocities at the solid walls are \( u=0 \) and \( v=0 \). This describes an impermeable wall and a no-slip condition, which means that the fluid sticks to the wall due to its viscosity. At the inlet and outlet the pressure is prescribed, with the pressure difference driving the flow. Additionally, at inlet and outlet the normal derivatives of the velocity components are set to zero, modeling fully developed flow. The latter condition requires the inlet and outlet to lie sufficiently far away from the interesting parts of the flow field.

The boundary conditions at the moving lip will be described below.

4.2.2 Mechanics

FEM is used to describe the mechanical behavior of the voice-producing element. Since the numerical model describing the aerodynamics is two-dimensional, the FEM model of the voice-producing element is two-dimensional as well. The thickness of the lip is small compared to its length, therefore it is allowed to approximate the lip using beam elements. In this way, the geometry in a cross-section is fixed, which would not
be the fact when a three-dimensional element type, like a shell, would have been chosen. The lip is divided in a number of beam elements, connected to each other in nodes. The beam element used has three degrees of freedom in every node: an axial and transverse translation, and a rotation.

The movement of the lip is restricted by the upper and lower wall of the voice-producing element (Fig. 2a). Therefore, during a cycle of the movement of the lip, contact between the lip and the upper and lower wall is described by a collision model that dissipates all kinetic energy.

The material of the lip has been chosen to be silicone rubber. This material is used for the shunt valves also. In the numerical simulations the silicone rubber lip is assumed to have a density of 8.6e3 kg/m^3, a Young’s modulus of 106 Pa, and a proportional damping of 0.01. The initially straight lip with a length of 7 mm is bend 90° to fit into the housing that has an inner height of 3 mm. The shape of the deformed lip and the corresponding forces to keep the lip in the deformed shape has been calculated with the non-linear algorithm of the FEM-program ANSYS 5.5 (SWANSON Analysis, USA). For all configurations of the lip, this additional calculation has been done. The forces that were needed to bend the lip were transformed to a pressure distribution along the lip and added to pressure distribution resulting from the aerodynamic calculations. In this way, the lip is assumed to behave linear in the final numerical model, moving around an equilibrium state that has been calculated using a non-linear method.

4.2.3 Interaction between aerodynamics and mechanics

To study the interaction between aerodynamics and mechanics, the two separate models are integrated by placing the models in a common computational grid (Fig. 3). The cells of this grid can contain a volume of air or a part of the geometry. Interaction between aerodynamics and mechanics is obtained by exchanging information in the common grid in every time step:

The velocity of the air is equaled to the velocity of the lips at the surface of the lip; in this way a continuous velocity field is obtained.

The Navier-Stokes equations compute a pressure field. From this pressure field, forces acting on the lip are calculated by integrating the pressure distribution in the cells adjacent to every beam element and sum the resulting forces in every node between the beam elements. In this way, continuity of the force field is achieved. These forces are used to compute the movement of the lip. The new positions and velocities of the lips form the input for the next time step of the aerodynamics.
4.3 Methods

4.3.1 Simulation
To achieve self-sustained oscillation, a pressure is prescribed at the inlet of the voice-producing element. In the simulations, pressure is increased from zero up to 3 kPa. During the pressure rise, the lip starts to oscillate at a certain pressure, defined as vibration threshold pressure, analogous to the phonation threshold pressure defined by Titze et al., 1992. From this point on to higher pressures, the relation between tracheal pressure, mean flow and F0 are investigated, up to the pressure at which no regular glottal waves are produced.

Several configurations of the lip have been implemented, see Table 1. A basic configuration has been chosen and relative to this configuration, three parameters have been varied: the thickness of the lip, the length of the lip and the height of the housing that the lip is placed in. From this parameter study, an indication of the desired value of the different parameters will be obtained.

4.3.2 Numerical verification
The numerical accuracy of the models of the aerodynamics and of the mechanics have been determined separately. The model of the aerodynamics is a second-order version of the symmetry-preserving method of Verstappen and Veldman (1998) that has been used extensively in simulating turbulent flow. The behavior of the mechanical part of the model is accurate when the description of the lip by the linear beam theory is a valid one. Because the quotient of thickness and length of the beam is very low, the lip can be considered to behave like the beam theory.

The numerical verification of the complete model, including the interaction between the aerodynamic model and the mechanic model, has been done by varying the density of the grid of the aerodynamic model from 80 x 25 to 150 x 50 and the number of elements in the mechanic model from 3 to 9. When finer grids and more elements do not result in other results, the grid size and number of elements is considered to be optimal.

4.3.3 Validation
To validate the numerical model of the lip principle and to obtain insight into the behavior of the lip principle, in-vitro experiments have been performed with a prototype of the voice-producing element in a test setup as described by Van der Plaats et al. (2000). Pressure is slowly increasing from the vibration threshold pressure to 3 kPa. The vibration threshold pressure, F0 as function of tracheal pressure and mean flow are registered.

Both pressure and flow values of the in-vitro experiments, mentioned above are decreased by 40% to correct for leakage along the lip and for leakage caused by torsion of the lip. This rather high amount of leakage is caused by the provisional way of manufacturing and corresponds to other experiences (Rakhorst et al. 2000).
TABLE I. Dimensions of the simulated configurations of the voice-producing element. Dimensions that differ from the basic configuration are indicated in bold.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>thickness lip (mm)</th>
<th>length lip (mm)</th>
<th>height element (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic configuration</td>
<td>0.25</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Long lip configuration</td>
<td>0.25</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>Short lip configuration</td>
<td>0.25</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>Higher opening configuration</td>
<td>0.25</td>
<td>7</td>
<td>4</td>
</tr>
<tr>
<td>Tapered configuration</td>
<td>0.25 (bottom) - 0.125 (free tip)</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Thin lip configuration</td>
<td>0.125</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

The corrected results of the in-vitro experiments are compared to the values determined with the numerical model. Also a comparison with values obtained in laryngeal phonation has been performed to demonstrate the correspondence between the voice produced by the voice-producing element and voice produced by the vocal folds.

4.4 Results

A basic grid of 100 cells in the flow direction and 30 cells in the transverse direction appeared to be optimal, because refining the grid did not change the results significantly, only increased calculation time. For the same reason representation of the lip by 6 elements appeared to be the best.

In Fig. 4, the relation between F0 and the simulated tracheal pressure is depicted. The vibration threshold pressure (Fig. 5) differs for the different configurations, as is presented in Table II. It can be derived that the thickness and length of the lip and the height of the opening in the housing are important for the determination of the vibration threshold pressure: glottal volume velocity waveforms are produced at 0.12 kPa for the thin lip, whereas 0.41 kPa is needed for the basic configuration. The short lip starts oscillating at a pressure of 0.81 kPa, whereas the basic and long lip start
oscillating at a pressure of 0.41 kPa and 0.44 kPa, respectively. The configuration with the highest opening has the highest vibration threshold pressure.

![Diagram showing the relationship between F0 and tracheal pressure for different configurations.](image)

**FIG. 4.** $F_0$ as a function of tracheal pressure, as calculated by the simulation model. Curves start at pressures where oscillation starts and ends at pressures where oscillations become non-periodical.

The range over which self-sustained oscillation occurs depends on the configuration: the thin lip has a phonation range of only 0.6 kPa, whereas the basic configuration has a range of more than 2 kPa.

The relation between $F_0$ and the mean flow is shown in Fig. 6. The mean flow through the simulated voice-producing element is almost linearly related to the tracheal pressure, therefore a minimum in the $F_0$ as function of the mean flow is present also. It must be noted that the higher opening configuration results in a considerable shift towards higher flow values.

The graphs that depict the relation between $F_0$ and tracheal pressure and between $F_0$ and mean flow both have a U-shape. A rise of tracheal pressure above the phonation threshold pressure, respectively flow results in a decrease of $F_0$ for all configurations until a minimum $F_0$ is reached. The minimum $F_0$ and the corresponding tracheal pressure and mean flow are presented in Table II. After that point $F_0$ increases slightly with increasing pressure, respectively flow. In Table II this increase is represented by the slope of the line connecting the two most
extreme $F_0$-values. The tracheal pressure at which the minimum $F_0$ occurs appears to be related with the vibration threshold pressure: a low vibration threshold pressure corresponds with a low pressure at which the minimum $F_0$ occurs.

**TABLE II.** Characteristics of the relation between $F_0$ and tracheal pressure and between $F_0$ and mean flow though the voice-producing element, resulting from the numerical simulation model.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Vibration threshold pressure (kPa)</th>
<th>Minimum $F_0$ (Hz)</th>
<th>Pressure at minimum $F_0$ (kPa)</th>
<th>Mean flow at minimum $F_0$ (L/s)</th>
<th>Increase of $F_0$ with mean flow (1/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic configuration</td>
<td>0.41</td>
<td>270</td>
<td>1.00</td>
<td>0.10</td>
<td>822</td>
</tr>
<tr>
<td>Long lip configuration</td>
<td>0.44</td>
<td>200</td>
<td>1.00</td>
<td>0.13</td>
<td>688</td>
</tr>
<tr>
<td>Short lip configuration</td>
<td>0.81</td>
<td>667</td>
<td>1.30</td>
<td>0.08</td>
<td>3533</td>
</tr>
<tr>
<td>Higher opening</td>
<td>1.00</td>
<td>300</td>
<td>1.50</td>
<td>0.26</td>
<td>87</td>
</tr>
<tr>
<td>configuration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tapered configuration</td>
<td>0.35</td>
<td>230</td>
<td>1.00</td>
<td>0.10</td>
<td>1067</td>
</tr>
<tr>
<td>Thin lip configuration</td>
<td>0.12</td>
<td>250</td>
<td>0.40</td>
<td>0.05</td>
<td>256</td>
</tr>
</tbody>
</table>

The relation between $F_0$ and mean flow, respectively tracheal pressure of the prototype used in the in-vitro experiments are shown in Fig. 7, together with the simulation results of the corresponding long lip configuration. In this way, the numerical model can be validated.

From Fig. 7a it can be seen that the two curves of the relation between $F_0$ and mean flow, resulting from the experiments and from the simulations correspond very well. The two curves of the relation between $F_0$ and tracheal pressure, resulting from the in-vitro experiments and from the simulations (Fig. 7b) correspond less, but still satisfactory. Both in Fig. 7a and 7b, it can be seen that the range in mean flow, respectively tracheal pressure as has been simulated is larger than the range found during the in-vitro experiments.
4.5 Discussion

The values of tracheal pressure needed for the production of glottal waves with the voice-producing element should fall in a range comparable to laryngeal voice production. For laryngeal phonation, Schutte (1980) measured a mean value for the tracheal pressure of 0.44 kPa. Overviewing the results of the simulations, it appeared that four of the six configurations oscillate at that pressure. However, the ideal relation between $F_0$ and pressure is not yet obtained, because increasing pressure causes a decreasing $F_0$. Patients that are able to put more effort in speech by producing higher pressures are able to reach the ideal part of the curve, where an increase in pressure causes an increase of $F_0$. Also, only for higher pressures the other two configurations (short lip and higher opening) start to oscillate.

In the ideal flow range of 0.1 – 0.3 L/s, three of the six configurations oscillate according to the requirements: increasing mean flow causes an increasing $F_0$. The higher opening configuration requires too much mean flow and as a results the speech duration will be shortenend too much. The thin lip and short lip configuration require a non-physiological low mean flow.

**FIG. 5.** Definition of vibration threshold pressure: during an increase of tracheal pressure, the moment of the start of self-sustained oscillation is determined. The corresponding value of the tracheal pressure at that time is known and will be called the vibration threshold pressure for that configuration.
Frequency requirements are only fulfilled for female patients. The lowest frequency obtained is 200 Hz for the long lip configuration. The required female frequency increase of 80 Hz per 0.2 L/s = 400 (1/L) is realized by the basic, long lip, short lip and tapered configuration. The prediction of the acoustical pressure variations resulting from the glottal waves is beyond the scope of this study, therefore the intensity requirements can no be checked.

The assumption of the incompressible flow is valid in our model because the airflow velocity did not exceed 50 m/s which is below Mach 0.2. At such low Mach numbers compressible effects can be neglected. The 2D-approach sometimes leads to unrealistic results. The pressure and flow range at which oscillations occur, was larger in the simulations than in the in-vitro experiments. The vibration threshold pressure appeared to be slightly higher during the in-vitro experiments; in the in-vitro experiments, oscillation above the maximum phonation tracheal pressure and mean flow appears to be non-periodic and not useful for voice production, whereas during simulation much higher pressures and flows still produce periodic oscillations. Most probably, 3D-effects like leakage along the lip and torsion of the lip (absent in the 2D-numerical model) cause the limited range found during the in-vitro experiments.
Because both in Fig. 7a and in Fig. 7b, the curves resulting from the simulation and from the in-vitro experiments show a large similarity, the numerical model appears to be a valid tool to study a relevant part of the pressure and flow range used for voice production.

**FIG. 7.** Relation between $F_0$ and mean flow (a) and relation between $F_0$ and tracheal pressure (b) of the long lip configuration, resulting from the numerical simulation model and from the in-vitro experiments. In-vitro results for mean flow have been scaled by 0.6 to account for leakage.
According to the considerations, mentioned above, the best female voice-producing element is the long lip configuration. More numerical simulation studies could improve this configuration even more by combining several successful configurations like the long and tapered one. A lower opening configuration could lead to a shift towards a lower flow range. With the developed numerical simulation model, these configurations can be tested very fast and reliably.

4.6 Acknowledgement

This research was supported by the Technology Foundation STW, applied science division of NWO and the technology programme of the Ministry of Economic Affairs.


