Chapter 1

Introduction
The ability of speech communication is essential for the interaction between human beings. In the process of speech production, several processes are involved simultaneously. One of these processes results in the production of the source sound for speech. This process is called phonation and is performed by the vocal folds. This source sound, or voice, passes the vocal tract, consisting of the air channels between the vocal folds on the one end and the lips and the nostrils on the other end. The geometrical configuration of this vocal tract can be adapted by movements of the articulators. The configuration of the articulators determines the resonance characteristics of the vocal tract. In this way, the vocal tract acts as an adjustable acoustic converter that converts the source sound to the desired speech sound. So speech is the result of simultaneously acting processes: the generation of the source sound by the vocal folds and the conversion of this source sound to speech by the vocal tract.

In some patients with laryngeal cancer a total laryngectomy is needed to be cured from the disease. In a total laryngectomy the vocal folds are removed. With this removal of the vocal folds, the natural way to produce voice is lost. In most patients, a reasonably good voice can be achieved by pressing air from the esophagus into the pharynx. The passing air starts the mucosa at the entrance of the esophagus to vibrate (Damsté 1958) leading to esophageal voice production (Mahieu 1988). Nowadays, a commonly used method to bring air in the esophagus is the application of a one-way valve between the trachea and the esophagus, so air from the lungs can be used for voice production (Hilgers and Schouwenburg 1991; Mahieu 1988; Nijdam et al. 1982; Singer and Blom 1980). The substitute source sound usually has a low fundamental frequency (60-90 Hz), (Blood 1984; Qi and Weinberg 1995; Qi et al. 1995; Robbins et al. 1984). This is considerably lower than in laryngeal voice production, where values for the fundamental frequency are about 110 Hz for male and 210 Hz for female. The low mean speaking fundamental frequency in esophageal voice is especially unpleasant to women. To overcome this disadvantage, a voice-producing element that produces a source sound with a higher fundamental frequency would be helpful. The developmental process of such a voice-producing element is the subject of this thesis.

To define the requirements for an artificial voice source laryngeal voice production needs to be studied. The requirements ultimately formulated are presented in chapter 6. The voice-producing element is composed of a single lip of silicone rubber in a square housing. The underlying principle is comparable to the vibrating lips of a musician playing a brass instrument (Adachi and Sato 1996; Sram 1989; Sram et al. 1983). In the voice-producing element, only one vibrating lip will be placed because it is easier to produce than an element consisting of two lips. In the neutral position, the lip is pressed against the wall opposite to the wall where the lip is inserted. When pressure is applied at the inlet of the voice-producing element, the lip starts to vibrate. When the initially closed lip opens, air flows along the lip. The right mechanical properties and geometry in combination with an appropriate pressure let the lip vibrate as result of aerodynamic and mechanic forces on the lip.
The production of voiced sounds in laryngeal phonation is a result of the cyclic opening and closing of the glottis. The glottis is the area between the two vocal folds. The process of opening and closing is a result of the interaction between the airflow, coming from the lungs through the glottis, and the tissue mechanics, occurring in the vocal folds. This interaction has been studied by numerous investigators: some of them focusing on the aerodynamics of phonation (e.g. Van den Berg et al. 1957; Scherer and Titze 1983; Schutte 1980), others focusing on the tissue mechanics (Baer 1981; Berry et al. 1994; Hirano 1974; Titze 1973; Titze 1974). The interaction between the airflow and the tissue mechanics is described by several investigators (Alipour and Titze 1996; Herzel et al. 1995; Ishizaka and Flanagan 1972; Pelorson et al. 1994; Story and Titze 1995). The last mentioned investigators all make use of numerical models of the interaction between the vocal folds and the airflow passing them.

By a stepwise improvement of these numerical models and by exchanging vocal folds to a voice-producing element, a numerical model has been developed that can be used for simulating and thus improving the behavior of the voice-producing element. Numerical models that are able to simulate vocal fold voice production consist of a simplified description of the vocal fold that interacts with a simplified description of the air that passes the vocal folds. The properties of the vocal folds are lumped together in masses, springs, and dampers. Several investigators presented a lumped parameter model using two masses, the so-called two-mass model (Herzel et al. 1995; Ishizaka and Flanagan 1972; Steinecke and Herzel 1995). These two masses move under influence of aerodynamic forces. These forces are calculated using a simplified description of the airflow in the glottis, which is based on the Bernoulli equation (examined in the glottis by Van den Berg (Van den Berg et al. 1957). This equation relates pressure and flow in a fluid. The interaction between the two-mass model and the aerodynamic forces were studied by several authors (Alipour and Titze 1996; Herzel et al. 1995; Ishizaka and Flanagan 1972; Pelorson et al. 1994; Story and Titze 1995).

For the two-mass model, Ishizaka and Flanagan adopted values for the lumped parameters (Ishizaka and Flanagan 1972). Two-mass models of the vocal folds presented after 1972 often make use of the same parameter values (Herzel et al. 1995). From studying the two-mass models of the vocal folds, it was obvious that these parameters were not determined by mechanical considerations but by pragmatic searching for the optimum glottal-wave generator. Therefore, we felt the need to have a method that makes it possible to determine the values for the parameters of the two-mass model which are based on mechanical considerations. In this way, a more realistic choice of the parameter values is obtained. In chapter 2, a first step of improvement of existing models has been made by determining new values for parameters governing the two-mass models of the vocal folds using a Finite Element Method (FEM) model of the vocal fold. By requiring an equal response of the two-mass model of the vocal folds and the Finite Element Method model of the vocal folds, the parameter values of the lumped parameter models can be tuned to more realistic values.
In the existing lumped parameter models of the vocal folds, the description of the aerodynamics is based on the Bernoulli equation. For the modeling of the voice-producing element, the use of the Bernoulli equation is not correct because the aerodynamic quantities (pressure and velocity) are not constant over a cross-section, as is assumed in the glottis in the lumped parameter models. Therefore, the second step of improvement is the introduction of a more accurate description of the aerodynamics. In chapter 3, this second step of improvement of the existing two-mass models is made by using a Navier-Stokes description of the glottal flow instead of a Bernoulli-based description. In this way, the results obtained with the two-mass models using the Bernoulli equation and the results obtained using the Navier-Stokes equation are compared.

At this stage in the process of simulating voice production, the voice-producing element that will be developed is used instead of the lumped parameter model of the vocal folds in the numerical model of chapter 3. Therefore, a third step of improvement is made (chapter 4), in which the voice-producing element is modeled by the FEM and brought into interaction with a Navier-Stokes description of the flow. With a FEM model, variations in the geometry and material properties of the voice-producing element can be evaluated much easier than with the earlier models based on lumped parameters.

With a FEM model and aerodynamics based on the Navier-Stokes equations, an accurate description of the voice-producing element including the fluid-structure interaction is obtained. With this model, the behavior of the voice-producing element has been predicted and can be improved.

To test the improved voice-producing element under realistic and reproducible conditions, an in-vitro setup has been developed which is described in chapter 5. In our case, realistic conditions are reached by giving the set-up the same acoustical properties as the structures involved in laryngeal phonation. These structures are the subglottal and supraglottal tract. To create a physical model of the human trachea and lungs with a comparable input impedance as has been measured by (Ishizaka et al. 1976), a numerical tool (Bergh and Tijdeman 1965) has been used. The physical, perspex model has been validated by means of an impedance tube measurement (Van der Eerden et al. 1998; Pierce 1989).

Using the same numerical tool, also three supraglottal tract vowel models based on a twin tube model (Mol 1970) have been developed. These models have been validated by means of a frequency sweep procedure. The acoustic influence of the subglottal vocal tract and of the supraglottal vocal tract on the functioning of the voice-producing element is examined. With the in-vitro setup, the voice-producing element has been tested under acoustical realistic conditions.

In chapter 6 an overview of the design process is presented. Requirements are discussed and the results of numerical modeling with the in-vitro experiments are compared. Based on the findings of the numerical model, nineteen prototypes have been manufactured and tested with the in-vitro set-up to select out the best male and female voice-producing element. Apart from single use the voice-producing element
could be integrated in an artificial larynx that is in development in the Eureka project ‘Artificial Larynx’.

This study has been financed by the Technology Foundation (STW).


