Annals of Biomedical Engineering (© 2006)
DOI: 10.1007/s10439-006-9196-3

Development of a Double-Membrane Sound Generator for Application in a Voice-Producing Element for Laryngectomized Patients

J. W. TACK,[1] G. J. VERKERKE,[1] E. B. VAN DER HOUWEN,[1] H. F. MAHIEU,[2] and H. K. SCHUTTE[1]

[1]Department of BioMedical Engineering, University Medical Center Groningen, University of Groningen, A. Deusinglaan 1, 9713 AV, Groningen, The Netherlands and [2]Department of Otorhinolaryngology, VU Medical Centre, De Boelelaan 1117, 1007 MB, Amsterdam, The Netherlands

(*Received 24 March 2006; accepted 30 August 2006*)

Abstract—For voice rehabilitation after total laryngectomy a shunt valve is usually placed in the tracheo-esophageal (TE) wall, thereby enabling the production of a TE voice. Some patients, however, are unable to produce a voice of sufficient quality. Furthermore, the TE voice is low pitched, which presents a problem especially for female laryngectomized patients. The voice quality after laryngectomy might be improved by introducing a voice-producing element (VPE) into the TE shunt valve. In this study a sound generator was developed that is suitable for application in such a VPE. This sound generator consists of two elastic membranes placed parallel inside a circular housing. A substitute voice source is created when the membranes start to vibrate via a constant flow of air passing between them. To determine the optimal membrane configuration for proper functioning under physiological conditions, up-scaled physical VPE models with different membrane geometries were evaluated using *in vitro* experimental tests. For certain membrane geometries the tests showed that a basic sound, containing multiple harmonics, could be successfully produced under physiological air pressure and airflow conditions. The fundamental frequency (60–95 Hz) and sound pressure level (57–78 dB, at 15 cm microphone distance) were regulated via changes in the driving pressure, thereby enabling the possibility of intonation in laryngectomized patients' speech. The obtained frequency range is considered appropriate for producing a substitute voice source for female patients. The geometry considerations in this study can be used for the development of a true scale VPE that can be evaluated clinically, to eventually replace the voice after laryngectomy.

Keywords—Laryngectomy, Voice rehabilitation, Shunt valve, Voice prosthesis, *In vitro* test, Self-sustained oscillations, Aero-acoustics, Voice sound.

Address correspondence to G. J. Verkerke, Department of BioMedical Engineering, University Medical Center Groningen, University of Groningen, A. Deusinglaan 1, 9713 AV, Groningen, The Netherlands. Electronic mail: g.j.verkerke@med.umcg.nl

INTRODUCTION

A total laryngectomy is indicated sometimes for patients with advanced laryngeal cancer, consisting of the surgical removal of the larynx, including the vocal folds and epiglottis (Fig. 1a). The trachea is separated from the larynx and sutured to an opening in the skin of the neck, thereby forming a tracheostoma through which the patient is able to breathe (Fig. 1b).

The most disturbing consequence of a laryngectomy is the loss of a natural way to produce voice. For voice rehabilitation a one-way shunt valve is usually placed in the tracheo-esophageal (TE) wall (Fig. 1b).[19,20,25] When the tracheostoma is occluded by either a tracheostoma valve or the patients' thumb, exhaled air flows through this valve into the esophagus, where remaining muscular structures and mucosa (the pharyngo-esophageal segment) can be set into vibration. The low-pitched sound that is generated in this way can be used as a substitute voice source for so-called TE speech.[18] The mean fundamental frequency of this TE voice is usually 60–90 Hz.[1,21,23] Such a low fundamental frequency is a problem especially for women, since in a normal female laryngeal voice this frequency has a mean value of about 210 Hz. Even for males this frequency is rather low; the mean fundamental frequency of the male laryngeal voice is about 120 Hz.[24] A frequency variation of about 7 semitones (≈50%) is used for intonation during normal phonation.[3,13] Besides the low pitch of the TE voice, for some laryngectomized patients another problem in voice rehabilitation is the low tonicity of the pharyngo-esophageal segment, which results in a breathy TE voice of poor quality.[2,17]

To improve voice quality in female laryngectomized patients and in patients with a hypotonic or atonic pharyngo-esophageal segment, recently several sound-generating prosthetic prototypes have been developed. Such devices, which we will call a voice-producing

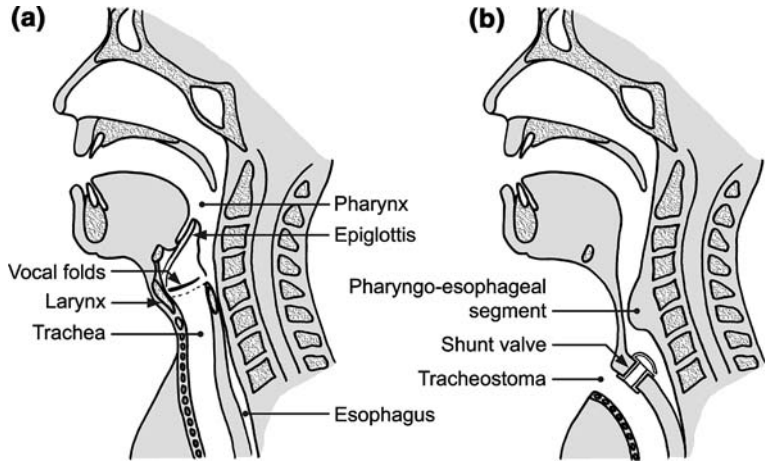

FIGURE 1. Sagittal section of the head and neck region (a) before laryngectomy, and (b) after laryngectomy with a shunt valve placed in the wall between the trachea and esophagus.

element (VPE), were placed inside the TE shunt valve. The VPE converts a constant flow of air from the lungs into a complex sound via self-sustained oscillations, thereby creating a substitute voice source.

Hagen *et al.*[12] and Herrmann *et al.*[14] presented results of clinical tests with a VPE prototype containing an inwards-striking metal reed, such as found in a mouth-organ. Using this device, patients were able to produce clear voiced sounds with a fundamental frequency higher than that normally found in TE speech. However, the disadvantage of these reed-based elements is that they produced a sound with a fixed frequency, leading to an unnaturally monotonous voice. Furthermore, the element appeared to be sensitive to blockage by tracheal secretions. De Vries *et al.*[4,5] developed a different kind of VPE, consisting of an outwards-striking silicone rubber lip placed in a square housing. In clinical tests, however, Van der Torn *et al.*[29] observed that the functioning of the silicone rubber lip was also sensitive to the mucus that entered the prototype. The lip sometimes stuck to the interior of its housing.

We have developed a new sound generating principle for the VPE, based on a double-membrane concept. This sound generator model consists of two weighted elastic membranes, placed parallel to each other inside a circular housing (Fig. 2). In comparison with the reed- and lip-based sound generators this membrane concept is expected to be less sensitive to blockage by mucus, since the exhaled air has to pass the lumen between the membranes, thus removing the mucus. Moreover, the membranes can be pushed away from each other to create a larger through-flow opening for the passing mucus, while afterwards the membranes will always return to their initial position because of their attachment.

The aim of this study is to realize a sound generator which is based on the double-membrane concept and is suitable for application in a VPE. Therefore, the sound generator has to operate under physiological air pressures and airflow rates, while the sound frequencies and intensities are suitable for producing voice. In this study we first considered the mechanism of sound production to indicate the membrane properties and geometries that are important for the flow problem considered. Subsequently, to determine the optimal membrane geometry several physical models with different membrane geometries were manufactured and tested *in vitro*. The physical models were manufactured on a 2:1 scale to allow for more accurate adjustments to the geometrical parameters evaluated. So the membrane dimensions were twice as large as compared to the anticipated final design. Consequently, the test results from the up-scaled models needed to be evaluated using scaling laws derived from a dimensional analysis of the double-membrane concept.[27]

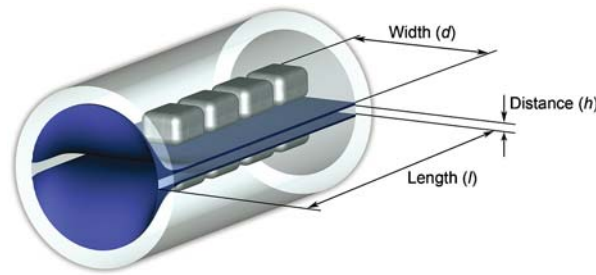

FIGURE 2. Drawing of the VPE based on the double-membrane concept, with the tested membrane geometries indicated.

DOUBLE-MEMBRANE CONCEPT

Mechanism of Sound Production

Figure 2 shows a drawing of the double-membrane VPE concept as viewed from the upstream side. Fixed inside a circular housing are two membranes, which start to vibrate and produce a sound under the influence of airflow. Each membrane is loaded with weights in order to vary the oscillation frequency.

The pressure level (p) that is required to drive the oscillations, and the resulting airflow rate (q), are mainly determined by the shape and the size of the membrane flow channel. This flow channel is specified by the membrane length l, width d, and membrane distance h, as indicated in Fig. 2. Since characteristic airflow velocities are very low (the Mach number is in the order of 10^{-1}), the flow is considered incompressible in first approximation. For characteristic membrane geometries the Strouhal number ($S_r = lf_0dh/q$) is, dependent on the membranes fundamental frequency f_0, of the order 10^{-1}. Hence, the flow is also considered quasi-stationary.[16] For a stationary situation, the flow rate is usually given by the Bernoulli equation:

$$q = dh\sqrt{\frac{2p}{\rho}}, \tag{1}$$

with ρ the air density. The driving pressure p is determined by the difference between the pressure in the trachea and the pressure just downstream to the element. It assumes – in the absence of sharp edges at the flow entrance – that the flow separates from the membrane at the end of the flow channel, forming a free jet, thus assuming the pressure in the flow channel to be equal to the pressure downstream. However, the use of this equation implies that friction can be neglected. While this is true for the average membrane aperture height $\bar{h}$ (with typical Reynolds numbers in the order of 10^3), friction cannot be neglected when the membranes collide during part of the oscillation cycle. In this case, when friction has to be accounted for, the pressure drops linearly with the length l of the flow channel, in accordance with expectations from a fully developed Poisseuille flow.

The membranes oscillate under the influence of several varying forces. Forces acting on the membrane surface are for example hydrostatic forces due to the pressure in the flow channel, and counteractive aerodynamic forces (''Bernoulli forces'') due to the airflow in the channel. The aerodynamic force has a dependence on the height h of the membrane aperture, but is not uniformly distributed over the membrane during oscillation because of a varying membrane distance along the membrane width d and length l. The external forces cause the membrane to deflect from its resting position, giving rise to elastic 'restoring' forces that are dependent on the amount of membrane strain during oscillation. Because of the complex interaction between the various forces and their – sometimes nonlinear – geometrical dependence, we will limit our analysis with a description of the membrane oscillation.

The oscillation of the membrane with a simplified model is illustrated in Fig. 3. In a zero-flow condition the membranes are in stable 'resting' position (Fig. 3a). Consider now the response of the membranes to an increase in pressure p from zero to a certain threshold value. Initially, the membranes are pushed away from each other because of relatively large hydrostatic forces, while the internal stiffness of the membranes result in an elastic force directed towards the resting position. The membranes' inertia causes them to move beyond the point of dynamic equilibrium, until the elastic and aerodynamic forces stop the movement and force the membranes back to the starting position and beyond, ultimately ending in the collision of the membranes and thereby closing off the airway channel. At this moment the motion is reversed, and through the kinetic energy from the collision, the elastic and hydrostatic forces, and the membrane inertia, the membranes move away from each other past their resting position again, thus completing one cycle of the self-sustained oscillation, consisting generally of an 'opening' and 'closing' phase (Fig. 3b). The magnitude of the membrane response that follows an excitation not only depends on physical membrane properties like the inertia and the elasticity, but is also dependent on the damping due to friction and heat transfer. However, since we are primarily interested in the frequency

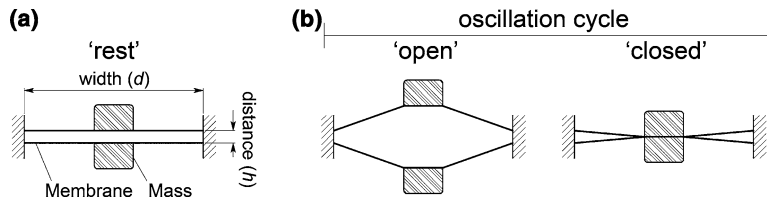

FIGURE 3. Illustration of the membranes oscillation, viewed from the direction of the airflow. Position of the membranes (a) at rest and (b) during oscillation when maximally opened, and closed.

of the self-sustained oscillations, and not the actual membrane oscillation amplitude, we did not consider the effect of damping in detail.

Because of the moving membrane surface the airflow is frequency modulated and sound pressure waves are produced. Besides the fundamental frequency, the sound also contains multiples of this frequency, the harmonics, which improve the naturalness of the artificial voice source. These harmonics result from the sudden change in airflow velocity at the moment of airway closure. The underlying principle is comparable to the oscillating lips of a musician playing a brass instrument,[9,10,26] and also to the avian vocal system, the syrinx, in which thin membranous sections at the junction of the two bronchi interact with the airflow from the lungs, producing a frequency-modulated sound.[7,11] Characteristic for these dynamic systems is that the oscillations have a wave-like propagation along the membrane.

The system of membrane oscillations with relatively large amplitudes behaves as a nonlinear dynamic system, similar to the nonlinear dynamics of the syrinx as described by Fee *et al.*[6] Accordingly, the membrane oscillation frequency is dependent on the aerodynamic forces exerted. An increasing driving (lung) pressure could, therefore, lead to an increase in the fundamental frequency of the sound produced, thus enabling intonation for speech.

Membrane Oscillation Frequency

We then considered important factors and geometrical parameters that influence the oscillation of the membranes. First, we considered the general conditions for initiating and maintaining oscillation, as discussed by Fletcher.[8] Fletcher defined these conditions for different types of simple pressure controlled valves. These are categorized via two signs. The first sign indicates if the valve opens (+) or closes (−) when a small overpressure is applied from the upstream side, the second sign indicates this valve behavior with a pressure applied from the downstream side. (Note that aerodynamic forces due to the airflow are omitted in this consideration.) As an example, the inwards-striking metal reed mentioned earlier has valve configuration (−,+), and the outwards-striking rubber lip configuration (+,−). According to this classification the double-membrane sound generator, like the valve in the avian syrinx, can be characterized by valve configuration (+,+), which implies that both a capacitive upstream impedance or an inertive downstream impedance can favor self-sustained oscillations.

It is expected that the membranes oscillate close to the membranes' first resonance frequency, which is an assumption that is usually valid for musical instruments. This natural oscillation frequency is mainly determined by the vibrating amount of mass and the internal membrane stiffness. We modeled the membrane as a string under tension (Fig. 3a) to derive a general indication of the parameters influencing the membrane oscillation frequency. In this model, the total mass of the added weights is concentrated in the middle, while in comparison the membrane is considered negligible. Although the actual membrane oscillations involve nonlinear dynamics, the oscillation frequency can be roughly estimated by considering the equation of motion of the free-vibrating string *in vacuo* with only small motions from the resting position. The resulting approximation of the natural oscillation frequency ω_n (rad/s) is:

$$\omega_n = \sqrt{\frac{4T}{md}}, \tag{2}$$

where m is the effective mass, consisting of the total amount of mass attached to the membrane, d is the membrane width (Fig. 2), and T is the elastic tension force used to stretch the membrane, defined as:

$$T = EA_l\varepsilon_d = El t\varepsilon_d, \tag{3}$$

with E the Young's Modulus of the membrane, A_l the cross-sectional area of the membrane in lengthwise direction with length l and thickness t, and ϵ_d the membrane's pre-strain that is applied along the membrane width. During oscillation, especially at increased levels of membrane excitation, the actual membrane tension T is expected to be higher than Eq. 3 suggests because of relatively large membrane oscillation amplitudes and colliding membranes, thus leading to an oscillation frequency higher than theoretically expected. However, by selecting specific membrane materials and geometries these equations can be used to estimate the oscillation frequency, and thus they can also be used to indicate the influence of scale on the frequency.

Physical Models

The physical models were manufactured to a geometrical scale 2:1 on the basis of VPE dimensions that allow placement inside a shunt valve in the TE wall. Typical shunt valves, like the commercially available Groningen Button[19] or Provox valve,[15] have an internal shaft diameter of 5 mm. The length of the shaft is 5–11 mm, depending on the thickness of the tissue wall between the trachea and esophagus.

In the up-scaled physical models a synthetic latex (Penrose tubes, Rubber B.V., Hilversum, The Netherlands) was used as elastic membrane. This material

has good mechanical properties and a relatively low modulus of elasticity (E = ca 2 MPa, t = 0.13 mm). The same amount of membrane pre-strain was applied in each model tested (ϵ_d = 1.6%). On top of each membrane, separate and equally shaped lead weights (0.19 g each) were attached, evenly distributed lengthwise along the membrane (Fig. 2). Through the segmentation of the mass into separate pieces, most of the membrane's flexibility was preserved, allowing for the occurrence of transversal membrane vibrations. The maximum total amount of mass that could be applied was limited by the space between the membrane and the housing available for the membranes oscillation amplitude. Furthermore, the number of applicable weights was dependent on the length of the membrane tested (m = 0.57 g for membrane length of 15 mm, and m = 0.76 g for length of 20 mm).

Several different membrane geometries were tested, see Fig. 2. The distance (h) between the membranes and the membrane length (l) are considered important parameters influencing, for example, the amount of airflow resistance experienced by the patient (Eq. 1), and the magnitude of the aerodynamic (Bernoulli) forces. Another geometric variable of interest is the width (d) of the membranes, which is related to the internal membrane stiffness (Eq. 2). A double-membrane configuration is defined by values for these geometric variables and is referred to as configuration $h_i l_i d_i$. Table 1 shows the values chosen for the variables h_i, l_i, and d_i, where the index i has a value of 0, 1, or 2 referring to the value in the corresponding column. For example, configuration $h_1 l_1 d_1$ had a distance of 0.5 mm between the membranes, a maximal membrane length of 20.0 mm and a membrane width of 8.0 mm. With the variables presented in Table 1, 18 unique combinations of $h_i l_i d_i$ could be formed ($3 \times 2 \times 3$), but in our tests we considered sixteen different double-membrane geometries. Two configurations, combining h_2 (1.0 mm) with d_0 (7.1 mm), were not tested because preceding tests indicated poor oscillation characteristics for these membrane configurations.

Given the relatively small size of the element, one of the main challenges in the design of the double-membrane sound generator as a VPE appears to be the generation of a sound with a sufficiently low fundamental frequency. For a male voice an f_0 of about 120 Hz is required, which transforms due to scaling to an f_0 of 30 Hz for the models. Considering Eq. 2, the lowest oscillation frequency is expected for membrane configurations that have width d_2 (9.0 mm). Substituting the membrane geometries and material properties in Eqs. 2 and 3, theoretically an f_0 of 35 Hz is expected for configurations with d_2 (9.0 mm). For configurations with d_1 (8.0 mm) this frequency is approximately one semitone (6%), and for d_0 (7.1 mm) two semitones (12%), higher.

Table 1. Different settings for the membrane geometries $h_i l_i d_i$ tested.

Variable	i = 0	i = 1	i = 2
h_i (mm)	0.0	0.5	1.0
l_i (mm)	15.0	20.0	not used
d_i (mm)	7.1	8.0	9.0

Figure 4 shows an experimentally tested VPE model with membrane configuration $h_0 l_1 d_1$, as seen from three different angles, from the side, upstream, and downstream, respectively. The housing of the physical models was made of PMMA. Via this housing two membranes, with a specific length and width, could be fixed at the proper in-between distance. A part of the membrane protruded from the housing at the upstream side, and was folded around the outside of the housing to create the inflow opening. Furthermore, the physical model provided the connection of the membrane housing to the airflow channel inside the experimental test setup.

METHODS

Quantitative Evaluation

The different membrane configurations were evaluated on their aero-acoustic parameter values. These parameter values are related to the physiological situation, namely the effort required from the patient to produce a voice in terms of driving lung pressure (p) and expiratory airflow rate (q). Furthermore, we considered voice characteristics, namely the fundamental frequency (f_0), the frequency variation during normal speech expressed in semitones range, and the sound pressure level (SPL) as measured at a fixed distance from the sound source. The sound quality of the VPE models was further quantified via determination of sound production efficiency and the noise-to-harmonics ratio (NHR).

Table 2 presents the physiological values of p, q, and f_0, as measured for a laryngeal voice. The theoretical parameter ranges that are valid for our VPE model, at the geometrical scale of 2:1, were calculated by taking into account the effect of scaling.[27] The parameters measured for a specific membrane configuration lie within these ranges in the most ideal situation. Therefore, the data presented in Table 2 were used as quantitative evaluation criteria to designate the most promising membrane configuration for clinical use.

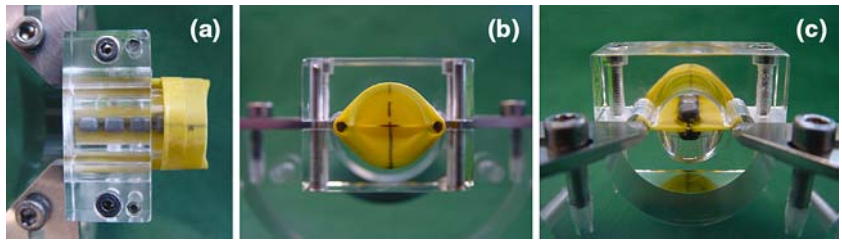

FIGURE 4. Pictures of membrane configuration $h_0l_1d_1$, with a view from (a) the side, (b) upstream, and (c) downstream.

Experimental Test Setup

A scheme of the complete experimental setup for measuring the physical models *in vitro* is shown in Fig. 5. Via an air cylinder a flow of air was supplied that was able to build up pressure inside a reservoir (volume = 0.23 m^3). This capacitive reservoir, coated inside with a sound-absorbing material, has one outflow opening, to which a VPE model was connected.

The airflow rate was manually adjusted via a Brooks flowmeter (Brooks Instrument GT1357, tube R-6–25-B, Veenendaal, The Netherlands), and also measured with a Lilly flowhead (Mercury Electronics, Glasgow, Scotland) that was connected to a differential pressure transducer (Honeywell 164PC01D76, Freeport, Ill., USA) and custom-built amplifier. The calibration of the flowhead was accomplished with the Brooks flowmeter mentioned, while the typical accuracy of the flow rate measurements was ± 5 ml/s. The air pressure inside the reservoir was measured in relation to the atmospheric pressure with a pressure transducer (Hewlett Packard 267BC, Waltham, Mass., USA), connected to an amplifier (Hewlett Packard carrier 8805B, Waltham, Mass., USA) and calibrated against a water manometer. The typical accuracy of the air pressure measurements was ± 0.02 kPa.

The sound was measured with a condenser microphone (Bruël & Kjær 4134, Copenhagen, Denmark), connected to an amplifier (Bruël & Kjær 2609, Copenhagen, Denmark). The calibration of the microphone's SPL was accomplished with an Acoustical Calibrator (Bruël & Kjær 4231, Copenhagen, Denmark). The microphone was placed outside the stream of air and close to the sound source, because the sound measurements were performed in an acoustically non-defined room of relatively small size. We used a 15 cm distance from the microphone to the double-membrane model. Theoretically, the SPL measurements in our setup were 6 dB higher as compared to measurements at a microphone distance of 30 cm.

The signals were digitally processed and recorded using a PC with custom-built data acquisition software (National Instruments LabVIEW™ 6.1, with 'Sound & Vibration' toolset 2.0). All signals were sampled real-time using a sampling rate of 8 kHz and a 4k sample block size. From the sound signal an exponentially averaged sound level was determined using a pressure reference of 20 μPa and a time constant of 125 ms. The sampling rate of 8 kHz allowed for the composition of a power spectrum of the sound signal for frequencies up to 4 kHz, which covers the domain of interest for voice analysis. We applied an A-weighting filter and RMS averaging while linearly weighting 25 averages, thereby reducing the signal fluctuations but not the noise floor. The fundamental frequency of the sound produced was derived from the power spectrum by means of a peak search algorithm in the 'Sound & Vibration' toolset of LabVIEW™. Furthermore, an acoustic efficiency was calculated as the ratio of acoustic power (dependent on the SPL) to the power provided to produce the sound ($p{\cdot}q$), according to the efficiency calculations performed by Van den Berg[28] and Schutte.[24] In addition, a sound sample was recorded (48 kHz sampling rate) and the NHR was determined using the Computerized Speech Lab in

Table 2. The physiological driving pressure (p), airflow rate (q), and fundamental frequency (f_0) ranges, and their multiplication factors for attaining the ranges applicable for a VPE model on a geometric scale (C).

	p (kPa)	q (ml/s)	f_0 male voice (Hz)	f_0 female voice (Hz)
Laryngeal voice during normal speech	0.2–1.5[a]	45–350[a]	Mean[a]: 120 Range[b]: 100–150	Mean[a]: 210 Range[b]: 180–270
Scaling factor (C:1)	$(C)^{-2}$	(C)	$(C)^{-2}$	$(C)^{-2}$
VPE model (C = 2)	0.05–0.38	90–700	Mean: 30 Range: 25–37.5	Mean: 52.5 Range: 45–67.5

[a]Average for 45 healthy adults (male and female), measured at normal intensity (70 dB(A) at 30 cm distance from the mouth); ref. 24.
[b]Range of about −3, +4 semitones with respect to the mean speaking fundamental frequency for a normal intonation pattern; refs. 3, 13, 24.

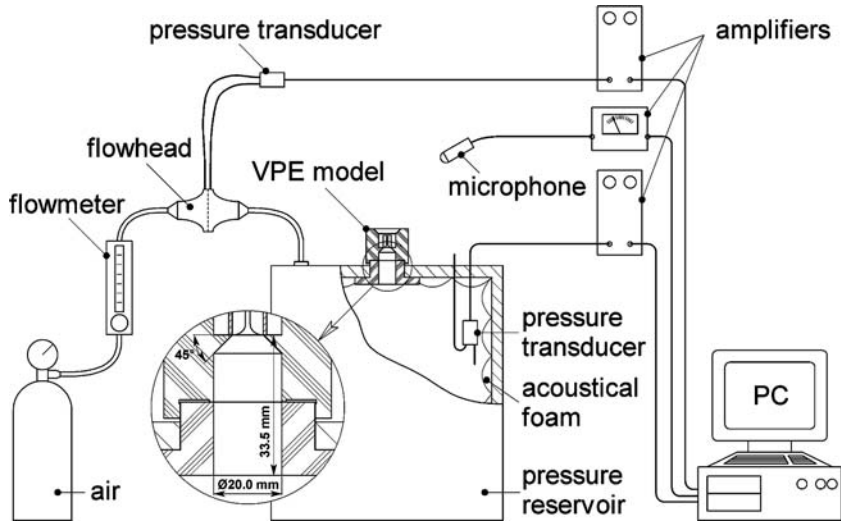

FIGURE 5. Schematic representation of the *in vitro* experimental setup. Components are not to scale. VPE = voice-producing element.

combination with the Multi-dimensional Voice Program (Model 5105, Kay Elemetrics Corp., Pine Brook, USA). The NHR is a general evaluation of the noise present in the analyzed signal. It is defined as an average ratio of the inharmonic spectral energy in the frequency range 1500–4500 Hz to the harmonic spectral energy in the frequency range 70–4500 Hz.

Measuring Method

After calibrating the equipment, the airflow rate was increased until the membranes of a VPE model started to vibrate and a sound was produced. All aero-acoustic parameters mentioned were recorded digitally at stable values of the mean air pressure and airflow rate. Subsequently, the airflow rate was increased by approximately 20 ml/s, and all parameters were recorded again at a stable air pressure and flow rate. This process – increasing the airflow rate and measuring the accompanying parameter values – was repeated until the masses on top of the membranes touched the inner wall of the housing due to large membrane oscillation amplitudes, or when the membrane oscillation became unstable. In addition to these measurements, video-stroboscopic recordings were made during the tests in order to observe the membrane oscillations from the downstream side.

At least two identical double-membrane configurations were manufactured and tested to determine the reproducibility of the results. When, at similar driving pressures, a difference larger than 20% was observed between the airflow rates for both configurations, a third identical configuration was also tested.

RESULTS

Table 3 shows the data measured for all membrane configurations $h_i l_i d_i$ represented in the left column. For the following parameters the lowest and highest values measured are shown respectively: the air pressure (p), the airflow rate (q), the corresponding fundamental frequencies (f_0), and SPL. Between parentheses the SPL measured at approximately 1 kPa ($SPL_{@1\ kPa}$) is presented whenever the pressure range measured for a specific configuration exceeded the value of 1 kPa, thereby enabling easier comparison with the other configurations. Furthermore, Table 3 shows the change in f_0 expressed in the number of semitones, the acoustic efficiency range, and the mean NHR. The shaded cells in the table indicate for each parameter (column) which three parameter values are considered best in view of the requirements.

The relations between the parameters are illustrated in Fig. 6 via the results of the three best-rated membrane configurations ($h_1 l_0 d_2$, $h_2 l_0 d_1$, and $h_2 l_0 d_2$). The relationship between p and q was approximately linear for all configurations (Fig. 6a). The f_0 and SPL increase with increased p was nearly linear for most of the configurations (resp. Figs. 6b and c). In some cases the rate of increase decreased when the pressures reached the upper part of the pressure range. During the measurements, while increasing the driving pressure, a sudden yet small decrease in f_0 or SPL was sometimes observed. A sudden increase was observed only occasionally. These changes were due to a change in the vibration mode of the membranes, as could be observed visually. Furthermore, the NHR data of some configurations showed an erratic course in relation to the pressure increase (Fig. 6d), while overall the

Table 3. Parameter values for all membrane configurations measured.

Membrane configuration			*p* (kPa)	*q* (ml/s)	f_0 (Hz)	SPL ($SPL_{@1\ kPa}$) (dB)	Semitones	Efficiency $\cdot 10^{-5}$	NHR[a]
h_0	l_0	d_0	0.74–1.09	21–41	90–112	59–71 (68)	4	0.7–4.1	0.23
		d_1	0.46–1.67	28–114	91–124	64–84 (75)	5	1.8–17.2	···
		d_2	0.43–0.90	24–57	75–93	57–74	4	0.8–6.3	0.28
	l_1	d_0	0.85–1.21	23–33	93–103	62–68 (65)	2	1.1–2.4	0.22
		d_1	0.60–1.36	41–130	80–122	63–76 (73)	7	0.8–5.4	···
		d_2	0.42–0.96	30–60	74–93	61–75	4	1.5–7.2	0.22
h_1	l_0	d_0	0.24–0.80	30–88	56–93	57–73	9	1.3–4.4	0.12
		d_1	0.30–1.14	45–141	72–101	61–76 (75)	6	1.1–3.9	···
		d_2	0.15–1.06	40–157	60–95	57–78	8	0.8–5.3	0.14
	l_1	d_0	0.27–1.02	33–91	71–114	59–71	8	1.2–2.1	0.17
		d_1	0.45–1.41	40–104	76–100	63–76 (72)	5	1.0–3.8	···
		d_2	0.21–1.09	23–118	60–94	57–74 (72)	8	1.5–2.5	0.17
h_2	l_0	d_1	0.16–1.19	53–184	60–97	60–79 (77)	8	1.2–5.3	···
		d_2	0.23–1.28	90–207	70–98	60–81 (79)	6	0.7–6.4	0.16
	L_1	d_1	0.25–0.85	62–138	62–80	62–71	4	0.7–1.5	···
		d_2	0.25–1.04	77–156	62–88	59–73	6	0.6–1.7	0.13

Shaded cells indicate the three best configurations for each parameter; light gray = good; medium gray = better; dark gray = best.
[a]NHR is the mean noise-to-harmonics ratio. Ratio could not be determined for membranes with width d_1 (see text).

NHR increased with increased pressure. The dotted line at 0.19 in Fig. 6d is the normative threshold value as set by the developers of the Multi-dimensional Voice Program (Kay Elemetrics Corp., Pine Brook, USA) after extensive field testing with normal and disordered human voices. A voiced sound is classified as 'normal'

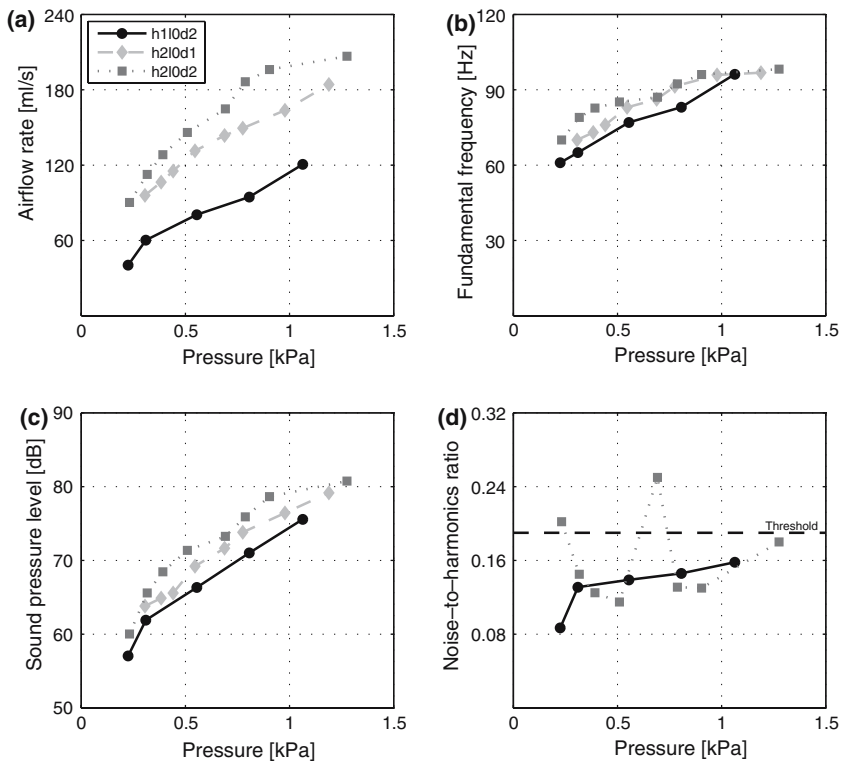

FIGURE 6. The relationship between the driving pressure and (a) the airflow rate, (b) the fundamental frequency, (c) the SPL, and (d) the NHR ratio, as measured for membrane configurations $h_1l_0d_2$, $h_2l_0d_1$, and $h_2l_0d_2$.

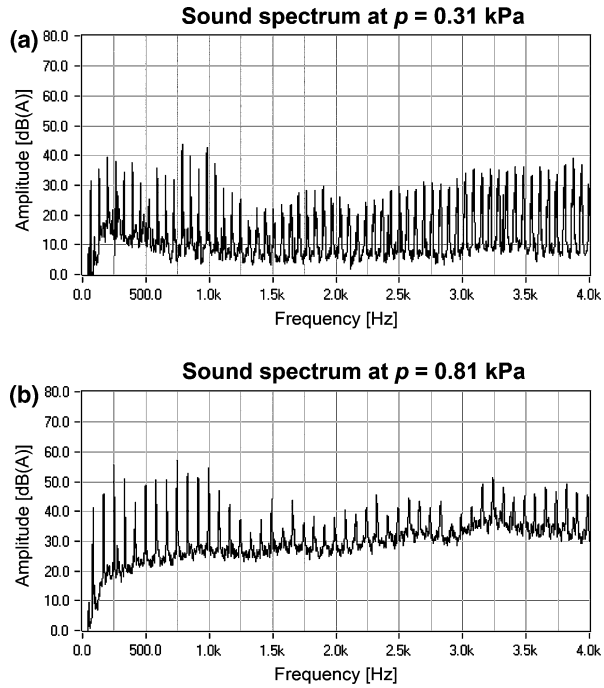

FIGURE 7. Sound spectrum measured for membrane configuration $h_1 l_0 d_2$ at a static driving pressure of (a) 0.31 kPa and (b) 0.81 kPa. SPL measured with a 20 μPa reference and A-weighting filter applied in the time domain; RMS averaging; number of averages = 25; sample rate = 8192 Hz; number of samples = 4096.

if the NHR value is lower than this threshold. Table 3 shows the mean values of the NHR measured. Unfortunately, for some configurations the NHR could not be determined because the recorded sound samples were of poor quality.

Figure 7a shows the sound spectrum for the membrane configuration $h_1 l_0 d_2$ measured at the relatively low driving pressure of 0.31 kPa, while Fig. 7b displays the spectrum for the same configuration at the higher pressure of 0.81 kPa. These spectra are characteristic of the double-membrane concept and illustrate the presence of harmonics, as well as the increased noise level established at higher driving pressures and flow rates.

The videostroboscopic recordings of the oscillating membranes showed the periodical opening and closing of the airway, during which time transversal waves propagated along the membranes from upstream to downstream. It was observed that an increased driving pressure caused the oscillation amplitude as well as the oscillation frequency to increase.

DISCUSSION

We use the data obtained (Table 3) to discuss the influence of the different membrane geometries on the functioning of the potential artificial voice source. Overall, the most marked characteristics of the double-membrane sound source are the high air pressures and low airflow rates required for sound production, in comparison with the requirements based on the physiological situation (Table 2). Furthermore, fundamental frequencies required for the up-scaled model for producing a male voice were not attained. The frequencies measured are higher than expected theoretically (Eq. 2), not only because of the nonlinear dynamics of the oscillating system, but also due to the fact that the metal weights attached to the membrane decreased the membrane's flexibility locally, thereby increasing the stiffness of these membranes.

Influence of Membrane Geometry

The most important geometrical influences on the ranges measured for the aero-acoustic parameters are discussed by comparing the results (Table 3) for the different membrane configurations with each other.

The distance (h) between the membranes (Fig. 2) is the geometrical parameter that mainly influences the size of the flow channel aperture. From the results in Table 3, the influence of the various h_i settings (Table 1) on the measured parameters can be derived by comparing the membrane configurations that have equal l_i and d_i settings. The flow dependence on the aperture height (h) valid for a stationary situation was described with the Bernoulli equation (Eq. 1). Although we are considering a dynamic system with a quasi-stationary flow, the results do reasonably confirm the predicted dependence, as the lowest q ranges were measured for configurations with h_0 (0.0 mm), while configurations with h_1 (0.5 mm) allowed more flow and with h_2 (1.0 mm) the highest q ranges were attained.

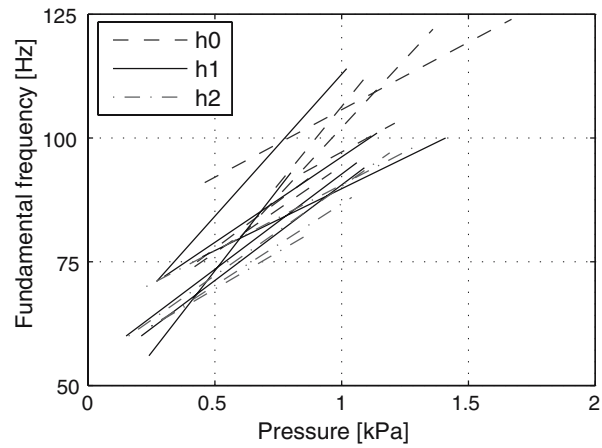

FIGURE 8. Visualization with straight lines of the driving air pressure and fundamental frequency ranges, grouped according to the three different membrane distances h_0 (0.0 mm), h_1 (0.5 mm), and h_2 (1.0 mm) applied.

Figure 8 visualizes the influence of the different h_i settings on the p and f_0 parameters. Configurations with h_0 (0.0 mm) appear to produce sound at higher p ranges than comparable configurations with h_1 (0.5 mm) or h_2 (1.0 mm). Since the h_0 (0.0 mm) membranes are initially in close contact over a large area, the higher driving pressures are most likely caused by higher friction forces and possibly also by higher elastic forces during oscillation, leading to an increased flow resistance. Furthermore, the results and Fig. 8 show that the h_1 (0.5 mm) and h_2 (1.0 mm) settings have a favorable effect on the f_0, as membranes with these configurations produced a sound with an f_0 lower than that produced by comparable membranes with an h_0 (0.0 mm) setting. It is hypothesized, supported by the stroboscopic observations, that during oscillation the available space between the two membranes (Fig. 3) allows them to move past their resting position before they collide, thus allowing for larger oscillation amplitudes and thus longer oscillation cycles at similar airflow velocities. During the experiments, we also observed an oscillation amplitude increase by increasing the airflow rate. As expected from similar nonlinear sound generators, like the brass players' lips or the avian syrinx,[6,9] the membrane excitation increased through a complex interaction of increased aerodynamic, elastic, and collision forces. Besides the oscillation amplitude, also the frequency f_0 increased with increasing flow rate q, which was measured for all configurations. Configurations with h_0 (0.0 mm), except for $h_0l_1d_1$, did not attain the required frequency range of 7 semitones, and their range was generally less than that measured for comparable configurations with h_1 (0.5 mm) or h_2 (1.0 mm).

During one oscillation cycle the airflow channel in the h_0 (0.0 mm) membranes is closed off completely, which hypothetically benefits the excitation of strong and harmonically rich sound pressure waves. Although higher efficiencies were measured, overall, the SPL values were lower and no increased harmonic richness was shown in the sound spectra. The mean NHR values are even the highest for the h_0 (0.0 mm) configurations measured.

When configurations with membrane length l_0 (15.0 mm) are compared with similar configurations comprising the longer membrane l_1 (20.0 mm), configurations with l_0 (15.0 mm) allowed slightly lower p and generally higher q values. This observation appears to support the assumption that friction is important (at least during a part of the oscillation cycle), as expected from the consideration of a Poisseuille flow. Furthermore, except for configuration $h_0l_0d_2$, the shorter membrane (15.0 mm) appears to have a positive effect on the SPL, the amount of semitones, and the efficiency of sound production.

Finally, we consider the membrane width (d) influences. The results in Table 3 show that for the h_0 (0.0 mm) and h_1 (0.5 mm) configurations, the pressure at which oscillation started was the lowest for the configurations comprising membrane width d_2 (9.0 mm). Most likely, the configurations with the largest membrane widths (9.0 mm) start to vibrate at a lower driving pressure because of their lower internal stiffness, as follows from Eq. 2. Furthermore, the membrane width relates to the size of the through-flow area, as shown by Eq. 1, and is therefore expected to influence the flow rate. The hypothesis that membrane width d_2 (9.0 mm) allows higher flow rates is supported by the measurements of the configurations comprised of h_1 (0.5 mm) or h_2 (1.0 mm), in which the q values measured are higher for the configurations with the larger membrane widths.

A lower f_0 was expected for the configurations which have the largest membrane width setting d_2 (9.0 mm), as shown earlier by Eq. 2. These lower f_0 values were measured for the h_0 (0.0 mm) and, to a lesser extent, for the h_1 (0.5 mm) configurations. Furthermore, the membrane width has an effect on the SPL measured; d_2 (9.0 mm) configurations reach a higher SPL than comparable configurations with smaller membrane widths. This increase in sound intensity can be related directly to the increased size of the sound source with increasing membrane width.

Best Membrane for Clinical Application

A VPE based on a double-membrane sound generator can only be used successfully as a new voice source for laryngectomized patients when the aero-acoustics conform to the physiological requirements. For all membrane configurations measured, the p, q and f_0 ranges do not perfectly match the requirements presented in Table 2. However, the individual membrane settings that yield characteristics close to the requirements are; membrane distance h_1 (0.5 mm) and h_2 (1.0 mm), membrane length l_0 (15.0 mm), and membrane width d_2 (9.0 mm). Using the data in Table 3, the three most promising membrane configurations could be selected; $h_1l_0d_2$,, $h_2l_0d_1$, and $h_2l_0d_2$. Measurement data of these configurations is also presented in Fig. 6.

An important requirement for the VPE put in action is that it does not require driving pressures that are higher than the human lungs normally can provide. In normal laryngeal phonation the pressure range is 0.2–1.5 kPa, with an airflow range between 45 and 350 ml/s (under normal speaking conditions).[24] However, if necessary values up to 3 kPa tracheal pressure can be reached. The lowest p range was measured for membrane configuration $h_1l_0d_2$ (Fig 6a). By taking the geometrical scaling factor of two into account, the

theoretically required tracheal pressures will be four times higher (0.6–4.2 kPa) for a down-scaled membrane configuration, see Table 2. These pressures are relatively high, but for the most part within the physiological range. As a related parameter, the down-scaled airflow range (20–79 ml/s) is low compared to laryngeal values. This low flow range is not expected to form a major problem, but it could mean that the patient needs more time to adapt to the new sound source. The other two membrane configurations mentioned, which have larger membrane distances, allow higher flow rates.

The lowest f_0 range was also measured for configuration $h_1 l_0 d_2$ (Fig 6b). Although we aimed for the lowest possible natural oscillation frequency for this up-scaled double-membrane system, we consider the frequencies measured to be too high for producing a male voice. The f_0 range for a down-scaled double-membrane sound generator (240–380 Hz), however, is promising for the restoration of female speech which typically has an f_0 range of about 180–270 Hz.[13] A laryngectomized patient should be able to alter the pitch of her substitute voice sufficiently via changes in the tracheal pressure, as the frequency range of the double-membrane configurations consists of about 8 semitones (Fig. 6b), while the SPL also increased sufficiently under the influence of an increasing driving pressure (Fig. 6c).

Besides the f_0, the multiples of this frequency (harmonics) must be present in the sound signal to enable the pronunciation of distinguishable vowels during speech. Across the whole pressure range, the sound spectra considered most optimal are those measured for membrane configuration $h_1 l_0 d_2$ (Fig. 7). However, at higher driving pressures the noise in the sound produced slightly increases, which is illustrated in Fig. 6d through an increased NHR with increasing driving pressure. Because the NHR values actually stay below the set threshold value, the quality of the sound is considered satisfactory.

Based on these considerations, it is advised that a sound generator design for clinical application uses a membrane geometry based on the up-scaled membrane configuration $h_1 l_0 d_2$. Since the materials used in the final design need to be biocompatible, an alternative material has to be found for the synthetic latex used in this study to manufacture the membranes. Applicable elastomeric materials that exhibit good biocompatible properties are silicone rubber and polyether-based urethane rubbers. These materials are listed in many handbooks (e.g. ref. 22). Silicone rubber is the closest alternative material to latex in terms of ultimate elongation (600–1100%). The ultimate elongation of polyurethane rubber is typically 400–700%, and its modulus of elasticity is generally higher compared to silicone rubbers. The polyurethane rubbers, however, have a much higher tensile strength (28–40 MPa) than silicone rubbers (5–10 MPa), and their tear resistance is also higher in general. Because of the higher strength the polyurethane is considered the most appropriate membrane material which can safely be used in a VPE. In order to generate the low frequencies required for a male voice, it is still necessary to increase the inertia of such a membrane by inserting a high density material like gold or tungsten into the membrane. The membranes should be attached in a tube-shaped housing made of a stainless steel or a biocompatible plastic. The design of the housing should enable a secure placement of the device into regular TE shunt valves.

CONCLUSIONS

Two membranes, placed parallel to each other inside a circular housing, perform self-sustained oscillations under the influence of a flow of air, thereby generating a sound. The double-membrane concept has sound characteristics that makes it suitable for replacing the voice after total laryngectomy. The sound produced contains multiple harmonics, and the fundamental frequency and SPL can be modulated via changes in the driving pressure, thus enabling intonation during speech. However, a VPE based on this sound generator, restricted in size by the shaft diameter of regular TE shunt valves, requires a relatively high driving air pressure, and produces sound with fundamental frequencies only suitable for female voices.

The membranes of the VPE should be fixed with a distance of at least 0.25 mm between them, otherwise too much air pressure is necessary for the sound production. Furthermore, the optimal membrane length is 7.5 mm, in combination with a membrane width of 4.5 mm. These settings allow for the lowest fundamental frequencies produced.

This study presents the feasibility of a VPE based on the double-membrane sound generator for female laryngectomized patients. Before *in vivo* experiments can be performed, a VPE should be manufactured on a true-scale, using biocompatible materials.

ACKNOWLEDGMENT

This study is part of the Eureka Newvoice 2614 project, financed by a grant from Senter (TSIN 1015).

REFERENCES

[1]Blood, G. W. Fundamental frequency and intensity measurements in laryngeal and alaryngeal speakers. *J. Commun. Disord.* 17(5):319–324, 1984.

[2]Cheesman, A. D., J. Knight, J. McIvor, and A. Perry. Tracheo-oesophageal 'puncture speech'. An assessment technique for failed oesophageal speakers. *J. Laryngol. Otol.* 100(2):191–199, 1986.
[3]Collier, R. Physiological correlates of intonation patterns. *J. Acoust. Soc. Am.* 58(1):249–255, 1975.
[4]De Vries, M. P., M. C. Hamburg, H. K. Schutte, G. J. Verkerke, and A. E. P. Veldman. Numerical simulation of self-sustained oscillation of a voice-producing element based on Navier–Stokes equations and the finite element method. *J. Acoust. Soc. Am.* 113(4 Pt 1):2077–2083, 2003.
[5]De Vries, M. P., A. Van der Plaats, M. Van der Torn, H. F. Mahieu, H. K. Schutte, and G. J. Verkerke. Design and in vitro testing of a voice-producing element for laryngectomized patients. *Int. J. Artif. Organs.* 23(7):462–472, 2000.
[6]Fee, M. S., B. Shraiman, B. Pesaran, and P. P. Mitra. The role of nonlinear dynamics of the syrinx in the vocalizations of a songbird. *Nature* 395(6697):67–71, 1998.
[7]Fletcher, N.H. Bird song – a quantitative acoustic model. *J. Theor. Biol.* 135(4):455–481, 1988.
[8]Fletcher, N. H. Autonomous vibration of simple pressure-controlled valves in gas-flows. *J. Acoust. Soc. Am.* 93(4):2172–2180, 1993.
[9]Fletcher, N. H. Excitation mechanisms in woodwind and brass instruments. *Acustica* 43:65–72, 1979.
[10]Fletcher, N. H. and T. D. Rossing. *The Physics of Musical Instruments*. Springer-Verlag, New York, 1998.
[11]Greenewalt, C. H.. *Bird Song: Acoustics and Physiology*. Smithsonian institute, Washington DC, 1968.
[12]Hagen, R., K. Berning, M. Korn, and F. Schon Stimmprothesen mit tonerzeugendem Metallzungen-Element–A experimentelle und erste klinische Ergebnisse. [Voice prostheses with sound-producing metal reed element–an experimental study and initial clinical results]. *Laryngorhinootologie* 77(6):312–321, 1998.
[13]'t Hart J., R. Collier, and A. Cohen. *A Perceptual Study of Intonation: An Experimental - Phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press, 1990, 212 pp.
[14]Herrmann, I. F., S. Arca Recio, and J. Algaba. A new concept of surgical voice restoration after total laryngectomy: The female voice. In: I. F. Herrmann (ed.), *The Second International Symposium on Laryngeal and Tracheal Reconstruction*, 1996, pp. 263–266.
[15]Hilgers, F. J. and P. F. Schouwenburg. A new low-resistance, self-retaining prosthesis (Provox) for voice rehabilitation after total laryngectomy. *Laryngoscope* 100(11): 1202–1207, 1990.
[16]Hirschberg, A., R. W. A. Vandelaar, J. P. Marroumaurieres, A. P. J. Wijnands, H. J. Dane, S. G. Kruijswijk, and A. J. M. Houtsma. A quasi-stationary model of air-flow in the reed channel of single-reed woodwind instruments. *Acustica* 70(2):146–154, 1990.
[17]Mahieu, H. F., A. A. Annyas, H. K. Schutte, and E. J. Vanderjagt. Pharyngoesophageal myotomy for vocal rehabilitation of laryngectomees. *Laryngoscope* 97(4):451–457, 1987.
[18]Mahieu, H. F. *Voice and Speech Rehabilitation following Laryngectomy*. Groningen: University of Groningen, PhD Thesis, 1988.
[19]Nijdam, H. F., A. A. Annyas, H. K. Schutte, and H. F. Leever. A new prosthesis for voice rehabilitation after laryngectomy. *Arch. Otorhinolaryngologie* 237:27–33, 1982.
[20]Panje, W.R. Prosthetic vocal rehabilitation following laryngectomy. The Voice Button. *Ann. Otol. Rhinol. Laryngol.* 90:116–120, 1981.
[21]Qi, Y. and B. Weinberg Characteristics of voicing source waveforms produced by esophageal and tracheoesophageal speakers. *J. Speech Hear Res.* 38(3):536–548, 1995.
[22]Ratner, B. D., A. S. Hoffman, F. J. Schoen, and J. E. Lemons (eds). *Biomaterials Science: An Introduction to Materials in Medicine*, 2nd edn. Amsterdam: Elsevier Academic Press, 2004, 864 pp.
[23]Robbins, J., H. B. Fisher, E. C. Blom, and M. I. Singer. A comparative acoustic study of normal, esophageal, and tracheoesophageal speech production. *J. Speech Hear Disord.* 49(2):202–210, 1984.
[24]Schutte, H. K. *The Efficiency of Voice Production*. Groningen: University of Groningen, PhD Thesis, 1980, 194 pp.
[25]Singer, M. I. and E. D. Blom. An endoscopic technique for restoration of voice after laryngectomy. *Ann. Otol. Rhinol. Laryngol.* 89:529–533, 1980.
[26]Šram, F. Diagnostik der Erkrankungen der Sprechorgane bei Blasinstrumentalisten. In: *Das Instrumentalspiel*. Prague: Verlag Doblinger, 1989, pp. 231–235.
[27]Tack, J. W. *Design of A Voice-producing Element*. Enschede: University of Twente, Doctoral Thesis, 2003, 77 pp.
[28]Van den Berg, J. W. Physiology and physics of voice production. *Acta Physiol. Pharmacol. Neerl.* 5:40–55, 1956.
[29]Van der Torn, M., M. P. de Vries, J. M. Festen, I. M. Verdonck-de Leeuw, and H. F. Mahieu. Alternative voice after laryngectomy using a sound-producing voice prosthesis. *Laryngoscope*. 111(2):336–346, 2001.