

An *In-vitro* Test Set-Up for Evaluation of a Voice-Producing Element Under Physiologic Acoustic Conditions

ARJAN VAN DER PLAATS,¹ HARM K. SCHUTTE,¹ FRITS J. M. VAN DER EERDEN,² MARTIN P. DE VRIES,¹
HANS F. MAHIEU,⁴ and GIJSBERTUS J. VERKERKE^{1,3}

¹Department of Biomedical Engineering, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands;

²Department of Mechanical Engineering, University of Twente, Enschede, The Netherlands; ³Department of Biomechanical Engineering, University of Twente, Enschede, The Netherlands; and ⁴Department of ORL, VU Medical Centre, Amsterdam, The Netherlands

(Received 20 April 2004; accepted 19 January 2006; published online: 22 April 2006)

Abstract—To improve the voice quality of laryngectomised patients, a voice-producing element has been developed. Prior to *in vivo* testing we constructed and validated an *in-vitro* test set-up, consisting of a physical model of the subglottal tract and three physical models of the vocal tract, for the vowels /a/, /i/ and /u/ to evaluate the voice-producing element under physiologic acoustic conditions. To meet acoustic conditions described in the literature, we determined the appropriate dimensions of these physical models, using a numerical model of the pressure perturbation in rigid tubes. The numerical model showed that an acoustic equivalent of the subglottal tract could be obtained with a three-tube system and an end impedance. Vocal tract models could be constructed using two- and four-resonator tubes. The physical models were built and successfully validated according to the human acoustic properties. The developed *in-vitro* set-up can now be applied to test voice-producing elements or vocal fold models under physiologic acoustic conditions.

Keywords—Laryngectomy, Acoustics, Vocal tract, Subglottal tract.

NOMENCLATURE

A_J	Cross-sectional area of tube J (m^2)
C_0	Speed of sound (m s^{-1})
i	$i = \sqrt{-1}$, imaginary number (-)
J, j	Index (-)
k	$k = \omega/c_0$, wave number (m^{-1})
L	Length of a tube (m)
M_V	Mass variation in a volume (kg)
n_V	Polytropic coefficient of volume V (-)
p	Pressure perturbation (Pa)
\hat{p}_A	Complex amplitude (Pa)
Q	Mass flow (kg s^{-1})
s	$s = l\sqrt{\omega\rho_0/\mu}$, shear wave number (-)

t	Time (s)
u	Velocity perturbation (m s^{-1})
V	Volume (m^3)
Z_L	$Z_L = \frac{1}{\rho_0 c_0} \frac{p}{u}$, Dimensionless acoustic impedance (-)
γ	$\gamma = C_p/C_v$, ratio of specific heats (-)
Γ	Viscothermal wave propagation coefficient (-)
ρ_0	Mean density (kg m^{-3})
ω	Angular frequency (rad s^{-1})

INTRODUCTION

Surgical treatment of a malignant tumor in the larynx sometimes requires a laryngectomy involving a reconstruction of the pharynx after removal of the larynx. The trachea is sutured to the skin allowing respiration via a tracheostoma. One of the consequences of this kind of treatment is that the vocal folds, as part of the larynx, are removed and voice production is no longer possible. Voice restoration is typically obtained by setting soft tissue structures located at the entrance of the esophagus (PE-segment or pseudoglottis) into vibration using air, which is brought into the esophagus. Nowadays, this air is directed from the lungs into the esophagus via a shunt valve (often incorrectly called voice prosthesis) situated in a surgically created fistula in the tracheo-esophageal wall. Closing the tracheostoma during exhalation directs airflow through the shunt valve into the esophagus. The flow pulses generated by the PE-segment serve as the basic acoustic excitation of the vocal tract cavities, consisting of the air space formed by the pharynx and oral and nasal cavities. The resonance frequencies of a specifically shaped vocal tract select a subset of frequencies from the frequency spectrum produced by the sound source. In this way the vowels of speech are formed.

The speech quality, however, is often poor. Speech produced in this way has a low fundamental frequency, which is, especially for women, undesirable. Also the ability for

Address correspondence to: Gijsbertus J. Verkerke, Department of BioMedical Engineering, University Medical Center Groningen, University of Groningen, P.O. Box 196, 9700 AD, Groningen, The Netherlands. Electronic mail: g.j.verkerke@med.umcg.nl

frequency variation diminishes. To improve speech quality, we have developed a voice-producing element that fits into the shunt valve and produces a sound with a more appropriate fundamental frequency.^{11,12}

One of the most important requirements for the voice source is that the resulting sound signal must be independent of the acoustic loads formed by the vocal and subglottal tract,^{1,13} in order to avoid involuntary fundamental frequency changes during normal speech. The subglottal tract is formed by the trachea and lungs. Since the glottis (the area between the vocal folds) is not present anymore the voice-producing element should be considered in this respect as neo-glottis to validate the use of the term ‘subglottal tract’. The acoustical influence of the esophagus can be neglected, since it is closed during speech.

To investigate the influences of subglottal tract and the various vowel configurations of the vocal tract on the functioning of a voice-producing element, or, alternatively, any other sort of artificial sound source, a test set-up with physiologically realistic acoustic representations of subglottal and vocal tract is required.

Goal of this study was to determine the appropriate properties of a physical model of the subglottal and vocal tract to represent realistic acoustical behaviour.

MATERIALS AND METHODS

General Approach

The physical models representing the vibrating column of air of both vocal and subglottal tract were considered to be a combination of connected tubes with rigid walls.⁷ The tubes have different geometry and end conditions. The appropriate geometry and end impedance was determined using a numerical model of the pressure perturbation in a series of connected tubes. The basic theory of this numerical model is described in the appendix.

Vocal Tract

The vocal tract is defined as the column of air formed by the pharynx and oral and nasal cavities. The acoustic load presented by any given vowel configuration of the vocal tract tends to force the fundamental frequency of the sound source (voice-producing element) towards the resonance frequency of the vocal tract.^{1,13} To investigate this load, three physical models of the vowels /a/, /i/ and /u/ were made. These three vowels are defined by their formant frequencies of the vowels /a/, being 700 and 1200 Hz, /i/, being 300 and 2300 Hz and /u/, being 300 and 700 Hz, based on the vowel data of Peterson and Barney.⁸

To determine the geometry of the three physical models, consisting of several concentrically connected tubes (see Fig. 1), a numerical model was developed, based on the theory described in the appendix.

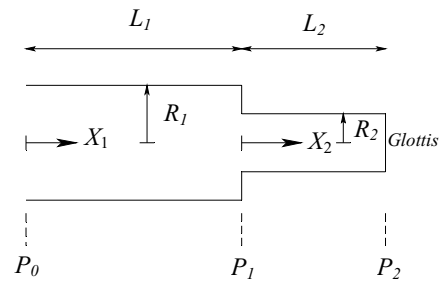


FIGURE 1. Vocal tract model (R_i : radius i th tube, L_i : length i th tube).

In a first approach, in the recursive formulation (see (A.9)) it was assumed that the lower side of the coupled tubes, the side where the voice producing prosthesis is situated, acted as a closed end, i.e. acoustically hard. The transfer function, being the ratio of the pressure perturbations at both ends of a coupled two-tube system, then becomes:

$$H_{\text{tot}} = \frac{p_2}{p_1} \frac{p_1}{p_0} \quad (1)$$

The two transfer functions $H_{10} = \frac{p_1}{p_0}$ and $H_{21} = \frac{p_2}{p_1}$ follow from the recursive formulation (A.9):

$$H_{21} = [\cosh(\Gamma kL)_2]^{-1} \quad (2)$$

$$H_{10} = \left[\cosh(\Gamma kL)_1 + \frac{A_2 \sinh(\Gamma kL)_1}{A_1 \sinh(\Gamma kL)_2} \right] \times \left[\cosh(\Gamma kL)_2 - \frac{p_2}{p_1} \right]^{-1} \quad (3)$$

The total transfer function H_{tot} becomes thus:

$$H_{\text{tot}} = [\cosh(\Gamma kL)_2]^{-1} \left[\cosh(\Gamma kL)_1 + \frac{A_2 \sinh(\Gamma kL)_1}{A_1 \sinh(\Gamma kL)_2} \right] \times \left[\cosh(\Gamma kL)_2 - \frac{1}{\cosh(\Gamma kL)_2} \right]^{-1} \quad (4)$$

The first two peaks found in the transfer function H must correspond with the first two formants or resonance frequencies of the vowel. By variation of the number of tubes and their geometric properties (radius and length), tube systems were determined having approximately the same resonance frequencies as the three considered cardinal vowels, /a/, /i/ and /u/ as found by Peterson and Barney.⁸

Validation

The vowel models were built and experimentally validated by applying a sinusoidal sound sweep at the side of the voice-producing element and measuring the resulting sound signal 5 cm above and 5 cm besides the outflow opening of the vowel. The sine sweep was applied with a condenser microphone (type 4132, Brüel & Kjær, Nærum,

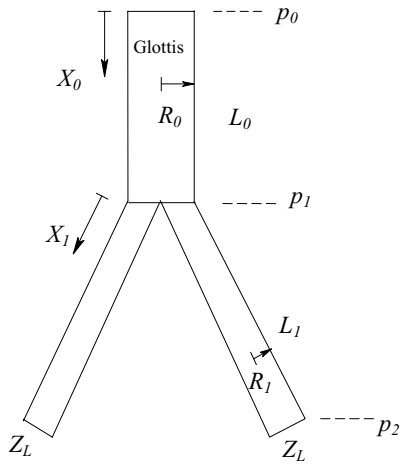


FIGURE 2. Description of the parameters used for the numerical simulation of the subglottal tract (for explanation of the parameters, see text).

Denmark) and a beat frequency oscillator (type 1022, Brüel & Kjær, Nærum, Denmark). The sound was recorded with a condenser microphone (type 4134, Brüel & Kjær, Nærum, Denmark) and data acquisition was performed with Lab-view for Windows 3.01 (National Instruments, Austin, Texas, USA) on a personal computer.

The resulting resonance frequencies were compared with the first two formants of the three vowels.

Subglottal Tract

The subglottal tract is defined as the tube system below the vocal folds, i.e. the trachea and the lungs. In combination with a vibrating sound source (vocal folds or voice-producing element), the subglottal tract can be considered as a vibrating air column. To determine the geometry of the physical model of the subglottal tract, a numerical model was built. The subglottal tract was assumed to be composed of several tubes; a main tube with length L_0 , representing the trachea, coupled with two symmetrical connected tubes with length L_1 . The latter tubes have an acoustically soft boundary condition (dimensionless end-impedance Z_L) which stands for the rest of the lungs, see Fig. 2.

The recursive formulation (A.9) of the acoustical behaviour in a system of coupled tubes as shown in Fig. 2, now becomes:

$$\frac{p_j}{p_{j-1}} = \left[\cosh(\Gamma kL)_j - \frac{\Gamma n_V \sinh(\Gamma kL)_j}{i\gamma A_j} \right. \\ \times \left(-2 \frac{i\gamma A_{j+1}}{\Gamma n_V \sinh(\Gamma kL)_{j+1}} \right. \\ \left. \left. \times \left\{ \cosh(\Gamma kL)_{j+1} - \frac{p_{j+1}}{p_j} \right\} \right) \right]^{-1} \quad (5)$$

The total transfer function can again be defined as:

$$H_{\text{tot}} = \frac{p_2 p_1}{p_1 p_0} \quad (6)$$

The boundary condition at the end of the last tubes is formed by an acoustical impedance Z_L . The transfer function of such a tube is:

$$H_{21} = \frac{p_2}{p_1} = \left[\cosh(\Gamma kL)_1 + \frac{1}{Z_L} \sinh(\Gamma kL)_1 \right]^{-1} \quad (7)$$

and the transfer function of the first tube becomes:

$$H_{10} = \frac{p_1}{p_0} = \left[\cosh(\Gamma kL)_1 + 2 \frac{A_2 \sinh(\Gamma kL)_1}{A_1 \sinh(\Gamma kL)_2} \right. \\ \left. \times \left\{ \cosh(\Gamma kL)_2 - \frac{p_2}{p_1} \right\} \right]^{-1} \quad (8)$$

The total transfer function can now thus be written as:

$$H_{\text{tot}} = \left[\cosh(\Gamma kL)_1 + 2 \frac{A_2 \sinh(\Gamma kL)_1}{A_1 \sinh(\Gamma kL)_2} \left\{ \cosh(\Gamma kL)_2 \right. \right. \\ \left. \left. - \left[\cosh(\Gamma kL)_1 + \frac{1}{Z_L} \sinh(\Gamma kL)_1 \right]^{-1} \right\} \right]^{-1} \quad (9)$$

The geometry of the tube system was chosen in such a way that its impedance Z_L of Eq. (9) fitted to the measurements of the acoustical properties of the subglottal tract that have been reported by Ishizaka *et al.*⁶ They measured the acoustic input impedance of the subglottal tract at the tracheostoma of Japanese laryngectomised patients. After correction for Western lung geometry they found peaks in the input impedance spectrum at 615, 1355 and 2110 Hz with impedance values of 4 kg acoustic ohms average, which is also in agreement with data found by Cranen and Boves.³

Validation

A physical model according to the specifications was built, and the impedance spectrum was validated using an impedance tube.^{4,9} An impedance tube applies a random noise signal to the physical model of the subglottal tract, and by measuring the transfer signal with two pressure transducers, the impedance of the model is calculated.

RESULTS

Vocal Tract

The geometry of the physical models for the vowels /a/, /i/ and /u/, was determined by the numerical simulation model to be a two-tube system for the /a/ ($L_1 = 98$ mm, $R_1 = 11.5$ mm, $L_2 = 80$ mm, $R_2 = 30$ mm), a two-tube system for the /i/ ($L_1 = 71$ mm, $R_1 = 13.5$ mm, $L_2 = 61$ mm, $R_2 = 5$ mm) and a four-tube system for the /u/ ($L_1 = 64$ mm, $R_1 = 17$ mm, $L_2 = 5$ mm, $R_2 = 7.5$ mm, $L_3 = 60$ mm, $R_3 = 17$ mm, $L_4 = 26$ mm, $R_4 = 5$ mm)

Table 1. Dimensions of the tube systems representing the three vowels examined.

/a/			L1=98	
			L2=80	
			R1=11.5	
			R2=30	
/i/			L1=71	
			L2=61	
			R1=13.5	
			R2=5	
/u/				
			L1=64; L2=55	
			L3=60; L4=26	
			R1=17; R2=7.5	
			R3=17; R4=5	

Note: all dimensions in 10^{-3} m.

(Table 1). Subsequently, the three physical models were built from polycarbonate, see Fig. 3. The results of the validation study are shown in Table 2, together with the simulated results.

Subglottal Tract

Using Eq. (9) for Z_L , the optimal geometry of the tube system depicted in Fig. 2 was determined as: $L_0 = 69 \times 10^{-3}$ m, radius $R_0 = 10 \times 10^{-3}$ m and $L_1 = 124 \times 10^{-3}$ m with radius $R_1 = 16 \times 10^{-3}$ m. The dimensionless impedance at the end of the system was determined as 0.2. This could be realised by a thin layer of glass wool (permeable for air) at the end of the tubes, as was confirmed by measurements with an impedance tube.^{4,9} Subsequently, a physical model was constructed from polycarbonate, see Fig. 4.

Figure 5 shows the simulated spectrum of the tube system representing a subglottal tract model (resonance fre-

quencies of 593 and 1300 Hz) and the measurement of this model with an impedance tube (resonance frequencies of 565 and 1288 Hz).

Complete Set-Up

The complete set-up for measuring the aerodynamic properties of an artificial sound source, in particular the voice-producing element, *in-vitro* is depicted in Fig. 6. It consists of a pressure vessel, coated inside with a sound-absorbing material to inhibit sound-reflections of the vessel wall. The physical model of the subglottal tract is situated inside this vessel. On top of the subglottal tract model a holder with the sound source, optionally situated inside a shunt valve, can be placed. The physical model of the vocal tract is placed over this holder.

Air is applied by means of an air cylinder connected to the vessel. The flow is directed through the subglottal tract model to the artificial sound source. The sound produced



FIGURE 3. Physical models of the vowels /a/, /i/ and /u/ (from left to right).

travels through the vocal tract model, which alters the sound signal into vowels.

DISCUSSION

As can be seen from Table 1 there are some minor deviations between the desired formant frequencies and the obtained ones. The differences can be explained by the difference in realistic and modelled boundary conditions at both ends of the coupled tubes. At the glottal side during validation a microphone membrane was present instead of the assumed acoustically hard wall in the numerical model. On the other side an open end was assumed, while in reality a radiation boundary is present. The results of the numerical model also match the description of the tube systems, described by Fant,⁵ so the three different physical tube systems are successfully validated. The numerical model that forms the basis of these tube systems can be beneficial for further studies, for instance to determine

the appropriate dimensions of physical models of other vowels.

It was shown that the acoustic load of the trachea and main bronchi could be described with just a three-tube system. The influence of the rest of the lungs, e.g. alveoli, could be approximated with an end impedance, here a thin layer of glass wool. The numerical description of the pressure distribution in the lungs has proven to be an appropriate tool for this kind of modelling.

Measurements of the physical model of the subglottal tract using the impedance tube showed that there is broad

Table 2. Measured and simulated resonance frequencies (in Hz) of the vowels /a/, /i/ and /u/.

		Measured frequency (Hz)	Simulated frequency (Hz)
/a/	F1	600	690
	F2	1000	1110
/i/	F1	250	295
	F2	2100	2300
/u/	F1	190	205
	F2	560	610



FIGURE 4. Physical model of the subglottal tract.

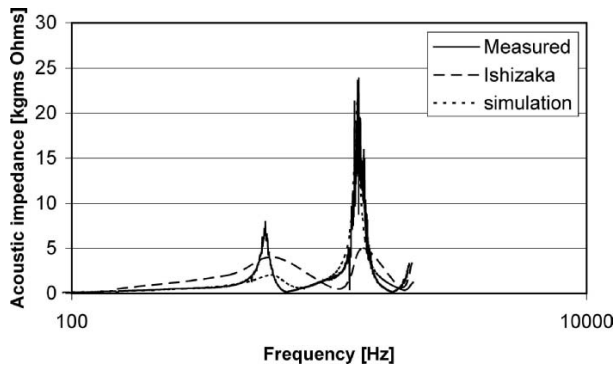


FIGURE 5. Impedance spectra of Ishizaka, simulated subglottal tract model and validation of the model with an impedance tube. The x-axis is given in logarithmic units.

agreement between the theoretical results and the measured results. The model gives a good description of the acoustic behaviour. Here the excitation at the glottis was constrained by the impedance tube instead of the assumed acoustically hard wall, resulting in a minor frequency peak shift. However, measurements and numerical model both show frequency peaks with a smaller bandwidth than the data of Ishizaka *et al.*⁶ (Fig. 5). This is probably due to the fact that both physical and numerical model assume rigid walls, where the data of Ishizaka obviously is based on tubes with elastic walls (human lungs). This difference

in damping may result also in lower resonance frequencies in the simulation and experiments compared to the peaks found by Ishizaka.

CONCLUSIONS

Using a numerical description of pressure perturbations in rigid tubes, we showed that a three-tube system with an end impedance presents an acoustical equivalent of the human lungs. In combination with physical models of different vowels formed by the vocal tract, we created an *in-vitro* measuring set-up, which makes it possible to analyse (artificial) sound sources under physiologic acoustic loads.

Appendix: Derivation of a Recursive Formulation of the Acoustic Behaviour of Coupled Tubes, Using the Theory of Bergh and Tijdeman² and Tijdeman¹⁰

We assume the one-dimensional pressure $p(x)$ and average velocity perturbation $u(x)$, including the viscothermal wave propagation for tubes, see Fig. A.1, to be:

$$p(x) = \hat{p}_A e^{\Gamma kx} + \hat{p}_B e^{-\Gamma kx} \quad (\text{A.1})$$

and

$$u(x) = -\frac{i\gamma}{\rho_0 c_0 \Gamma n} (\hat{p}_A e^{\Gamma kx} - \hat{p}_B e^{-\Gamma kx}) \quad (\text{A.2})$$

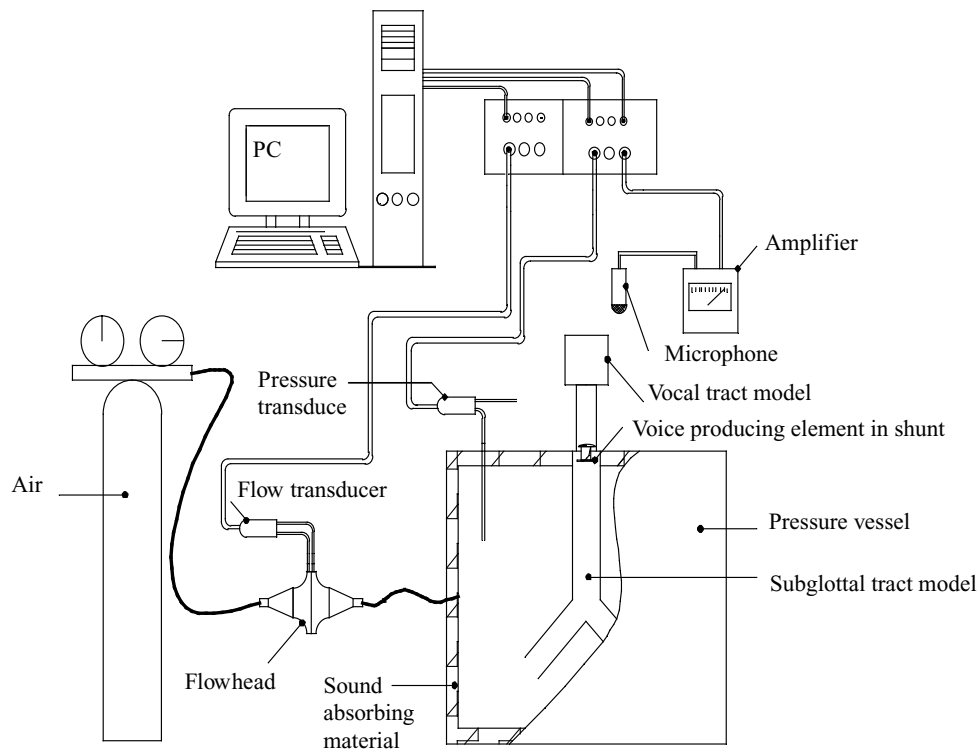


FIGURE 6. The *in-vitro* test set-up.

- ⁷Mol, H. *Fundamentals of Phonetics: II. Acoustical Models Generating the Formants of the Vowel Phonemes*. The Hague: Mouton & Co, 1970, 131 p.
- ⁸Peterson, G. E., and H. L. Barney. Control methods used in a study of the vowels. *J. Acoust. Soc. Am.* 24:175–184, 1952.
- ⁹Pierce, A. D. *Acoustics; An Introduction to its Physical Principles and Applications*. New York: McGraw-Hill, 1989, 678 p.
- ¹⁰Tijdeman, H. On the propagation of sound waves in cylindrical tubes. *J. Sound Vib.* 39:1–33, 1975.
- ¹¹Torn, van der M., I. M. Verdonck-de Leeuw, J. M. Festen, M. P. de Vries, and H. F. Mahieu. Female-pitched sound-producing voice prostheses—initial experimental and clinical results. *Eur. Arch. Otorhinolaryngol.* 258:397–405, 2001.
- ¹²Vries, M. P. de, A. van der Plaats, M. van der Torn, H. F. Mahieu, H. K. Schutte, and G. J. Verkerke. Design and *in-vitro* testing of voice producing element for laryngectomised patients. *Artif. Org.* 23(7):462–472, 2000.
- ¹³Wien, M. Über die Rückwirkung eines resonirenden Systems [On the counteraction of resonating systems]. *Ann. Physik. Chem.* 61:151–189, 1897.