TAMPERE UNIVERSITY OF TECHNOLOGY

HERKKO MATTILA
Tracheal Breath Sounds
- instrumentation and origin

LICENTIATE THESIS

Examiners: Prof. Hannu Eskola and
Docent Päivi Piirilä
Subject approved by the faculty council
of Science and Environmental Engi-
neering 23.6.2010

TIIVISTELMÄ

Tässä työssä tehtiin kirjallisuuskatsaus trakeasta kuultavien hengitysäänien instrumentointiin, normaaleihin ja epätavallisiin trakeasta kuultaviin ääniin. Katsauksessa havaittiin, että hengitysäänten rekisteröintiin käytettävien sensoreiden suunnittelulla voidaan vaikuttaa rekisteröityihin ääniin. Erityisesti ilmakytkettyjen kontaktimikrofonien ilmakammiolla voi olla suuri vaikutus mitattuihin ääniin. Normaalien trakeahengitysäänien alkuperä on aeroakustinen. Ilmavirtauksen nopeuden kasvaessa laminaarinen virtaus irtautuu henkitorven seinämästä ja virtaukseen syntyy pyörteitä. Virtaus on muuttunut turbulenttiseksi. Kun nämä pyörteet kulkeutuvat kapeikon läpi, mikä voi olla pelkkä muutos henkitorven poikkipinta-alassa, osa pyörteen liike-energiasta muuttuu akustiseksi häiriöksi. Tämän lisäksi kirjallisuuskatsauksessa tutkittiin epätavallisten trakeasta kuultavien hengitysäänten syntyä. Avoimen, osittain ahtautuneen ja kokonaan ahtautuneen hengitystien tapaukset tarkasteltiin erikseen Sterling-vastusmallin avulla. Kuorsausta ja kudos värinää havaittiin tilanteessa, jolloin hengitystie oli osittain ahtautunut. Pirkanmaan sairaanhoitopiirin unilaboratoriossa käytetylle trakeaäänien rekisteröintilaitteistolle tehtiin laatumittauksia. Mittauksissa havaittiin, että mikrofonissa oli hieman harmonista vääristymää ja käytetyn äänikortin taajuusvaste ei ollut suora, kun näytteenottotaajuutena oli 11kHz. Yhteenvetona voidaan sanoa, että eri tutkimusryhmien tulisi käyttää laadultaan korkeatasoisia mittalaitteita, kun trakeasta kuuluvia hengitysääniä rekisteröidään. Jos näin toimittaisiin, korkealaatuinen trakeahengitysääni voisi tuoda uutta tietoa unen hengitysfysiologiaan liittyen. Nykyisessä tilanteessa eri tutkimusryhmien mittaukset eivät ole täysin vertailukelpoisia keskenään ja tämän vuoksi tieteellistä luotettavuutta on vaikea saavuttaa.

ABSTRACT

This thesis consists of a literature review of tracheal breath sound instrumentation and a research on the origin of normal and adventitious tracheal breath sounds (TBS). It was discovered that the design of the tracheal sound sensors may significantly affect the measured sound. Especially air coupled contact sensors may have several air chamber configurations which affect the measurement results. The creation mechanism of normal tracheal breath sounds was found to be aeroacoustic. When the velocity of airflow increases, the laminar flow is separated from the tract wall and generates vortices. The flow becomes turbulent. When these vortices convect through a constriction, which may only be a change in a cross-sectional area of the respiratory tract, a part of the vortex kinetic energy is dissipated and acoustic disturbance is created. In addition, generation mechanisms of adventitious sounds were studied. The cases of an open tract, a partially collapsed tract and a totally collapsed tract were explained with the Sterling resistor model. The snoring sounds and flutter were found in the partially collapsed case. Quality measurements for the tracheal sound recording system used at the Sleep Laboratory of the Pirkanmaa hospital district (PSHP) were performed. It was discovered that microphones have some harmonic distortions and that the frequency responses of the sound card used was not flat when a 11kHz sampling rate was used. As a conclusion it can be said that different research groups should use high fidelity instruments when recording tracheal breath sounds. If this is done with high quality, tracheal breath sound recordings might give some new information related to the breathing physiology of human beings during sleep. In the present situation, the measurements of different research groups are not totally comparable with each other, and therefore scientific reliability is hard to achieve.

ACKNOWLEDGEMENTS

CONTENTS

NOMENCLATURE

AMG acoustic myography, page 5

AR   acoustic rhinometry, page 91

ARR   autoregressive, page 89

BAI   bioacoustic insulator, page 35

BATT bioacoustic transducer tester, page 29

CR   stereo cross-talk, page 97

dB   decibel, page 15

DR   dynamic range, page 97

DSP   digital signal processing, page 1

EMG   electro myography, page 5

FFT   Fast Fourier Transform, page 1

FR   frequency response, page 16

FS   full-scale, page 97

HeLSA Helsinki lung sound analyzer, page 94

HL   hearing level, page 113

IDW   initial deviation width, page 31

IMD   intermodulation distortion, page 97

LPC   linear predictive coding, page 89

M   Mach number, page 39

MMG   mechanomyography, page 5

NAM   non-acoustic murmur, page 6

OSA   obstructive sleep apneas, page 89

OSAS obstructive sleep apnea syndrome, page 79

PCG   phonocardiogram, page 5

PSHP Pirkanmaa hospital district, page 95

RMAA RightMark Audio Analyzer, page 95

RR   respiration rate, page 130

SDB   sleep disordered breathing, page 79

SNR   signal-to-noise ratio, page 14

SPL   sound pressure level, page 110

TAR   tissue-to-air ratio, page 35

TBS   tracheal breath sounds, page 37

THD+N total harmonic distortion plus noise, page 97

TUT   Tampere university of technology, page 95

Z   acoustic impedance, page 3

1. INTRODUCTION

This thesis presents a combination of article reviews and a series of measurements related to tracheal sound recording instruments and the origin of tracheal breathing sounds. The properties of sound, breathing sounds and sound transducers[1] are introduced. Chapter 2 is a review of respiratory sound transducers. The parameters affecting the sensitivity of transducers are studied in particular. Chapter 3 is an article review of normal tracheal breath sounds. The creation mechanisms and characteristics of tracheal breath sounds are studied. Chapter 4 is also an article review; the adventitious tracheal sounds are studied instead of normal tracheal breath sounds. As in the previous chapter the origin and characteristics of the adventitious sounds are reviewed. The methods of mathematical modelling, *in vitro* measurements and *in situ* measurements have all been used in reviews. Chapter 5 presents the tracheal sound recording system HeLSA, which has been us at the Sleep Laboratory of Tampere University Hospital. The results of the quality measurements done with the system are presented and some samples of tracheal sounds measured with the HeLSA are presented as well.

The method of auscultation with a stethoscope has been used since the 19th century.[1] The major steps of development were taken in the 1950's,[1, 2, 3] when the auscultated sounds were recorded electronically and in 1970 when these sounds were recorded digitally[1]. Since the 1980's, digital signal processing (DSP) has rapidly evolved, enabling data analysis with different methods, such as spectrum analysis by using fast Fourier transform (FFT). While this quantitative analysing has reached large popularity, the importance of measuring instruments has increased as well. Therefore, there is a need for deeper understanding of how these recording instruments work and what kinds of issues affect their properties. Unfortunately, the majority of the research groups publishing their data are not familiar with the properties of their instruments. Researchers may even draw wrong conclusions from their observations if the physiological signal is mixed with the effects of the measuring instruments. In general, the amount of published articles related exclusively to the properties of breath sound sensors is limited.

---

[1]Etymology: Latin transducere to lead across, transfer

2.   A REVIEW OF RESPIRATORY SOUNDS TRANSDUCERS

2.1   Sound in a human body and breath sound transducers

The basics of sound production can be most easily understood by studying the theory of musical instruments. They can be roughly divided into two major groups based on the method that they use to produce the sound. The first group consists of those instruments that produce sound when the material is vibrated mechanically, like drums, the violin or the guitar. When the mass is set to vibrate, the air molecules accelerate periodically. As a result, a density gradient occurs and it starts to propagate as a pressure wave. The second group consists of instruments which manipulate the passing air flow, like the organ and the flute. In a flute the flow is disturbed and turbulence is created. A part of the kinetic energy of the turbulent flow is transferred into acoustic energy, thus creating a sound. The human body uses these same mechanisms to produce sounds. For example heart sounds are produced by mechanical vibrations whereas breathing sounds are a result of the effects of air flow in the respiration tract. [4]

The sounds observed by humans are gradients of a pressure wave in the air. The air is a mixture of gases and therefore the idea of sound is commonly associated with cases in which the pressure wave is travelling in a gaseous phase. The pressure wave can propagate across the boundary layer of gaseous and solid material. When the wave is in a solid material, the concept of sound is treated as a mechanical wave. However, the same physical phenomenon can be observed with the propagating pressure wave. When the pressure wave arrives to the surface of a solid material it may cause vibrational movement in the material. This is an inverse phenomenon of when a sound is produced by a vibrating string; a sound may also make a string to vibrate. [4]

When a sound propagates in a medium, one can observe the same kind of phenomenon as in the case of a light wave, even if the light is an electromagnetic wave which doesn't need a medium for propagation. The boundary layer of materials with different densities affect the propagation of sound waves. Let the wave propagate first in perpendicular to the boundary. At the boundary, a part of the wave is transmitted over the boundary and the other part is reflected backwards. This

phenomenon is discussed later in the thesis. In the case of a non-perpendicular wave the transmitted wave may also be refracted. While propagating the energy of the wave can be also absorbed by the material or scattered by small inhomogeneities of the material. Even dispersion, diffraction and interference of the sound waves are possible. These phenomena are more familiar in the field of optics.[5]
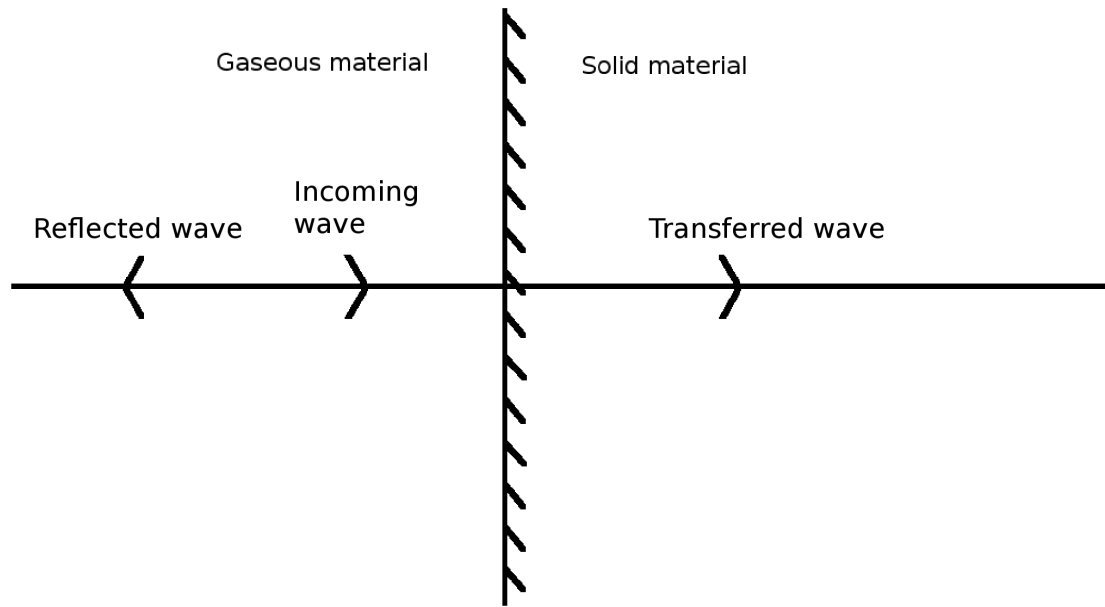
Figure 2.1: Boundary layer of gas and solid material.

The condition for highly efficient acoustic and mechanical wave transmission is a certain impedance relationship between two different materials. The acoustic or mechanical impedance of a material is defined as a product of material density and pressure wave (sound) propagation speed in this material, i.e.[6, 7]

$$Z = \rho C, \qquad (2.1)$$

where Z is the acoustical impedance, ρ is the density of the material and C is the speed of the propagating wave in the material. There is no actual difference between acoustic or mechanical impedance, because the pressure wave propagating in the gas is heard as a sound whereas the same pressure wave in solid or liquid material is observed as a mechanical vibration. Nevertheless, the pressure wave is the same. [6]

The acoustic impedance of water is [6]

$$Z_w \qquad = \rho_w C_w \qquad = 0.994 g/cm^3 \times 1.467 \times 10^5 cm/s$$
$$= 1.458 \times 10^5 \frac{g}{scm^2}.$$

For air the acoustic impedance is (sic!) [6]

$$Z_A \qquad = \rho_A C_w \qquad = 0.0012 g/cm^3 \times 1.467 \times 10^5 cm/s$$
$$= 1.458 \times 10^5 \frac{g}{scm^2}.$$

When the acoustic wave travels over the boundary of the layers of different materials, a part of the wave is reflected backwards (see fig 2.1). The amount of energy reflected with the wave is determined by the acoustic impedances of the materials. The phenomenon of the reflection arising from the impedance mismatch is commonly known in different fields of science e.g. in theory of electromagnetism.For water-air interface the transmission coefficient is [6]

$$\alpha_t = \frac{4 Z_A Z_W}{(Z_A + Z_w)^2}. \tag{2.2}$$

Similarly the reflection coefficient is [6], [7]

$$\alpha_r = \frac{(Z_W - Z_A)}{Z_W + Z_A}. \tag{2.3}$$

When the numerical values are substituted, the following is obtained $\alpha_t \approx 0.0011$ and $\alpha_r \approx 0.9989$ [6]. This means that over 99% of the total energy is reflected and only 0.1% is transmitted. It can be seen from equations 2.2 and 2.3 that as the acoustic impedances approach each other the transmission coefficient approaches one and the reflection coefficient approaches zero.

The density of the tissue can be estimated to be $\rho_T = 1.06 g/cm^3$ and the velocity of sound in a tissue $C_T = 1.54 \times 10^5 cm/s$. This gives the impedance value $Z_T = 1.632 \times 10^5 \frac{g}{scm^2}$, which leads the transmission coefficient to be α_t 0.00097 and the reflection coefficient to be α_t 0.99903 [6]. The situation is somewhat similar to the case of air-water interface.

There are many different sound sources in the human body. In clinical use the breathing and heart sounds are the most studied. The flow chart of an acoustic signal originating from a human body is presented in figure 2.2. In an ideal case all of the sounds would come from the body, but in reality the signal may be distorted by external noise.

In the 1960's and 1970's, the method of phonocardiography was studied, because the ultrasound imaging devices were not available for ordinary clinical use. The phonocardiogram (PCG) is a plot of recorded heart sounds which can be used when the functioning of the heart is studied. [8] The heart sounds give information related to the physiology of the heart. The same principles as in traditional stethoscopic

```
┌─────────────────────┐      ┌──────────────────────┐      ┌───────────────────────┐
│ Source of the       │      │ Transducer           │      │ Recording device      │
│ sound               │      │                      │      │                       │
│                     │ ──▷  │ Microphones,         │ ──▷  │ Amplifiers, filters,  │
│ The human body:     │      │ accelerometers, etc. │      │ A/D converters,       │
│ surface of the chest,│     │                      │      │ recording computer, etc.│
│ trachea etc,        │      │                      │      │                       │
└─────────────────────┘      └──────────────────────┘      └───────────────────────┘
                                        △
                             ┌──────────────────────┐
                             │ Noise                │
                             │                      │
                             │ External room noise, │
                             │ noise of measuring   │
                             │ instrument, etc.     │
                             │                      │
                             └──────────────────────┘
```
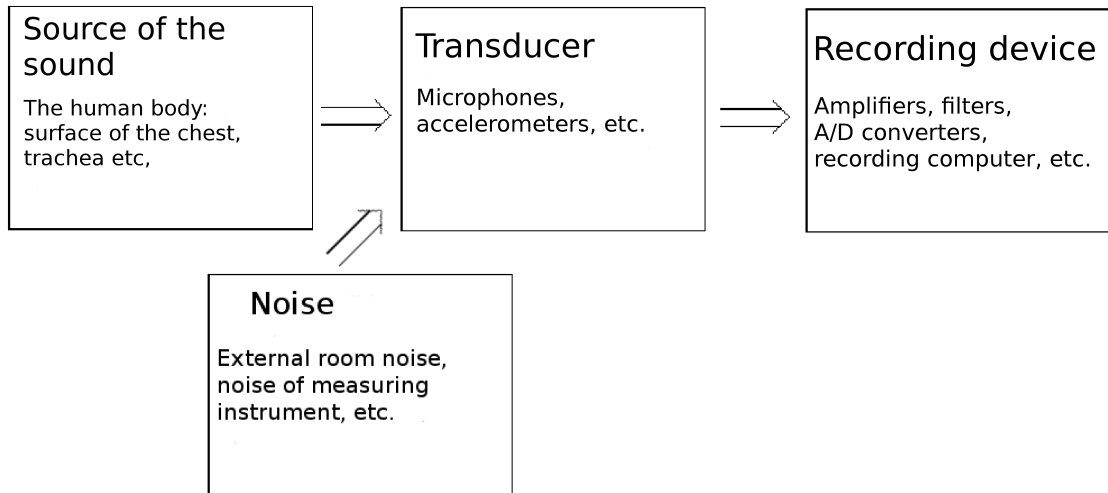
Figure 2.2: The flow-chart of the acoustic signal. The source of sound is the human body. The sound is detected with a transducer which converts the acoustic energy into electrical energy. Finally the signal is amplified, filtered, recorded and analysed.

auscultation are valid, but the methods of digital signal processing can be used. Usually an air-coupled electret microphone or an accelerometer have been used as a transducer[9, 10].

The novel method of recording sounds of blood circulation in the brain veins has been presented recently.[11] A sensor is placed over the closed eye lid and sounds from the brains propagate to the sensor through the eye. Aneurysms and stenosis can be discovered based on the changes in the spectrum of the sound. The resources of the imaging methods of the brain veins are limited among large populations, therefore the method of sound recording might become an alternative for screening a larger population.

The constriction of muscle cells produces sound. The method of recording these sounds is called acoustic myography (AMG) or mechanomyography (MMG) [12, 13]. This method can be combined with electro myography (EMG) in order to provide more information on voluntary contraction of patients[12]. The clinical use of AMG is not widely adopted.

An interesting application of sound recording is recording ordinary speech that conducts through the human body. Modern technology has enabled the human voice to control different kinds of devices. A problem with this technology is that environmental noise may interfere with the voice control. Another problem is that like in the case of passwords, the environment should not be able to hear the produced sounds. Human beings can produce whispers or voices called murmurs. Non acoustic murmurs are produced when vocal cords do not vibrate. Researchers have discovered that this non acoustic murmur can be detected from the surface of the

human being with special non-acoustic murmur microphones (NAM) .[14, 15, 16] The best location to record these sounds was determined experimentally and this location was discovered to be behind the speaker's ear [14].

In order to detect body sounds, the pressure waves travelling in tissue must be transmitted to the transducer. Several types of transducers have been used to record body sounds and many of these have been used to detect breathing sounds as well.

Skin vibration can be presented with three physical quantities which can be derived from each other. These are the displacement, the velocity and the acceleration of the skin. Let dx be the displacement of the skin and dt time. Then the velocity of the skin is [4]

$$v = \frac{dx}{dt}.$$ (2.4)

Similarly the acceleration of the skin is

$$a = \frac{dv}{dt} = \frac{1}{dt}\left(\frac{dx}{dt}\right) = \frac{d^2x}{dt^2}.$$ (2.5)

Therefore, the unit of the measured signal should be identified if commercial measuring systems are being used.

The most commonly used sound sensors are electret microphones with acoustic couplers and piezoelectric accelerometers. In figures 2.3 and figure 2.4, the schematic figures of sensors are presented.
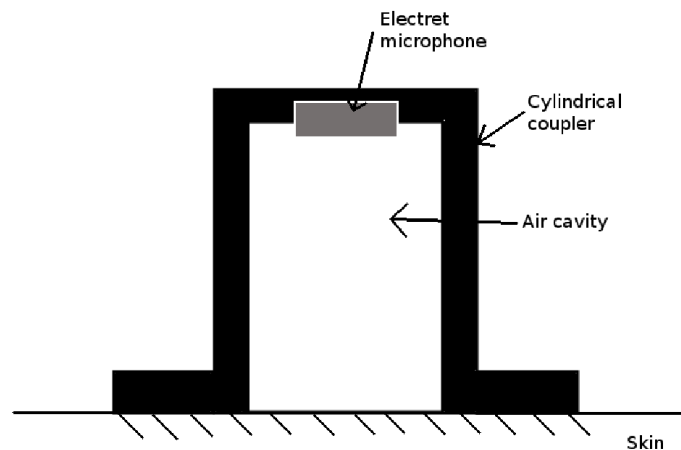
Figure 2.3: Air coupled electret microphone.

One basic problem with the sensors is that microphones are designed to be used in a free field (without mounting them into the cavity) and accelerometers are designed to be used at rigid surfaces. Both of these assumptions are violated when these sensor are used as body sound sensors.

Air coupled microphones have been used at least in PCG, AMG and breath sound detection. Watakabe[13] has shown that an air coupled microphone only
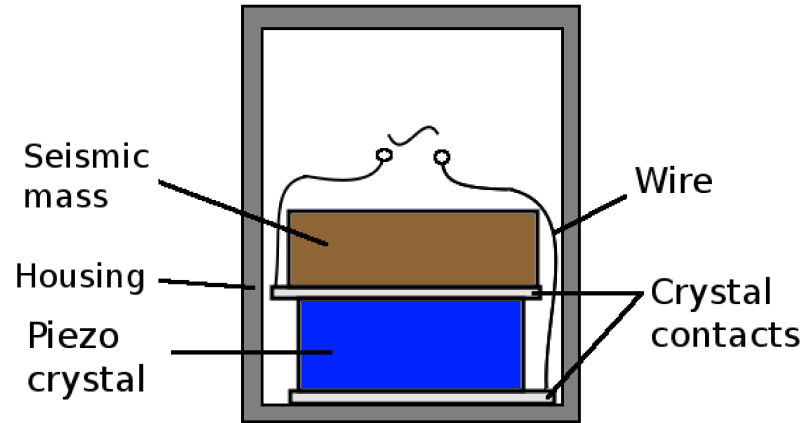
Figure 2.4: Piezoelectric accelerometer. When the housing is accelerated the inertia of the seismic mass produces force on the crystal and a voltage difference is produced into the wires.

measures the displacement of the skin, not the acceleration or the velocity [13]. The impedance of the air coupled microphone is discussed later in this thesis. Air coupled microphones must be acoustically coupled to the skin. Otherwise impedance mismatch occurs, as previously noted. The coupler is an enclosed air chamber, of which one end is sealed against the skin and the microphone is mounted to the other end [17, 2, 3]. Generally two types of chambers are used; conical and cylindrical ones (see figures 2.10 and 2.3).The chamber acts as an acoustic impedance matching device; high impedance and low vibration of the skin are changed to relatively low impedance and strong vibration of the diaphragm [17, 3].

A conical shaped acoustic coupler amplifies the sound coming to the transducer. An explanation for this can be given by using the principle of energy conservation at each end of the coupler. Let us assume that a conical shaped coupler has two ends with different cross-sectional areas called A_1 and A_2, see figure 2.5. When the
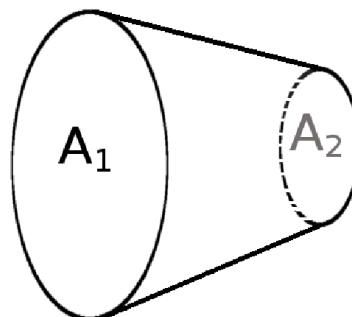
Figure 2.5: Gain effect of conical coupler due to difference in cross sectional areas.

pressure wave arrives to the coupler it produces a force $F = PA_1$ on the surface A_1.

If the displacement of the pressure wave is x_1, the total energy can be expressed as $E = Fx = PA_1x_1$. The total energy must be conserved and therefore it must be the same at both ends of the coupler [6]. This can be formulated as [6]

$$P_1A_1x_1 = P_2A_2x_2. \tag{2.6}$$

When approximated that the air inside a coupler is incompressible, it can be said that that $P_1 = P_2$ and the equation 2.6 gives [6]

$$g = \frac{x_2}{x_1} = \frac{A_1}{A_2} = \left(\frac{r_1}{r_2}\right)^2, \tag{2.7}$$

where g is the displacement gain of the coupler and r is the radius of the cross-sectional area. For example if the gain is 4 it means that when the front face moves one unit, the displacement of the rear face is four units. [6]

Accelerometers have been used in breath sound detection, in PCG and in blood flow noise detection. The transducer which measures the blood flow noise of the brain through the eye, uses a tent shaped piezoelectric film. In the case of accelerometers the coupler is not necessarily needed because the sensing piezoelectric rod is in an enclosure which is attached straight to the vibrating skin [2]. NAM-microphones do not use traditional piezo-crystal. Instead they have a plastic coupler similar to air coupled electret microphones. The cavity is filled with solid material, usually with silicone[15]. Therefore NAM-microphones may be categorised as a kind of accelerometer. Researchers have found out that by replacing the silicone with urethane elastomer, the sensitivity of the microphone increases[18]. This implicates that the cavity has an important role in the sensitivity issues of NAM microphones.

There are two types of contact microphones i.e accelerometers. The first type is a heavy mass microphone and the other one is a light-weight type microphone. A problem with the accelerometers is that the sensor sensitivity is a strong function of its mass [9]. In addition, some models may need special signal conditioning and power supplies in order to function properly. [2]

The heavy weight accelerometer has steady housing which contains a piezoelectric crystal and a partially steel rod with a knob. One end of the rod is attached to the piezo crystal and the outcoming knob is attached to the skin. When the skin moves, only the rod moves with it while the housing remains at rest. In the case of the light weight microphone, the whole housing is in motion and the piezo crystal is attached to the wall of the housing. [9] The contact microphones used in tracheal sound analyses are accelerometers i.e. the light weighted microphones. The oscillating mechanical circuit representing the accelerometer is presented in figure 2.6, where the symbols are the same as in the equation 2.8. The impedance of the
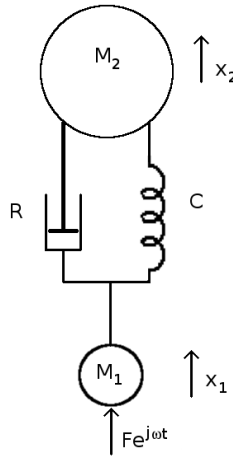
Figure 2.6: Mechanical circuit of the accelerometer

accelerometers can be expressed as [9]

$$Z_m = j\omega M_1 + j\omega M_2 \frac{1 + j\omega RC}{1 - \omega^2 M_2 C + j\omega RC},$$ (2.8)

where R is the mechanical resistance of the crystal, C is the mechanical compliance of the crystal, M_1 is the mass of the pickup in contact with the chest wall, M_2 is the seismic mass, $\omega = 2\pi f$ is the angular frequency and $j^2 = -1$ is the imaginary unit. F is the force exerted on M_1, x is the displacement of the accelerometer and e is the exponential constant. In the case of the light-weight microphone, M_1 is the mass of the housing.

The mechanical impedance of the contact microphones is a complex quantity determined by mechanical resistance and mechanical impedance. Experimental results of the mechanical impedance of the light-weight microphone is presented by Van Vollenhoven [9] It was observed that the impedance of the accelerometer decreases in the range of 100-1000Hz. Van Vollenhoven presents that mechanical resistance and reactance are in a slow increase in the range of 400Hz-1000Hz. These figures show that the impedance mismatch between the tissue and microphone impedances increases in the range of 400-1000Hz. An issue affecting the impedance characteristics is the mass loading effect. It means that when the contact force of the vibration transducer is increased, the vibration characteristics of the tissue changes. The same effect is observed when the mass of the sensor resting on skin is increased, because the increasing mass increases the contact force.[9]

The mechanical impedance of the body has been determined by Franke (1951) [19], Takagi et al. (1964)[20] Van Vollenhoven [9] and Katz (2000) [21]. Franke and Takagi measured the impedance of the arm, Van Vollenhoven the impedance of the chest and Katz measured from all previously mentioned places and in addition from

areas of the forehead, the cheek and the stomach.

Van Vollenhoven [10, 9] has studied the impedance of the human body from the point of view of the phonocardiogram. The mechanical impedance of the chest wall is difficult to determine, because there is intersubject variability and the measurement results depend on the location of the sensor on the chest wall, the contact area between the sensor and the skin and the static pressure applied to the sensor. Van Vollenhoven used the method of Franke where a moving piston was used. Skin moves in the same way as a piston when a contact sensor is applied on it. According to the author this assumption is valid when holding breath and the mechanical impedance in the range of 50-1000Hz should be determined during a short period of time (7s). [9]

Van Vollenhofen found out in his experiments that four factors affect the determination of mechanical impedance of the human chest. The first one was intersubject variability. The second one was the differences among the locations of measuring places i.e. intercostal spaces. It was discovered that above the bones the resonance frequencies (the frequency where reactance is zero) increased. A third observation was that the contact area of the microphone strongly affects the measured impedance. The fourth observation was that the static force with which the vibrating piston is applied to the skin affects the measured impedance. Usually this force is equal to the weight of the microphone when it is resting on the chest. The mechanical resistance and the mechanical reactance increase as the contact force increases. This happens because the applied contact force increases the stiffness of the chest wall. As a result the resonance frequency increases. [9] The conclusions of Van Vollenhoven are similar to the ones of Franke[19]. In contrast, Katz [21] found out that the human body is acoustically rigid in the range of 1kHz-6kHz, i.e. the impedance is relatively flat and the reflection coefficient is high in the frequency range studied. Katz explained that the difference with the experiment of Franke is that Franke used a mechanical vibrator as the source of vibration. Katz presented that because the vibrator (piston) was physically connected to the skin it affected the results, whereas the author used sound pressure to excite the skin. The connected piston would produce curvatures and tensions to the skin which would have affected the measurements. [21] Van Vollenhoven [9] observed that the microphones were loading the chest wall and in addition he discovered that the three different microphone models had totally different impedances. Two microphones had a smaller impedance than the chest wall. When the application force was increased, the relative amplitude level of these microphones decreased. One microphone had a larger impedance than the chest wall. When the force of application was increased for this microphone, the relative amplitude increased. The impedance of the contact microphone changes when it is applied to the skin, but simultaneously the mechan-

ical impedance of the chest wall changes as well.[9] In his publication Vollenhoven pointed out that the frequency response of the microphones differ mostly at low frequencies and that higher frequencies look rather similar. The author also points out that the characteristics of the frequency response at a high frequency range are defined by the low pass filters of the system rather than by the coupling effects. Therefore impedance measurements may have described the filter effects rather than the real impedance at high frequencies. In his research the author critisizes the manufacturers of phonocardiographs by saying: "Although it does not seem too difficult to account for the frequency response of the microphone by adapting the filter response to the type of microphone used this is not done by the manufacturer. The microphone is just plugged into the phonocardiograph and no correction is made with respect to the difference in frequency response of the microphones". Finally, the author states that the calibration of microphones is significant for a frequency analysis of heart sounds and murmurs. [9]

In 1970 Van Vollenhoven [10] studied the calibration of air coupled microphones in phonocardiography. Different methods were investigated and the calibration was done in the range of 50-1000Hz. It was observed that the best results were achieved with a pistonphone. The author concluded that the method is simple and accurate, but that it requires a quite expensive experimental setup. However, the pistonphone did not produce a good sinusoidal output and therefore it is not suited for determining frequency responses. It was observed that by using the earphone, similar results were achieved as with the pistonphone in the range of 50-600Hz. For higher frequencies the curves did not agree. It was also observed that the air volume in the coupler should be as small as possible. The same kind of efficiency of small cavities was also found by Wodicka et al [17]. Details of this study is reviewed later. Vollenhoven also noted that the earphone produced better results than an electrodynamic microphone. In their research the authors presented that there are two reasons why the mechanical impedance of the chest wall is different when measured with an air-coupled microphone compared to contact microphones. Firstly, the air coupled microphone touches the skin only with its rim. Therefore, the central part of the skin is loaded only by the mass of the air column enclosed to the chamber of the coupler. Secondly, the pressure inside the coupler varies between positive and negative, whereas the contact microphone exerts a constant positive pressure to the skin. [9]

Vollenhoven says that for calibrating the PCG microphone the concept of mechanical impedance is essential. The mechanical impedance is a complex ratio of the applied force to the resultant velocity. For calibration purposes the mechanical impedance of both the microphone and the chest wall should be determined. [9]

Schwartz et al. [6] have studied the effects of acoustic impedance matching in

phonocardiogram systems. The authors presented improvements to phonocardiogram systems based on acoustic matching and gain effect of conical couplers. Furthermore, the acoustic coupler was filled with an aqueous coupling medium. As a result, the amplitude increased by 100. The transducer type was chosen to be PZT piezoelectric crystal and the cantilever type configuration was chosen. A conical coupler was used and the displacement gain was adjusted to 16. After that the crystal material and vibrational configuration were chosen. The original resonant frequency of the crystal was [6]

$$\omega = 1/\sqrt{mc_m} = 42\text{kHz}, \tag{2.9}$$

where ω is the angular frequency, m is the mass of the crystal and c is the inverse force constant of the crystal. The crystal was connected to a preamplifier and the conical cavity was filled with water. Both of these implementations caused the resonant frequency to decrease to 92Hz. In order to fill the conical cavity with water, a thin Mylar film was mounted over the front end. In the rear end, a silicone rubber membrane was used to seal the cavity. One end of the crystal was fixed, but the other end was positioned as a cantilever. It was explained in the article that the shape of the horn (linear, exponential, catenoidal) is not important as long as the dimensions of the horn are less than 0.1 times the shortest wavelength recorded. This enables plane wave approximations. The mechanical impedance of the measuring system is [6]

$$|Z_M| = \sqrt{R_m^2 + (\omega m' - 1/c_m\omega)^2}, \tag{2.10}$$

where R_m is the energy dissipating term, ω is the angular frequency, m' is the effective mass and c_m is the inverse force constant of the crystal. The R_m is mainly determined by the input resistance of the amplifier unit. This resistance was chosen to be 100kΩ which gives the impedance the value $|Z_m| \approx 4.36 \times 10^6$g/s. The specific acoustic impedance of the water column which has the equal size of the microphone cross-sectional area was calculated to be [6] 4.58×10^5g/s. This gives a theoretical mismatch between the water column and the measurement system to be ≈ 9.5 [6]. The measured value of the mismatch was 6.67 [6]. The quality factor is defined as a ratio of total energy to energy dissipated per one cycle, i.e. $Q = 2\pi \times \frac{\text{Total energy}}{\text{Energy loss}}$. The Q of the microphone is defined as $Q = \omega_0 m'/R_m = 4.32 \times 10^{-4}$ [6], which means that the recording system was heavily damped. If the input resistance of the microphone amplifier is increased (affects to R_m), the frequency dependency of the mechanical impedance Z_m increases [6]. Reduction in the input resistance decreases the frequency dependency, but causes an increase of impedance mismatch, which in turn leads to a reduced signal amplitude [6].

2.2   Respiratory sound transducers

In this section, the properties of respiratory sensors will be discussed in the light of various articles and by comparing the information they contain. The articles are introduced mainly in chronological order as to provide a general view of how present knowledge has evolved. Many differences between research results concerning sensor types exist, because sensor knowledge has improved during time. Articles written by small research groups are presented from time to time in this review of respiratory sound transducers. This is because there are only a few articles in which pure technical aspects are favoured in order to obtain as accurate characteristics of breath sound sensors as possible. Unfortunately, these technical articles are written by few and the amount of published articles is limited. Therefore, the scientific reliability is not as good as it should be. In some articles only *in situ* measurements are done, but methods of mathematical simulation, *in vitro* and *in situ* measurement have all been used. Many other researchers have also reported the properties of sensor use, but these have been based on the findings on breath sounds, not on the comparison of sensors.The difference is that without a technical comparison the phenomena that are thought to be caused by the sensor may be actually be effects of breath sounds, measuring setups, or noise from the environment. Unfortunately even today there are many researchers who study breath sounds and are not familiar with these technical studies or they have not referred to these kinds of articles in the least. detector comparisons should also be done by larger groups of researchers in order to verify or cancellate the results obtained by others.Research groups would benefit from this themselves, because an increase in sensor knowledge would direct them to use sensors which are better for their applications and as a consequence this would accelerate breathing sound sensor development worldwide. In this context it should be pointed out that the optimal breathing sound sensor has not yet been invented.[22]

The ability to detect respiratory sounds electronically was developed in 1950's by McKusick and it was adopted by others in the 1960s'and 1970s'. The aim of the research was to obtain more information on physiological or pathological processes. [3] Some electrical stethoscopes are commercialized, but they are mainly designed for ordinary clinical use [3]. The advances in microprocessor technology have made respiratory sound analysis possible in clinical laboratories worldwide. The choices of sensors, measurement setup, sound signal sampling and processing and the presentation of the results varies around the world.[22] The different performances of sensors is a problem when a sound spectrum or the shape of waveforms are analyzed. Different practises of the sensors attached to the skin affect also the spectrum and waveforms, which complicates the analysis even more.[3] Usually breathing sound

sensors have been attached to the thorax or on the neck area, where the tracheal site has been most popular.

The sound recording properties of sensors vary greatly. The general idea is that contact sensors, for example piezoelectric vibration sensors, perform better than the much cheaper air coupled electret microphones. [3] A comparison

of sensors has been made by Pasterkamp et. al 1993[22] and Kraman et. al 2006 [3]. In these articles, contact and air coupled electret microphones were chosen to be compared, because they have been the most common sensor types used in breath sound research.

2.2.1   Sensor comparison

Three air-conductive and four contact microphones were studied by Pasterkamp et al. in 1993 [22]. Detailed information of the sensors is shown in figure 2.1. The

Table 2.1: Tested sensors [22]

Name	Manufacturer	Weight (g)	Diameter (mm)
Air coupled			
Sony ECM 155	Sony Corp.	1.7	5.6
Sony ECM 77	Sony Corp.	1.5	5.6
Radio Shack No. 33-1052	Tandy Corp.	2.0	7.6
Contact			
HP 21050	Hewlett-Packard	52.2	14.0
Siemens EMT25C	Siemens	15.4	28.0
PPG No. 201	Technion University	9.9	28.0
FYSPack2	University of Brussels	2.1	20.0

weight of the contact sensors varied more[1] than with air coupled sensors and they were very different in size as well. The subjects studied sat in an acoustic chamber and breathed through a pneumotachograph. The target air-flow was adjusted to $2 \pm 0.5 \mathrm{l/s}$. The sensor was attached to the posterior lower chest to the place where the maximum sound was heard when auscultated. Only inspiratory sounds were recorded. The authors stated that measurements from the tracheal site would have provided a greater signal-to-noise ratio (SNR) and signal bandwidth. Lung sounds were amplified and pass-band filtered between 100 and 2000 kHz. Low frequencies (< 100Hz) were rejected, because strong spectral components of heart and muscle sounds dominate in a low frequency range. This is a typical situation and a comparison of breath sound sensors in the low frequency range is impossible with *in vivo* measurements [13]. The authors had calculated the average spectra by using

---

[1]As was previously noted the sensitivity of the accelerometers is a strong function of it's mass.

2048 point FFT. The Hanning window was chosen and only those sound samples that occured with flows between 2 ± 0.5l/s were used. The background noise spectra was calculated by using samples at flow rates of 0.0 to 0.1 l/s during the late expiration. The bandwidth was defined from the points where the signal level was less than 3dB above the background noise level. SNR was calculated in decibels(dB) as $10 \times \log$(signal power/noise power). Frequency F_{Hi} was the highest frequency where the signal amplitude was just above the background noise level. Attenuation as a dB/octave in the range of 300-700Hz was achieved from a log-log plot of the spectrum. [22] The results of the sensors from three subjects is presented in table 2.2.

Table 2.2: Performance of the sensors.[22]

Sensor name	S/N			Slope			F_{Hi} (Hz)		
	1#	2#	3#	1#	2#	3#	1#	2#	3#
Air coupled									
Sony EMC155	41.4	29.8	24.6	-25.2	-22.3	-21.1	1250	930	1105
Sony EMC77	42.2	28.3	23.1	-26.7	-25.4	-23.8	1010	785	850
Radio Shack	38.6	25.3	23.3	-16.9	-19.1	-17.6	1370	655	925
Contact									
HP21050	42.0	24.5	28.9	-15.9	-18.1	-14.2	1010	1280	1880
Siemens EMT25C	33.0	16.4	22.3	-16.1	-14.3	-15.3	1560	560	920
PPG #201	41.1	25.0	26.2	-16.0	-17.5	-15.3	> 2000	955	> 2000
FYSPac2	40.3	26.0	25.5	-11.4	-12.4	-13.7	> 2000	765	935

The calculated averages of the table 2.2 are presented in table 2.3.

Table 2.3: Performance averages of tested sensors [22]

Sensor	Average		
Air coupled	S/N	Slope	F_{Hi}
Sony EMC155	31.9	-22.86	1095
Sony EMC77	31.2	-25.3	881
Radio Shack no 33-1052	29.1	-17.9	983.3
Contact			
HP 21050	31.8	-16.1	1560
Siemens EMT25C	23.9	-15.2	1013.3
PPG no 201	30.8	-16.3	1651.7
FYSPac2	30.6	-12.5	1233.3

It can be seen from table 2.3 that there was no great difference in SNR between contact and air coupled sensors. Figure 2.7 shows a frequency response (FR) of five used sensors. Air coupled microphones have a steeper curve in the range of 300-700Hz than contact microphones. Therefore, contact microphones have a broader
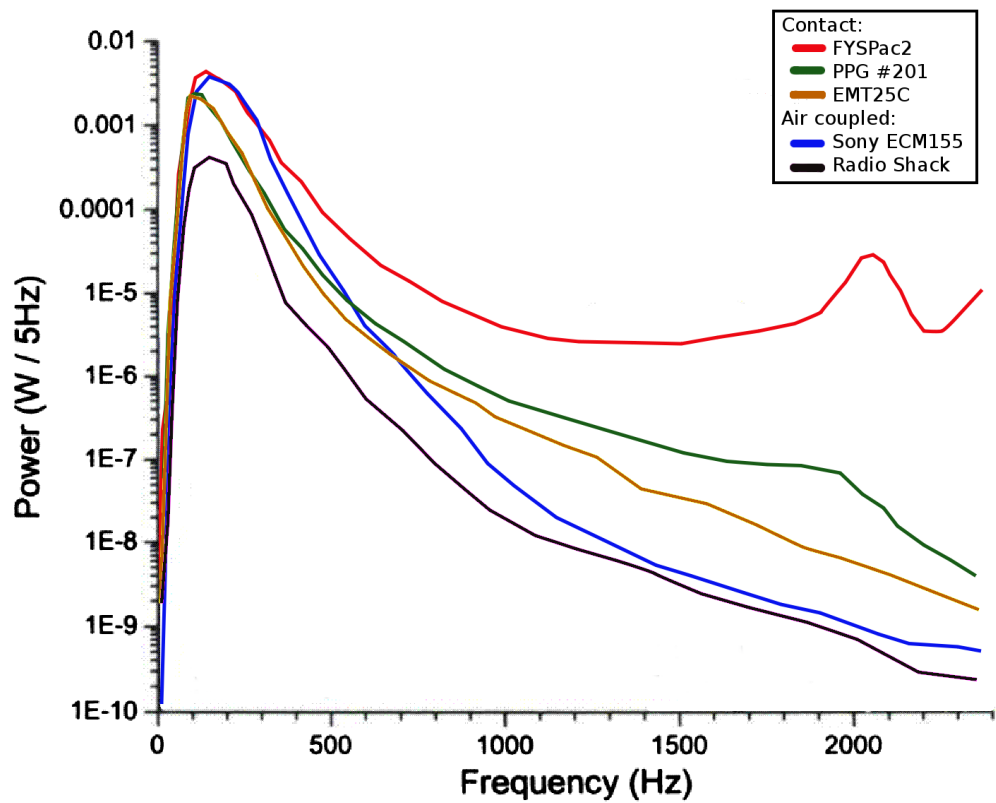
Figure 2.7: Frequency response of the sensors. The background noise spectra has been subtracted from the lung sound signal spectra. Modified from [22]

bandwidth and they can be used to record sounds with spectral components over 2000Hz. In fact, the authors found spectral components over 2kHz with PPG contact sensors at high flow rates from two studied subjects. However, the bandpass filter attenuated the signal very efficiently and an analysis of high frequency components was impossible to make. According to the authors, a problem with the contact microphones is the mechanical resonance where the sensor starts to oscillate, as can be seen with the FYSPack2 sensor at 2000Hz in figure 2.7. Other two issues affecting the signal are cable noise and the attachment of the contact microphone. [22]

In figure 2.8 a major difference between two contact microphones within the same model exists. The authors explain that this is due to the characteristics of the
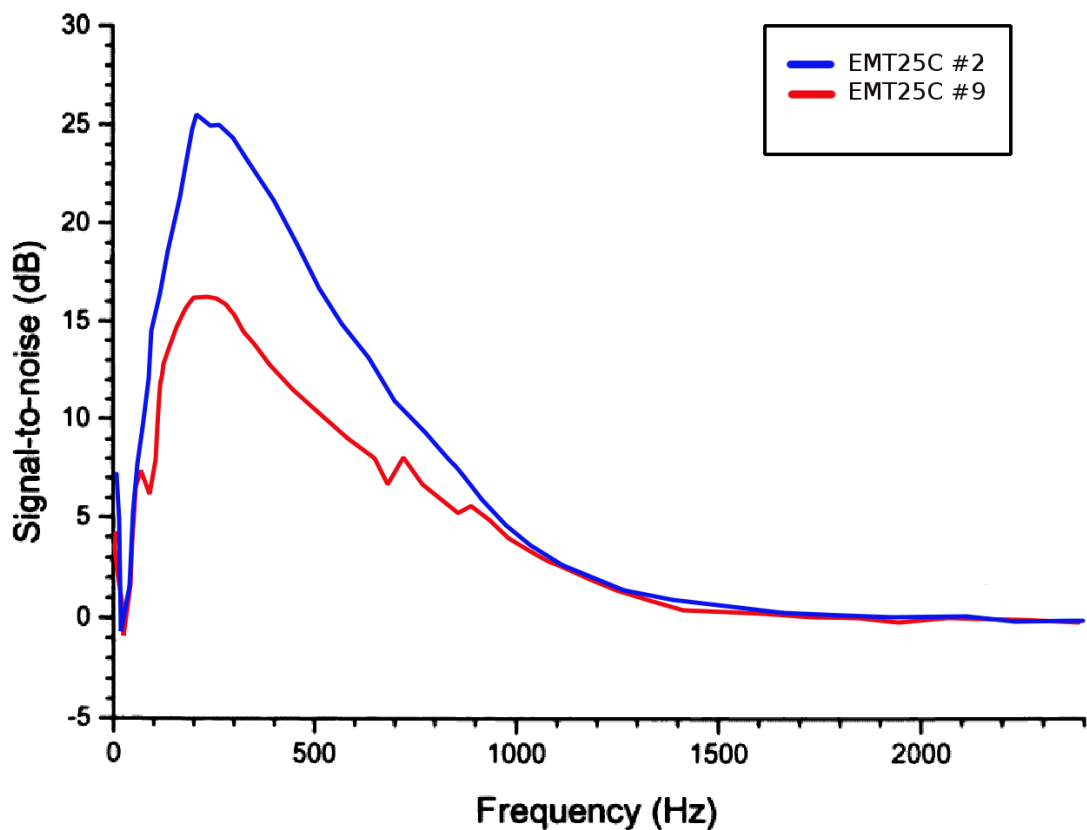
Figure 2.8: Difference between two similar contact microphones. Modified from [22]

sensor or to damage during transportation. This implicates that there is a need for a quality test before any measurements are made. Pasterkamp et. al suggest that the weaker FR of air coupled microphones was due to damping of the high frequencies of the sound in the coupler[2]. They do agree that a better design of the cavity would improve the results of air coupled microphones. The manufacturer's specification of the electret microphones used stated that the FR extends beyond 10kHz in free field. Druzgalski et al.[23] put the microphone in the cavity and found no effect on

---

[2]See the effect of cavity depth on page 18

the FR below 2500Hz in free field. [22] Pasterkamp et al. used vented cavities for air coupled sensors. The present knowledge implies that this might have reduced the sensitivity at high frequencies[3].

The authors state that "Until the ideal sensor for respiratory acoustical measurements has been found, different types and models are going to be used." They strongly believe that quality control has to be done and that standardization is becoming increasingly important. [22]

2.2.2 Effects of microphone cavity depth

In 1994 Wodicka et al. studied the effects of air coupled microphone cavity depth [17]. They were investigated in three ways: 1) An acoustical model of the chest wall-air cavity-microphone interface was developed for understanding the importance and action of the cavity. 2) *In vitro* measurements with artificial chest walls were performed with different cavity depths and 3) *in situ* measurements were done with different cavity depths in order to quantify the effects of the moving chest. The study subject was one of the investigators of the in situ measurements. It was assumed that the highest breathing sound frequency was 1500 Hz and that the velocity of the sound was the same as in the air, 35400cm/s. It can be calculated that [17]

$$\lambda = \frac{v}{f} = \frac{35400\text{cm/s}}{1500\text{Hz}} \approx 23.6\text{cm}, \tag{2.11}$$

where λ is the wavelength of the sound, v is the velocity of the sound and f is the frequency of the sound. In tissue wavelength it is even longer due to the increased velocity of the sound. The wavelength is longer than the thickness of the chest wall, the depth of the acoustic coupler or the thickness of the microphone diaphragm. [17] Therefore, resonances in these places are impossible because there is no place for standing waves.

The lumped acoustic element model assumes that the acoustic system can be modelled with infinitely small discrete elements which are located in certain positions relative to each other. The acoustic properties of the chest wall, the cavity and the microphone are modelled with the help of an acoustic circuit, see figure 2.9. In acoustics, a certain analogue with electricity can be found. Sound pressure p (N/m^2) is analogous to electrical voltage and volume velocity U (m^3/s) is analogous to the electrical current. The acoustical impedance can be stated as[17]:

$$Z = \frac{p}{U}. \tag{2.12}$$

Electric resistance is analogous to acoustic resistance, electric capacitance is analo-

---

[3]See page 32

gous to acoustic compliance, electric inductance is analogous to acoustic inertance, etc. Therefore an acoustic circuit can be treated with the same means as an electric RLC circuit. The acoustic circuit model of the measuring setup is presented in figure 2.9[17]. With a cavity diameter of 8mm the following estimates for the chest wall
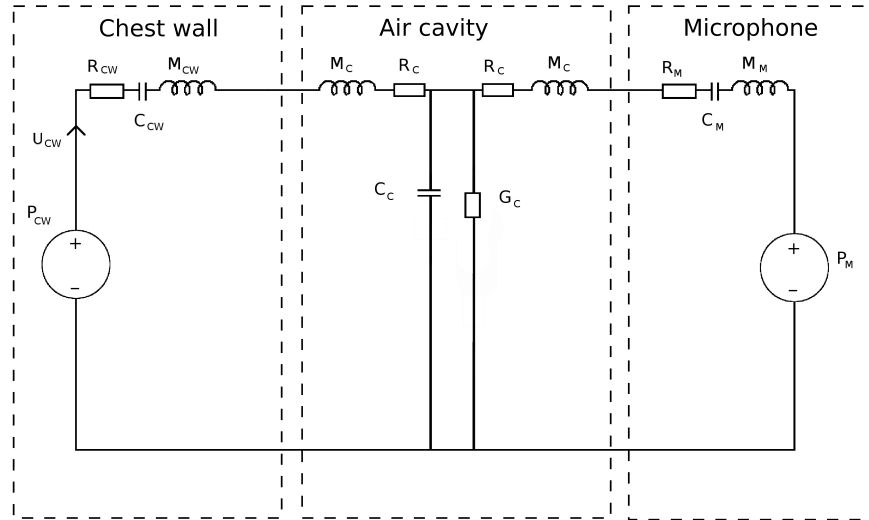
Figure 2.9: Acoustic circuit

are the following: [17]

$$
\begin{array}{rl}
\text{acoustic resistance} & R_{\mathrm{CW}} = 8 \times 10^{8}\,\tfrac{\mathrm{Ns}}{\mathrm{m}^{5}} \\
\text{acoustic compliance} & C_{\mathrm{CW}} = 2.5 \times 10^{-12}\,\tfrac{\mathrm{m}^{5}}{\mathrm{N}} \\
\text{acoustic inertance} & M_{\mathrm{CW}} = 2 \times 10^{5}\,\tfrac{\mathrm{kg}}{\mathrm{m}^{4}}
\end{array}
$$

With these values the series of resonance for the chest wall is

$$
f = \frac{1}{2\pi\sqrt{C_{\mathrm{CW}}M_{\mathrm{CW}}}} = 225\mathrm{Hz}. \tag{2.13}
$$

Below the resonance frequency, the compliance dominates Z_{CW} and above, the inertance is dominating.

The impedance values of the microphones (Z_M) for Sony EMC-155 are the following according to the manufacturer's specifications:

$$
\begin{array}{rl}
\text{acoustic resistance} & R_{\mathrm{M}} = 1 \times 10^{8}\,\tfrac{\mathrm{Ns}}{\mathrm{m}^{5}}, \\
\text{acoustic compliance} & C_{\mathrm{M}} = 3.6 \times 10^{-14}\,\tfrac{\mathrm{m}^{5}}{\mathrm{N}}, \\
\text{acoustic inertance} & M_{\mathrm{M}} = 3.3 \times 10^{2}\,\tfrac{\mathrm{kg}}{\mathrm{m}^{4}},
\end{array}
$$

The authors report that these values were quite similar to other microphones from different manufacturers. [17] The series resonance for the microphone is

$$f = \frac{1}{2\pi\sqrt{C_M M_M}} = 46176\text{Hz},\tag{2.14}$$

which is a much higher value than the one for the resonance of the cavity wall.

The component under interest in the acoustic circuit is the microphone cavity. It consist of the following components:[17]

$$\text{acoustic resistance}\quad R_C = \frac{S}{2A^2}(\pi f \rho \mu)^{1/2} l = 1.3 \times 10^6 l \tfrac{\text{Ns}}{\text{m}^5},\tag{2.15}$$

$$\text{acoustic compliance}\quad C_C = \frac{Al}{\rho v^2} = 3.5 \times 10^{-10} l \tfrac{\text{m}^5}{\text{N}},\tag{2.16}$$

$$\text{acoustic inertance}\quad M_C = \frac{\rho l}{2A} = 11.4 \times 10^3 l \tfrac{\text{kg}}{\text{m}^4},\tag{2.17}$$

$$\text{acoustic conductance}\quad G_C = \frac{0.4S}{\rho v^2}\left(\frac{\pi f \gamma}{\rho \alpha}\right)^{1/2} = 1.77 \times 10^{-8} l \tfrac{\text{m}}{\text{Ns}},\tag{2.18}$$

$$\tag{2.19}$$

where
A = cross sectional area of the cavity = $\frac{\pi d^2}{4} = 5 \times 10^{-5}\text{m}^2$,
ρ = density of the air = 1.14kg/m^3,
S = circumference of the cavity = $\pi d = 2.5 \times 10^{-2}\text{m}$,
f = frequency of the sound,
μ = viscosity coefficient = $1.86 \times 10^{-5}\text{Ns/m}^2$,
γ = coefficient of heat conduction = $0.55 \times 10^{-1}\text{cal/cm}°\text{C}$ and
α = specific heat of air at constant pressure = $0.24 \times 10^{-3}\text{cal/kg}°\text{C}$.

The authors note that when the cavity depth is $l = 2.5\text{mm}$ the compliance of the cavity is $C_C = 8.75 \times 10^{-13}\tfrac{\text{m}^5}{\text{N}}$, which is roughly half of the compliance of the chest wall. Similarly, the inertance of the cavity is $M_C = 2.85 \times 10^1 \tfrac{\text{kg}}{\text{m}^4}$, which is smaller than the inertance of the microphone. The authors explain that the compliance of the small air cavity is close to the compliance of the chest wall while the inertance is very small when compared to the chest wall and microphone. The authors state that the most important effect that the air cavity has on the frequency response of the system is that it matches the acoustic compliance that converts the tissue vibrations to sound pressure, which can be detected with a microphone diaphragm. The resonance of the cavity can be calculated by using the equation 2.13. When the cavity depth is $l = 2.5\text{mm}$

$$f = \frac{1}{2\pi\sqrt{C_{\text{CW}} M_{\text{CW}}}} = 31871\text{Hz}.\tag{2.20}$$

The cavity resonance can be modelled with the help of the Helmholz resonator[5, 7,

13]. The Helmholz resonator is a cavity where the air column resonates. This is a familiar sound when one blows air across the top of an empty bottle. The frequency of the sound changes if liquid is added to the bottle or different sizes of bottles are being used. When the air is forced into the cavity the pressure inside increases until it causes a force that cancels out the external force which originally caused the flow direction inwards. When these two forces are equal the flow is zero. Due to the inertia of the air, the pressure change overcompensates and the air from the cavity flows back to the external space until the pressure inside the cavity has reached its original value. The pressure overcompensation happens also to this direction and as a result we have air flow resonators. The condition for the existence of the resonator is that a constant force is applied from the outside to the cavity.In the following the conical cavity resonator is investigated. The volume of the cavity is V, the cross-sectional area A and height L. It follows that the volume $V = A \times L$. If the force F is applied to a column of air with a cross sectional area A the pressure change dP has a relation

$$F = AdP = A\frac{dP}{dV}dV = AV\frac{dP}{dV}\frac{Ax}{V} \tag{2.21}$$

where V is the volume of the air and x is the vertical offset of the air column. Because the vibrations are rapid and small, the pressure changes will not produce heat and the process can be treated as adiabatic [13]. For an ideal gas, the adiabatic bulk modulus is

$$K_s = -V\frac{dP}{dV} = \gamma P, \tag{2.22}$$

where γ is the heat capacity ratio and P is the undisturbed pressure. By substituting equation 2.22 in equation 2.21 it gives

$$F = ma = -\frac{A^2\gamma P}{V}x = \rho AL\frac{d^2x}{dt^2}, \tag{2.23}$$

where a is the acceleration, ρ is the density of gas and L is the height of the air column. Rearranging terms gives a differential equation:

$$\frac{d^2x}{dt^2} + \frac{A\gamma P}{V\rho L}x = 0. \tag{2.24}$$

The frequency of the oscillator is

$$f = \frac{1}{2\pi}\sqrt{\frac{A\gamma P}{V\rho L}} = \frac{v}{2\pi}\sqrt{\frac{A}{VL}} = \frac{v}{2\pi}\sqrt{\frac{1}{L^2}}, \tag{2.25}$$

where $v = \sqrt{\gamma\frac{P}{\rho}}$ is the velocity of sound in the medium. When the height of the electret microphone cavity (2.5mm) is substituted in equation 2.25 the following is

obtained

$$f = \frac{v}{2\pi}\sqrt{\frac{1}{L^2}} = \frac{340m/s}{2\pi}\sqrt{\frac{1}{(2.5 \times 10^{-3}\text{m})^2}} \approx 21645Hz. \tag{2.26}$$

This is much less than than the equation 2.20 predicted. Moreover, this kind of resonance is not observed in a breath sound spectrum. It seems that the Helmholz resonator is not a valid model because the cavity is closed and the open surface does not exist.

In article [13] an equation is derived for the pressure change as a function change of air column height. It has been stated that for adiabatic gas [13]

$$PV^\gamma = \text{constant}, \tag{2.27}$$

where P and V are the pressure and volume of the gas, respectively and γ is a heat capacity ratio. By differentiating the equation 2.27 we get [13]

$$V^\gamma dP - \gamma PV^{\gamma-1}dV = 0. \tag{2.28}$$

The equation 2.27 can be applied to the cavity of the air coupler. Figure 2.10 presents a cylindrical cavity, where V is the volume of the cavity, d is the diameter of the cavity and L is the height of the cavity, which is the same as the height of the air column. By rearranging the terms of the equation 2.27 it follows [13]
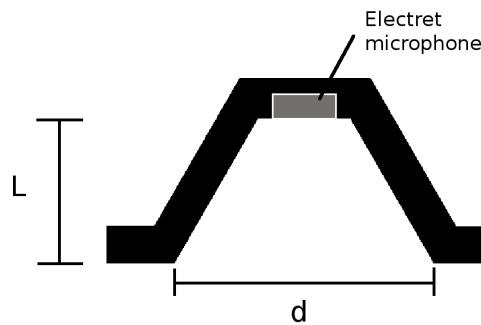

Figure 2.10: Width and height of the cavity

$$dP = -\gamma\frac{P}{V}dV, \tag{2.29}$$

that P and V are the static pressure and volume of the cavity, respectively. In the case of a cylindrical cavity and if assumed that the diameter of the cavity is constant, the equation 2.29 can be expressed as [13]

$$dP = -\gamma\frac{P}{L}dL, \tag{2.30}$$

where L is the height of the air column and dL is the change in the air column length caused by the vibrations of the skin. Equation 2.29 shows that when the height of the air column is increased, the pressure change decreases and the sensitivity of the microphone drops. This is in agreement with figures 2.12 and 2.13.

Theoretical sound pressure in the microphone was simulated by Wodicka et al. [17] for different cavity depths, based on the model of the acoustic circuit in figure 2.9. The authors calculated the pressure levels at the depths of 2.5mm, 5, 10 and 20mm. The pressure level at the depth of $l = 2.5$mm was chosen to be the reference level. The results are shown in figure 2.11. The referential values were obtained at the depth l as

$$\mathrm{dB} = 20 \times \log_{10} \left(\frac{p_{\mathrm{M}}, l = 2.5, 5, 10, 20\mathrm{mm}}{p_{\mathrm{Ref}}, l = 2.5mm} \right). \tag{2.31}$$

It was noticed that the pressure decreased as the cavity depth increased.[17]

The *in vitro* measurements were done with an artificial chest in a sound-proof chamber. The measurement setup of Wodicka et al. was the following[17]: a long horn of 10 cm was connected to a Peavey 22A loudspeaker. The proximal diameter of the horn was 2.2cm and the distal diameter was 4.3cm. A wet sponge (thickness 2cm) was sealed into a latex pouch and the sponge was stretched over the distal opening of the horn. As a result, a flat horizontal surface was achieved and the wet sponge simulated the acoustical properties of a chest wall. The authors noticed that no attempt was made to exactly match the surface impedance to that of the estimated impedance of the chest wall. An air coupled electret microphone, a Sony EMC-155, was mounted to cavities at depths of 2.5,5,10 and 20mm. The plastic coupler was sealed to the surface with double sided tape along the cavity circumference. Every cavity was vented with a needle measuring 3 cm and with a lumen diameter of 0.43mm. In each measurement a signal between 200Hz and 600Hz with equal energy was fed to the loudspeaker. The sampling frequency was 10kHz and the bandpass filter was adjusted between 100 and 2000Hz. The spectrum was obtained with a 2048 point FFT transform and the Hanning window was used. The spectra adopted and modified from [17] are shown in figure 2.12. The same decrease in sensitivity at higher frequencies with increasing cavity depths can be seen when these results are compared to the simulated results (see figure 2.11).

In situ measurements were done by Wodicka [17] for one of the investigators in a soundproof chamber. The subject breathed through a pneumotachograph and the target air flow was set to 2l/s. A microphone was attached to the right anterior chest to the place where the maximum breathing sound was found when auscultating. The microphone was attached with doubled sided tape as in the *in vitro* measurements. The signal was processed in the same way as was done in the in vitro measurements. Data was selected and analyzed when the inspiratory flows were in the range of
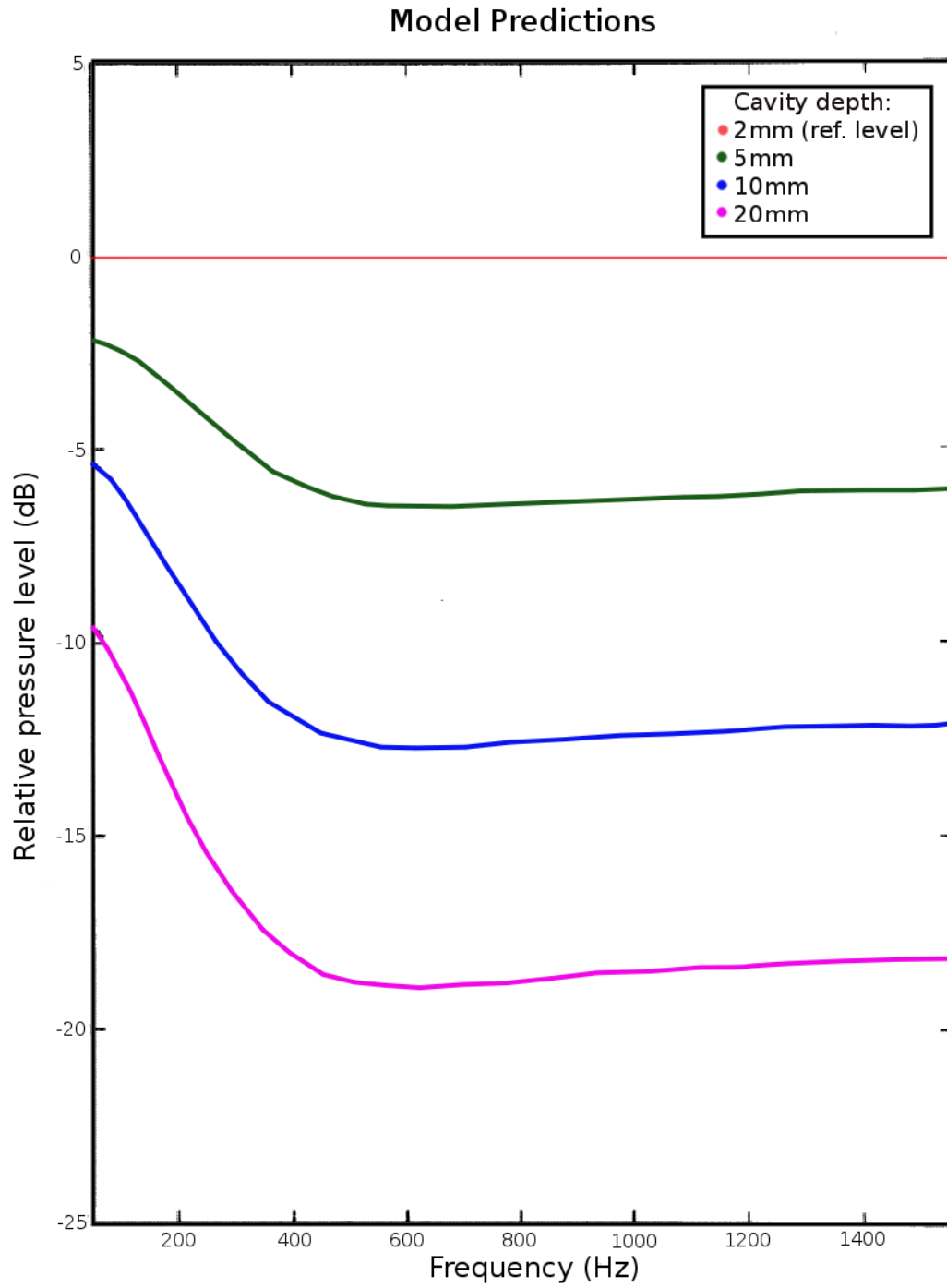
Figure 2.11: Simulated relative pressure levels. Modified from [17]
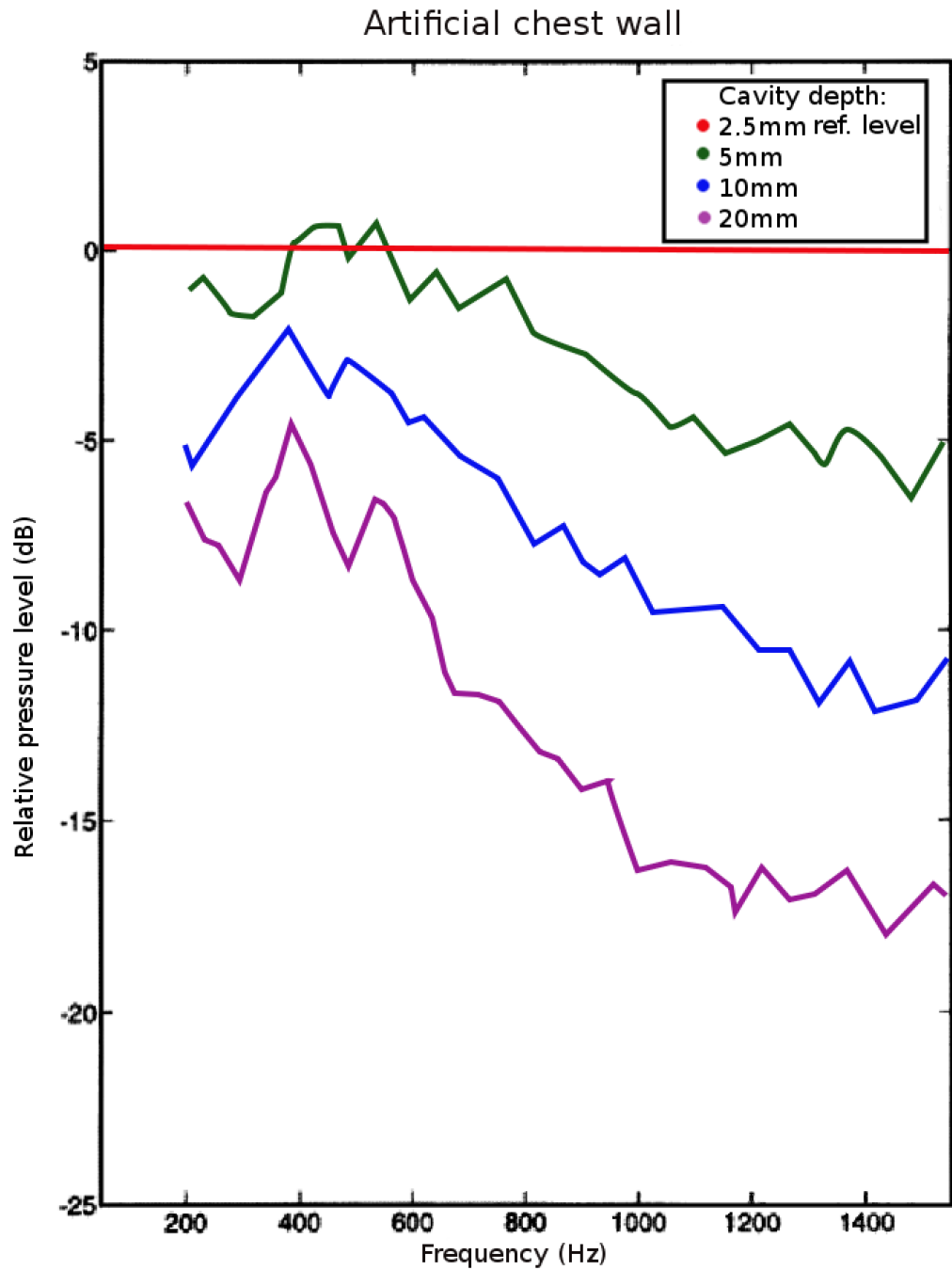
Figure 2.12: Results of the *in vitro* measurements. Modified from [17]

1.5l/s to 2.5l/s. The background noise level was measured when flow levels were between 0.0 and 0.1l/s. The resulted spectra are shown in figure 2.13.
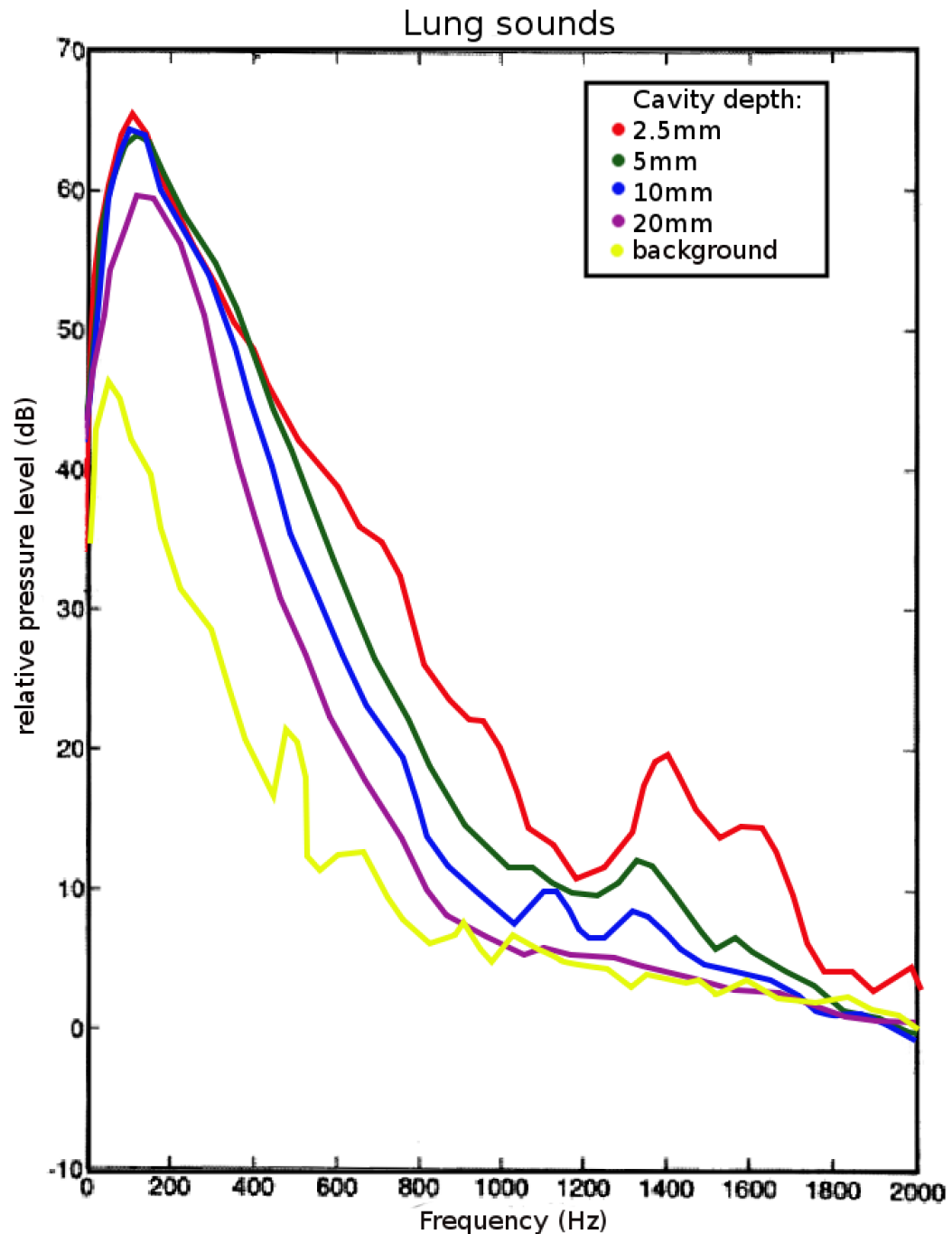
Figure 2.13: Results of the *in situ* measurements. Modified from [17]

In a cavity depth of 2.5mm the noise level was achieved approximately at 1700Hz and in a depth of 20mm the noise level was achieved at 800Hz. This finding supports the assumption that a smaller cavity depth increases sensitivity at high frequencies. The authors point out that for higher frequencies there are multiple peaks instead of plateaus and they may represent heterogenities of the sound source or the chest

wall. However, the best performance of the 2.5mm cavity was evident also in *in situ* measurements.

2.2.3   Effects of cavity width, shape and venting

In 1995 Kraman et al. [24] made *in situ* measurements with an aircoupled electret microphone. The electret microphone (Sony EMC-155) was kept the same through all the measurements, but a variety of plastic couplers was used. The effects of cavity width, shape and venting were studied. The cross-sectional schematic diagram of the conical coupler is presented in figure 2.10. The cavity width is another presentation of the surface area of the cavity. Mostly two kinds of shapes for cavities have been used in breath sound research; cylindrical and conical cavities (sometimes called bell shaped cavities). The cavity was vented with a needle which allows excess pressure to dissipate while preventing the entrance of sound into the chamber. The diameter of the chamber varied from 5 to 15mm, but the depth of the chamber was kept constant at 2mm. The studied subject was one of the authors and the measurements were done in a soundproof chamber. The microphone was attached to the lower chest with a double-sided adhesive ring. The subject breathed through the pneumotachograph and the target flow was 2 l/s. The same signal sampling conditions as in article [17] were used. Only the breath sounds with flows of 2 ± 0.5l/s were used for analysis. In figure 2.14 there is a comparison of spectra recorded using conical and cylindrical chambers with different chamber diameters. The background spectrum is also presented.

Measurements with altering cavity width and shape were made without cavity venting. In the frequency range of 500 to 1500Hz, the conical couplers were generally more sensitive, but below 500Hz only a small difference was observed.[24]

The effects of venting were studied exclusively with expiratory breathing sounds. Different sizes of needles were tested with a coupler size of 10mm in diameter. Between 800Hz and 1400Hz the signal contamination of external sounds became evident when the width of the vent increased and the length decreased (compare to figure 2.17 C).

The authors noted that the effects of cavity width are much smaller than the effects of cavity depth. They also mentioned that this is not surprising because this was predicted by the mathematical model adopted from their previous article [17, 24].However, the authors studied the model only from the point of view of the cavity. The effects of cavity width were not discussed by the authors, even if they exist in the equations presented by them [17]. The surface area A and the circumference S was introduced in equation 2.19, where the terms of acoustic resistance, compliance, inertance and conductance were introduced. The surface area and the circumference affect all these terms. What might have affected the
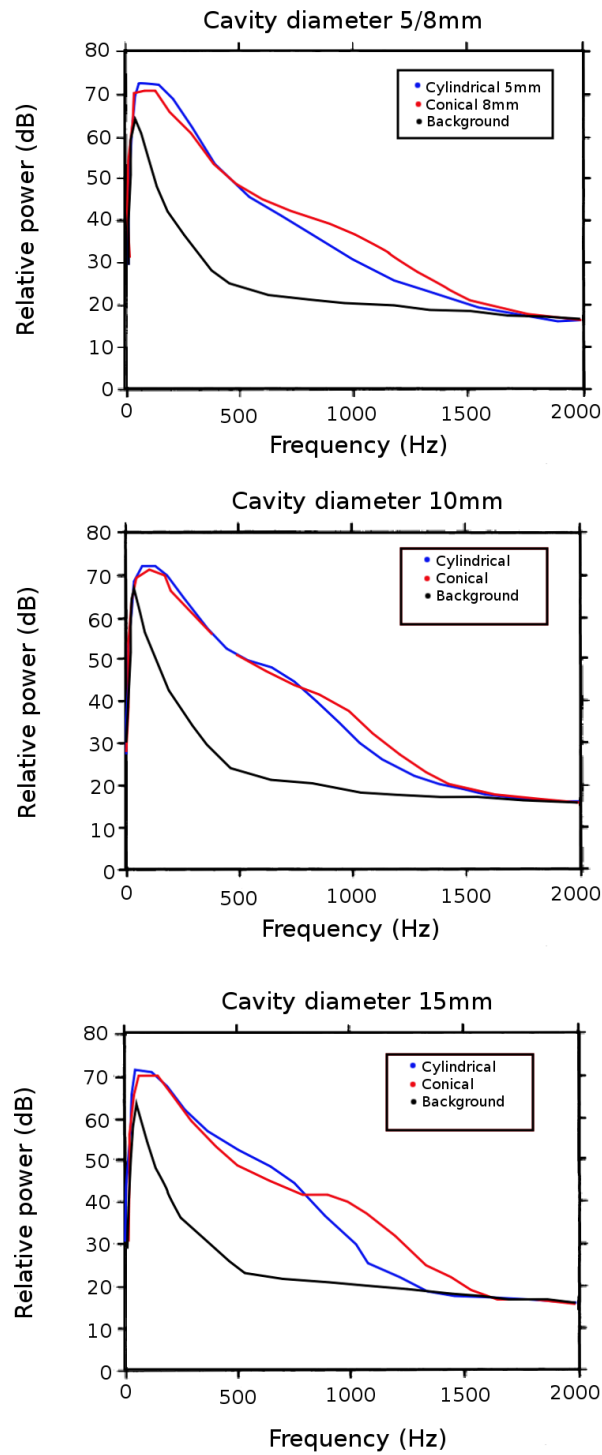
Figure 2.14: A comparison between cylindrical and conical couplers. The lower amplitude spectra in the figures is the back ground noise level measured at zero flow. Modified from [24]

measurement results is that when the effects of width (i.e. the surface area) were studied, the cavity diameter was altered while the height was kept constant. It followed that the absolute value of the cavity volume was altered. In their previous article the authors noticed that the cavity volume affected the acoustic compliance and it was shown that a small cavity provides the best impedance matching. When the diameter of the cavity was increased, the authors observed no great changes even if their cavity volume was increased at the same time. If their model is correct, it might implicate that the effects of the cavity diameter and the circumference on acoustic compliance were cancelled by changes in acoustic resistance, inertance and conductivity.

The authors also noticed that the conical coupler was better than the cylindrical one (compare to figure 2.14). The theoretical explanation is that the ratio of the surface-area and the air-chamber volume is larger for conical couplers than for cylindrical couplers. [24]Other explanations can be given by the displacement gain of the coupler presented in equation 2.6 and by a reduced cavity volume expressed in equations 2.29 and 2.19. The authors also state that when cavity venting is applied, great care is recommended and they even suggest it would be done best without venting, because usually double sided adhesive tape is not air proof, therefore excess pressure does not exist. [24]

2.2.4 The BioAcoustic Transducer Testing system

The **BioA**coustic **T**ransducer **T**ester (**BATT**) was designed by Kraman et al. in 2006 [2]. It was introduced because it is difficult to compare different kinds of sensors with the human body and the results are often inaccurate. The BATT system is made of a viscoelastic polyurethane elastomer material called Akton®. The authors reported that Akton simulates the acoustical properties of the tissue well, although the results of the comparison were not presented. A schematic diagram of the BATT system is presented in figure 2.15. Below the polyurethane layer there is a cavity called the antechamber where the loudspeaker of the headphone and the electret condenser microphone are placed. The idea of the test is to produce white noise with the loudspeaker and record it simultaneously with the antechamber microphone and with the sensor that is being tested and which is attached to the surface of the Akton. The microphone in the antechamber acts as a reference microphone and the sensor being tested is placed above the Akton layer. The signal from the tested microphone is compared to the signal coming from the reference microphone. The more the signals resemble each other the better the sensor is. By measuring one sensor at a time they can be compared with each other according to their acoustical properties.[2]. In the same year Kraman et al. [3] made a comparison of lung sound transducers using BATT. The FR and the wave form reproduction of the transducers
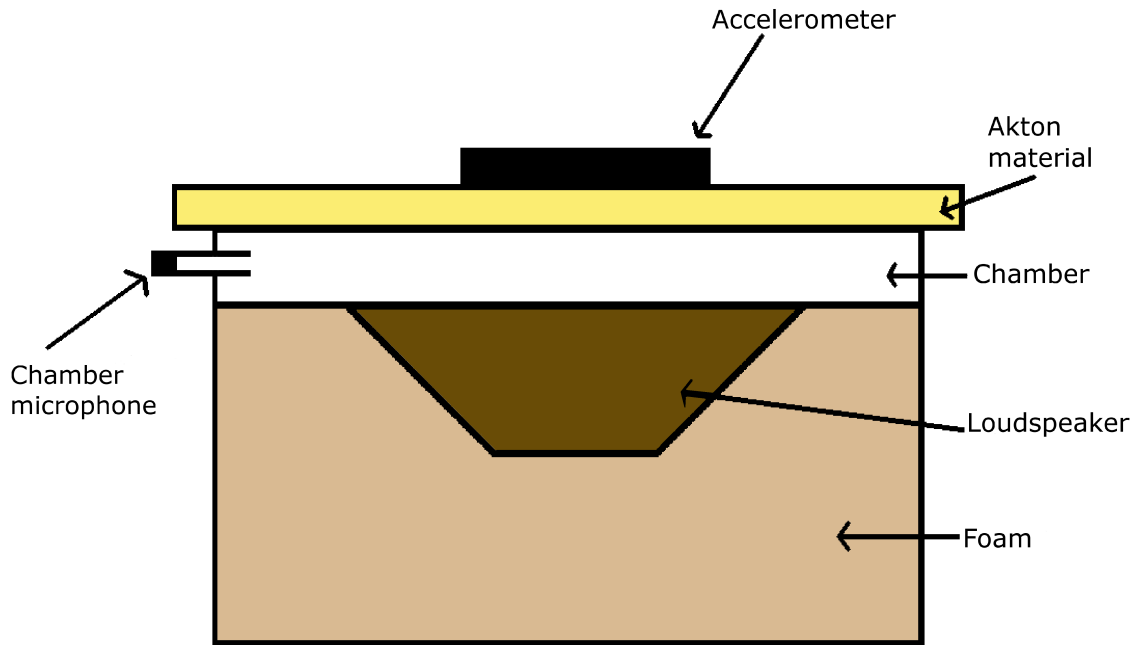
Figure 2.15: A schematic diagram of the BATT system. Modified from [2]

were investigated. For the FR a flat (±2dB) white noise between 30Hz and 4000Hz was supplied to the speaker of the BATT system. The volume at the surface was adjusted so that for every transducer the SNR was 60dB.The results of the FR measurements are shown in figure 2.16. The frequency and amplitude scales are the same, but there is an offset on the y-axis. The minimums of the signals are adjusted to the same level.

The authors state that all the transducers had good sensitivity in the range of 200Hz to 1000Hz, where the major components of the lung sounds are. All the sensors had some sensitivity up to 2000Hz, except for the Andries Tek. The PPG had a relatively high sensitivity up to 4000Hz, but from 1kHz to 200Hz the sensitivity dropped by 12dB.[3] Without the peak at 1000 Hz the PPG would have been clearly the best sensor tested. The sensitivity of the air chamber coupler was the smoothest of all the transducers in the range of 200Hz to 1000Hz. From 1kHz to 2kHz the sensitivity decreases by 30dB, but from 2kHz to 3kHz it decreases only by 5dB. After 3kHz the sensitivity reduces again rapidly. In figure 2.17 the effects of the shape, width and venting of the cavity for the electret microphone are presented. The authors state that in figure 2.17 the differences between different cavity widths are insignificant (< 3dB). A small trivial difference can be seen also to the benefit of the conical cavity in the range over 1kHz in figure 2.17 B. The corrupting effect of the vent is evident from figure 2.17 C. In table 2.2.4 the impulse response of the transducers are investigated with the means of initial deviation width (IDW)
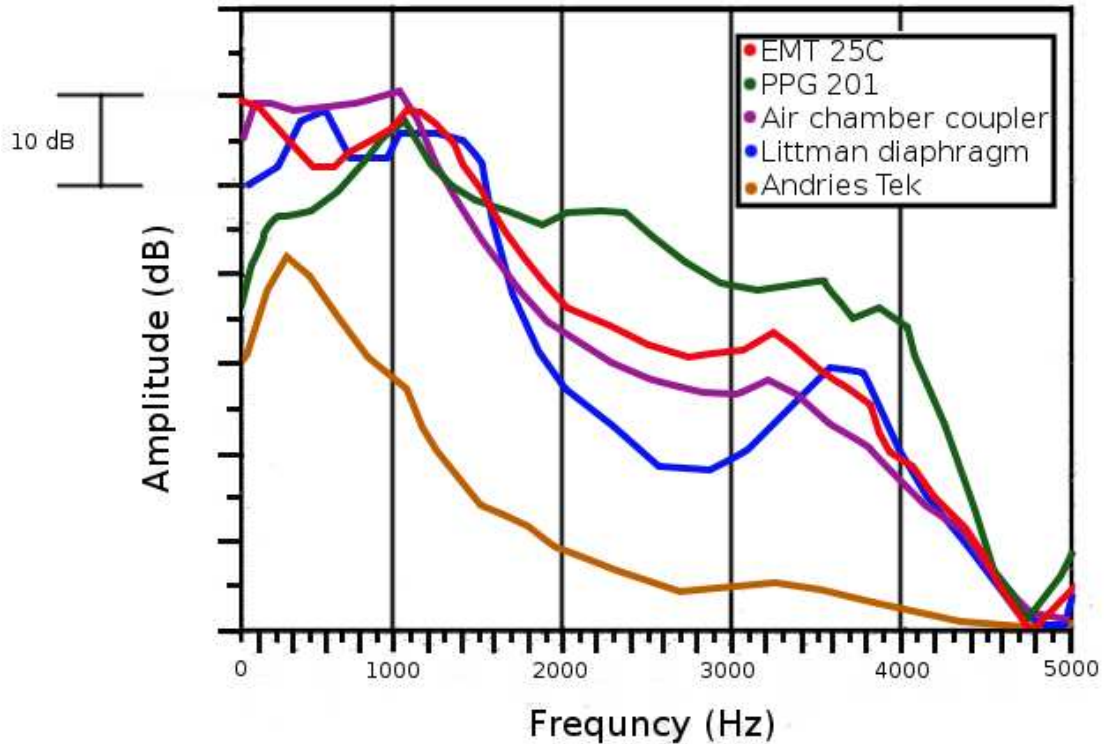
Figure 2.16: The frequency responses of the transducers in a relative y-scale. The minimums of the signals were adjusted on the same level. Modified from [3]

and cross-correlation. The IDW is the duration of the first half cycle of the 1kHz signal[25], which was used to drive the loudspeaker.

It can be seen that the PPG has the best and the Andries Tek has the worst correlation with the signal from the chamber microphone. The air coupled Sony EMC 150 was the best after the contact transducers.

In a spectrum comparison the authors showed that the spectra of the PPG 201 transducer and the chamber microphone were quite similar to each other[3][4]. The difference between the spectrum of the PPG and that of the Siemens EMT 25C accelerometers was the greatest at frequencies over 1200 Hz. The performance of the PPG 201 was very good. Its spectrum was almost identical to that of the chamber microphone up to 4800 Hz. Littman's stethoscope is an ordinary stethoscope for clinical use. Littman's cross-correlation coefficient was the second worst, which implies that it should not be used in quantitative measurements. The authors observed from the BATT measurements that at frequencies over 400Hz Littman's bell configuration had a higher attenuation compared to the diaphragm configuration. According to the authors, the width of the chamber is insignificant. This is con-

---

[4]The PPG sensor is an accelerometer specially designed for breathing sound detection
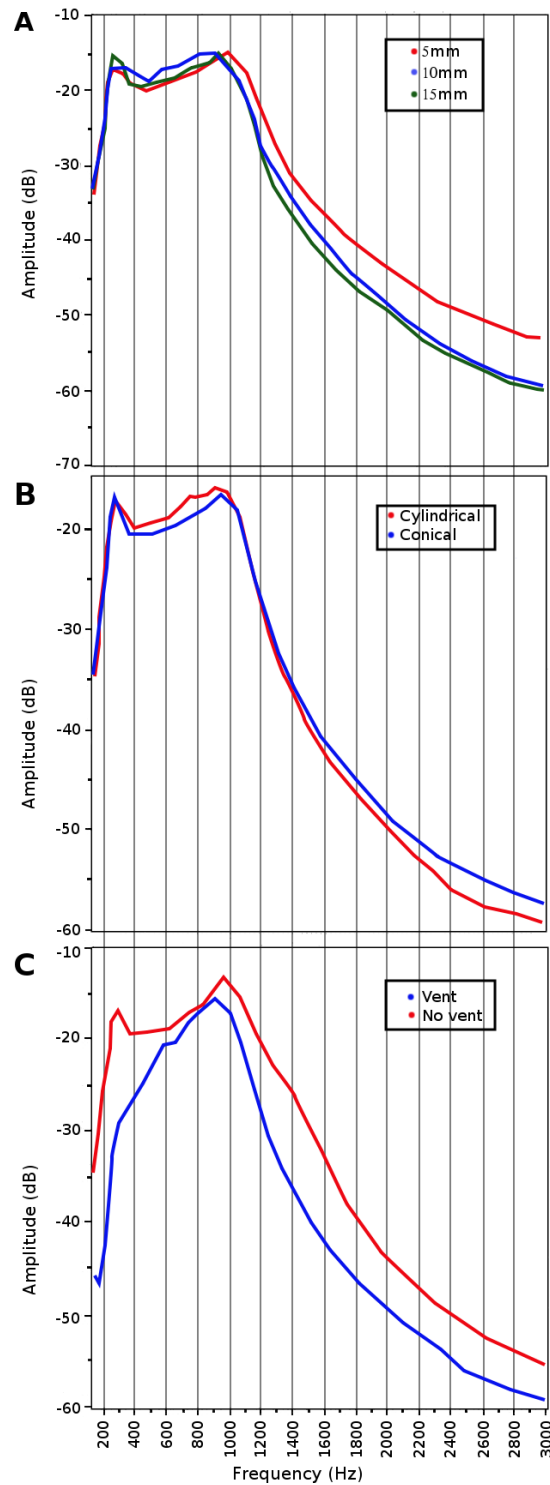
Figure 2.17: A) The effects of cavity width while cavity depth was kept constant at 2mm. The widths of 5, 10 and 15mm were investigated. B) The effects of cavity shape. Conical and cylindrical cavities were compared. C)The effects of a badly chosen venting needle. A larger needle was chosen than in previous articles from the same authors. Modified from [3]

Table 2.4: The response of the sensors to the pulses. The loudspeaker in the chamber was driven with 1kHz and the second row from the top shows the signal from the chamber microphone. The rest of the signals are compared with a chamber microphone signal by normalized cross-correlation.[3]

Sensor model	IDW (ms)	Cross-correlation coefficient (comparison to chamber microphone)
Chamber microphone	0.48	1
PPG 201	0.50	0.98
Siemens EMT 25C	0.61	0.93
Sony ECM 150	0.64	0.90
Radio Shack 33-1052	0.64	0.82
Littman diaphragm	0.57	0.77
Littman bell	0.62	0.68
Andries Tek	0.64	0.47

sistent with previous articles. However, their own measurements show that a 5mm cavity has the best sensitivity at high frequencies, although the advantage is quite small (see fig. 2.17). The theory and the model presented in article [17] predicts the same kind of results. The decrease in the width of the cavity should improve the acoustic coupling of the microphone. An increase in the cavity width increases the volume of the cavity at the same time, which was previously suggested to reduce the performance of the sensor. It would have also been interesting to compare the air coupled electret microphone Sony ECM 150 with

the Siemens contact sensor with this same method because as shown in table 2.2.4 they had quite similar correlation coefficients.

2.2.5 Effects of ambient respiratory noise on tracheal sounds

Pasterkamp et al. [26] have studied the effects of air transmissions of breath sounds from the mouth and nose to the sensors on the chest and trachea. Louder pathological sounds like snoring were not included in their study, but it was assumed that louder sounds produce at least as much corruptance as normal breathing sounds. The breathing sounds were recorded in a sound proof dual chamber which was constructed in a way that a smaller chamber was placed into a larger one. The sounds were recorded in two phases. During the first phase the breathing was directed outside from the inner chamber and during the second one the breathing was redirected to the inner chamber. A nose clip and targeted flow rates were used in all measurements. The ambient noise was measured with two electret microphones suspended in the air and they were placed close to the contact microphone and the air coupled microphone, which were mounted on the skin. Background spectra were recorded during breath holds. [26]

The ambient sounds during expirations were significantly reduced while the expirations were directed outside from the inner chamber. No changes or only small changes were observed during inspiration and breath hold. A spectral peak at 960Hz was observed during the ambient sound in the power spectrum. The authors also found this peak from the lung and tracheal sound transducers, which indicates that ambient sounds may have a corruptive effect on contact microphones. The authors also state that air-coupled microphones were affected more easily by ambient sounds than contact microphones. The same phenomenon is observed with phonocardiogram microphones [8]. At the same time they found other peaks in the spectrum which were not related to externally coupled breathing sounds.[26]

Air-borne and tissue-borne sensitivity of the bioacoustic sensors has been studied also by Zanartu et al [27]. Both light-weight accelerometers (contact microphones) and air-coupled microphones were investigated. The air-coupled microphone was a Sony EMC-77B mounted on a plastic coupler and the accelerometers were light weight Siemens EMT25C's and very light weight Knowles BU-7135's. According to the authors, air-borne sensitivity of bioacoustic sensor in relation to tissue-borne one has not been consistely investigated. They also state that the most typical air-borne sound sources are breathing and room noise. In some applications the study subject is required to speak and this is a bigger challenge. [27]

Tissue-borne sensitivity was measured using the BATT system [27]. For air-borne sensitivity measurements the microphone was mounted onto the chest of the study subjects. This was done because it was discovered that the human body provides a large and unique mounting configuration. The true conditions of skin mounting are difficult to mimic artificially because the body determines the amount of acoustic energy which is transferred to transverse the vibrations of the skin. A loudspeaker was placed 26" from the anterior chest. High acoustic pressure, 60dB over noise floor, was used. Air-borne sensitivity was measured with three sensors attached to a second intercostal space over the anterior upper lobe, to the suprasternal notch and to the posterior right upper lobe. The colouring effect of room resonances was compensated by using a near-skin microphone. The face of the microphone was situated perpendicular to the loudspeaker and the distance to the skin was 1". The concept of tissue-to-air ratio (TAR) was introduced in the article. For an ideal bioacoustic transducer the TAR is high, which means that the sensitivity of the sensor is high for tissue-borne sounds but low for air-borne sounds. It is similar to the SNR except that in this case the noise is associated with the airborne component and the signal is associated with the tissue-borne component of the sound. In addition, the idea of a BioAcoustic Insulator (BAI) was introduced. The BAI is a device (e.g. earmuffs) which can be placed over the transducer to insulate the sensor from arriving air-borne sounds. [27]

It was observed that the highest TAR curve was found to be with the Knowles accelerometer whereas the air-coupled microphone showed the smallest TAR curve. The air-coupled microphone has high sensitivity for both air-borne and tissue-borne sounds. Therefore, air-borne sounds may easily corrupt the recorded sound signal. The best result for the BAI can be achieved when the air-coupled microphone is used. The 3" PVC cap was proved to be the most effective. In the case of accelerometers the BAI can amplify or reduce certain frequencies, which makes it less suitable for accelerometers. It was pointed out that the BAI can reduce tissue-borne sensitivity at low frequencies. For the air-coupled sensor the reduction was -10dB below 315Hz. On the other hand the heart sounds increased when the BAI was present. The authors note that the BAI is best suitable for air-coupled sensors. [27]

As a conclusion the authors note that the sensitivities presented are based on a radiation of voiced sounds on a sternal notch. They suggest that the method presented could also be used in other applications. The gain of the bioacoustic sensor affects the air-borne and tissue-borne sensitivities, but does not affect the TAR. The best results for the TAR and the sensitivities was achieved by Knowles' accelerometer. However, this sensor is not so sensitive to tissue-borne sounds, meaning that it should be used with higher vibration amplitudes like when recordings of voiced sounds on a sternal notch are made. The Siemens accelerometer performs best at frequencies below 200Hz. The TAR performance of the air-coupled microphone can be greatly improved when using the BAI. An improvement is seen especially at mid-high frequencies. However, the authors note that handling an air-coupled microphone that has been improved with the BAI can be quite demanding. [27] The authors suggest that the corrupting effect of the air-borne sound could explain some discrepancies and uncertainties observed in studies where bioacoustic transducers are being used. [27]

Summary The transducer used for tracheal sound measurements has a great effect on sounds recorded. Bad design attenuates the high frequencies first, because the amplitude of the sound decreases rapidly when the frequency increases. The acoustic impedance should be matched between the skin and the transducer. However, the acoustic impedance of skin is difficult to measure because the location and sensor used for the measurements affect the results. In general, contact transducers had a better performance than air coupled transducers. However, if the coupler of the air coupled transducer is properly designed the performance of the transducer can be improved significantly. As a general rule the height of the cavity should be as small as possible. A bioacoustic testing system which can be used for analytical comparison of different sensors was presented.

A general testing system available for all researchers is needed because research

groups should be able to test their own measuring systems. Also, this would help them to compare their results with those of other research groups.

3. TRACHEAL BREATH SOUNDS

Since the 1990's, several articles have been written concerning tracheal breath sounds. Both normal sounds and sounds indicating disorders have been covered in articles. Tracheal breath sounds(TBS) have always had a special role in breath sound research, because both upper and lower airway sounds are coupled in the tracheal region. TBS have been used in lung sound research, because TBS are louder than lung sounds, they are less filtered by the chest wall than chest sounds and they correlate with lung sounds recorded from the chest wall[28]. Properties of tracheal breath sounds have been studied in many papers, but in many cases TBS are investigated from a lung sound study point of view. Recently, the TBS have received more attention, because it has been discovered that they are related to the geometry of upper airways and that sound amplitude is a function of breath flow. Therefore, it is suggested that it would be possible to use TBS as a diagnostic tool for upper airway obstructions [29, 30, 31]

Characteristic patterns in tracheal sounds have been associated with conditions affecting airway patency such as asthma, obstructive sleep apnea, infections, airway edema, malformations and tumours. Tracheal sounds have usually been measured from the neck at the suprasternal notch over the extrathoracic area. [32] TBS are relatively large in amplitude and they have close relations with tracheal airflow. They have a wider spectrum than lung sounds and the spectrum is similar to a broadband noise source with shallow spectral peaks ranging from 80Hz to 1500Hz. [32]

Normally the tracheal sound amplitude increases when the position is changed from sitting to supine. Pasterkamp and his colleagues showed that subjects with sleep apnea cannot properly dilatate their pharynx even while they are awake.[31] It was discovered that apnetic patients with obstructions had a larger amplitude increase than healthy people. Yonemaru [33] measured tracheal sounds from the right side of the neck at the level of thyroid cartilage. They demonstrated that patients with tracheal stenosis presented an increase at the peak of the spectral power at about 1kHz and an overall increase in the spectral power in the range of 600 to 1300Hz when compared to normal subjects. [32, 34] Sanchez et al.[35] discovered a significant difference in tracheal sound spectra between adults and children. In the spectrum of tracheal sound, natural resonance frequencies and

sound amplitude were higher with children than with adults. This was explained to be caused by the different dimensions of the upper and central airway dimensions. They also concluded that similar measurements might be of value when studying abnormalities in the upper and central airways.[32] These findings suggest that TBS might contain information about the state of upper airways.

3.1 The origin of tracheal breath sounds

There are two components in a tracheal sound measured from the surface of the neck. Both are due to turbulent airflow in the trachea. In certain conditions the flow produces an acoustic sound wave. This phenomenon can be explained with terms of physics in fluid mechanics. Because the trachea and its surrounding tissue have large acoustic impedance, the acoustic wave travels through the tissue and produces vibration which can be detected with a transducer on the surface of the skin [36, 3]. Another source of vibration is pressure fluctuation inside the trachea, which are caused by the turbulent flow in the trachea [36, 37]. These mechanical waves are observed as vibration as well. In the next subsection the production of acoustic waves by turbulent flow is investigated in detail by referencing Micheal Krane [37].

3.1.1 Aeroacoustic sound generation

Investigators of breath sounds have measured breath sounds and compared their results with the theoretical ones. However, a theoretical explanation of flow-induced sound has not yet been presented. Speech science is interested in sounds produced by the human respiratory system. However, they have focused on lumped-element models of the motion of the air in vibrating vocal folds and even the necessity of turbulent airflow for producing sounds has been known since the 1960's.[37]

In 2005 Krane [37] wrote an article on unvoiced sound production in speech. The article presents results from the aeroacoustic theory and describes how airflows can produce sound. The focus of the article is in flow vorticity, which is essential to comprehend in order to understand the dynamics of turbulent flow and aeroacoustic production of sounds. The paper presents the key ideas of fluid dynamics and aeroacoustics that are necessary for understanding turbulent sound production. They are used to predict sound level and spectral characteristics. The studied environment is the vocal tract, but the presented results of fluid dynamics are also valid for flows in the respiratory tract. This is mainly because the flow speed is approximately the same in both cases.

Because airflow dynamics is a very complex phenomenon, approximations are needed. Reasonable formal expressions are provided when it is assumed that the

nonacoustic motion behaves like an acoustic source. As a result, it can be noticed that an exact representation of the acoustic source is unnecessary. [37].

According to the author, in order to understand flow-induced sound production there are two phenomena that must be understood properly. [37].

1. a description of the production and evolution of jet vorticity and

2. a description of how this motion produces sound.

Every tracheal breath sound can be expressed by two modes of motion: flow and sound modes. The flow mode is called "convective" or "incompressible" and the sound mode is called "propagative" or "irrotational". In the propagative mode the energy and momentum are transmitted through air particles by the propagation of waves of compression and expansion. Fluid particles do not experience any net displacement or rotational movement since the pressure forces of the sound field are acting uniformly on the fluid particles. By contrast, the convective mode is characterized by the transfer of energy and momentum through actual displacements and by the rotation of fluid particles. The net displacement might be rather large, as when air is displaced out of the lungs and expelled out of the body. The flow speed in speech is observed to approach 40m/s (M=Mach number=0.12) and in breathing 20-50 m/s, when M=0.06-0.15[36]. This means that for speech and breathing, the mode of motion is essentially incompressible, because of the low Mach number. This means that the mathematical model of turbulent sound production must [37]

1. support air particle rotation and

2. be incompressible.

A schematic diagram of air flow in the respiratory tract is presented in figure 3.1. When the lungs are compressed the flow $U_j(t)$ is forced through the constriction. While in the constriction the airstream decelerates and the flow *separates* from the tract wall and the *jet* is formed. The jet is a high-momentum region in the center and it is surrounded by stagnant air. Between the flow center and the stagnant air there are the so called *shear layers*. The particles in the shear layer are in translation and rotational motion. The *vorticity* is a sort of measure of rotational motion. In the shear layer the vorticity tends to diminish into *coherent structures* which may or may not be turbulent and which are convected from left to right in figure 3.1. The shear layer diffuses in the direction perpendicular to the flow causing the jet to spread in the transverse direction. This spreading of the jet is due to two reasons: the diffusion of the shear layer and the recirculation in the stagnant region. This structure is shown in figure 3.2. [37]. In the first mechanism the width of the shear layer grows spatially with the distance from the origin of the jet. This diffusion
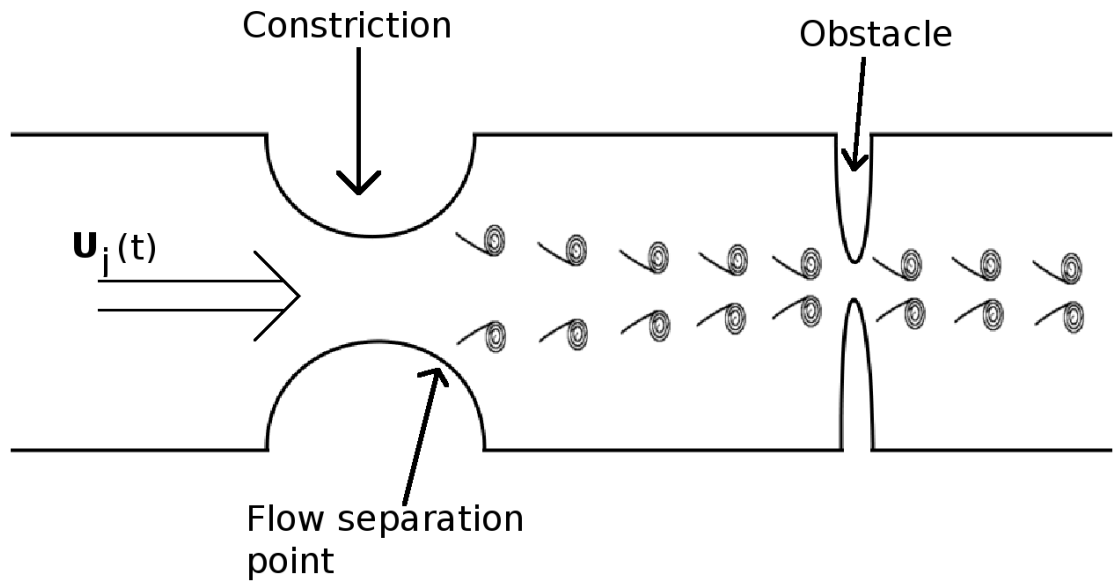
Figure 3.1: Air is flowing at a velocity $U_j(t)$ from the lungs from the left. Downstream from the constriction, the flow separates from the wall and the jet is created. In the boundary of the jet the vortices are convecting at the local flow speed which is a fraction of U_j. When the vortical structures are convecting in the tract they produce unsteady forces to tract walls, which produce sound. The sound production is particularly intense if the cross-sectional area of the tract changes rapidly, like in the case of an obstacle. Modified from [37]
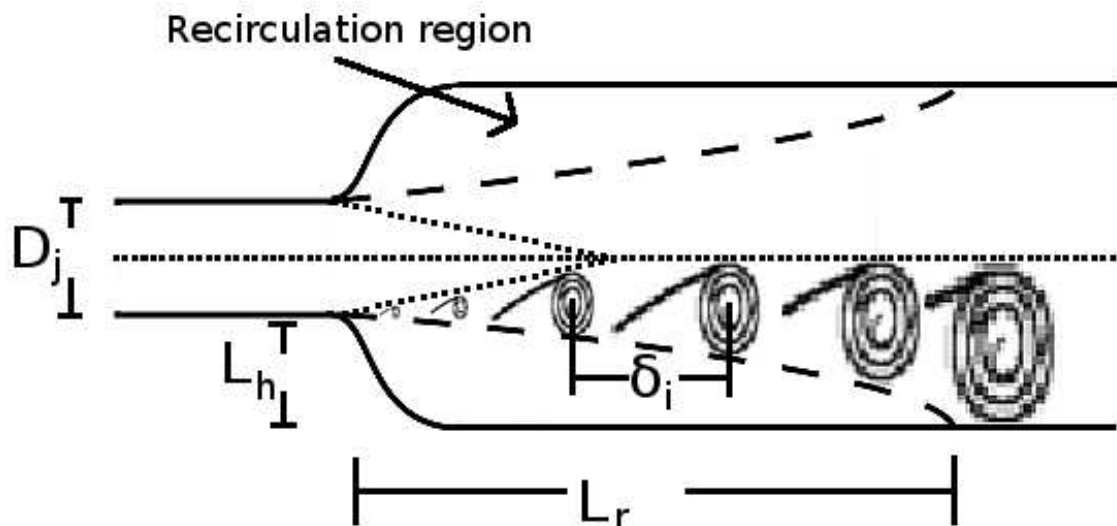
Figure 3.2: A schematic diagram of jet spreading due to diffusion and recirculation.

transfers the momentum of the jet to the stagnant region. The rate of diffusion is a function of shear layer thickness; the thinner the shear layer the higher the rate of diffusion. [37].

The second mechanism of jet spreading is recirculation in the stagnant air. Because the flow transfers its momentum to the stagnant air a part of the stagnant air is lost to entrainment. This entrained air moves with the jet along the jet boundary. The amount of stagnant air lost to entrainment must be replaced with new stagnant air. This produces weak recirculation in the "stagnant" region. The recirculation and diffusion are affected by each other and formulate a nonlinear problem in fluid dynamics. Eventually the jet is spread over the whole cross-sectional area of the tract. [37].

The vorticity is localized to the boundary layer of the flow before the point of separation. The vorticity is formed due to drag forces in the boundary layer which is a thin viscous layer near the tract wall. In the separation point the boundary layer injects rotational motion to the flow and the jet is created. This effect is called *boundary layer separation.* After the point of separation in fig 3.1 the jet goes through the obstacle, which is for example a change in the cross-sectional area of the duct. The vortical structures of the jet produce unsteady forces in the wall of the obstacle as they pass by it. These forces excite the wall to vibrate and a part of the energy is transmitted to irrotational mode, which is actually the acoustic sound. The sound propagates along the tract but the end of the tract does not allow perfect transmission. Therefore, a part of the sound is reflected back, and the acoustic energy is accumulated in the form of standing waves. The standing waves may transmit oscillation inside a resonator (tract) to the outside world through the mouth and tissues. Only a small fraction of the kinetic energy of the vorticity is transmitted to irrotational mode. Most of the kinetic energy is dissipated by viscosity with the aid of turbulence or convected out of the mouth by airstream. [37].

Acoustic pressure of aerodynamic sources The schematics of aeroacoustic problems are presented in figure 3.3. There is an infinitely long pipe with a uniform cross section area A. In the middle of the pipe there is a short constriction. The steady airflow convects the vortexes through the constriction and unsteady forces are applied to the constriction walls, which produces the acoustic disturbance and a sort of virtual sound source is turned on. The sound starts to propagate along the pipe. If the shape of the duct, the velocity of the flow U, and the vortical disturbances are known, the expression for sound pressure field at the observation point $\mathbf{x}$ can be written when the sound source is in location $\mathbf{y}$. The solution for this kind of a problem is the domain of aeroacoustics. Howe (1975,1998) has shown that
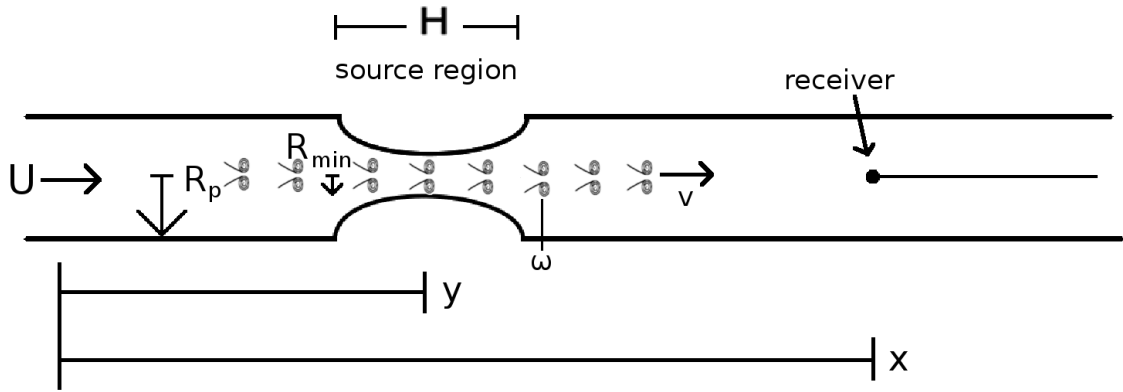
Figure 3.3: The aeroacoustic problem. The task is to find out what the sound pressure at observer x is, when aeroacoustic sound is produced at the location y. The pipe is infinitely long, and its radius is R_p. The minimum radius of the constriction is R_{min} and its axial length is H. ω is the vorticity and $\mathbf{v}$ is the local velocity of the vortices. Modified from [37].

concise formulation can be derived when the acoustic total enthalpy B' is used as a variable. The acoustic total enthalpy is defined as $B' \approx p'/p_\infty + \mathbf{U} \cdot \mathbf{u}'$, where $\mathbf{U}$ is the steady flow speed, $\mathbf{u}$' is the particle velocity of acoustic disturbances, ρ_∞ is the ambient undisturbed air density and p' is the sound pressure fluctuation. Far from the regions where the sound is produced, U is small and $B' \approx p'/\rho$. Howe also showed that for low Mach number flows (observed during breathing and speaking) the convected wave equation for the low-frequency acoustic field variable B' is [37].

$$\frac{D^2 B'}{Dt^2} - c^2 \frac{\partial^2 B'}{\partial x^2} = \frac{\partial}{\partial x}(\omega \times \mathbf{v})_x, \tag{3.1}$$

where c is the speed of sound and DB'/Dt is the convective or material derivative of B'. For fluid particle moving through space

$$\frac{DB'}{Dt} = \frac{\partial B'}{\partial t} + U \frac{\partial B'}{\partial x}, \tag{3.2}$$

where the first term on the right-hand side is the time rate of the accumulation of total enthalpy in a fixed location in space and the second term time is the rate of change of total enthalpy due to convection of total enthalpy fluctuations past a fixed point in space. In other words, the second term describes the convection of sound waves by the motion of a medium. If the right-hand side of the wave equation 3.1 is zero, the acoustic disturbance will propagate in the way that its energy is conserved. If the right-hand side is nonzero the disturbance will gain or lose energy. In the case that the term is nonzero, but the disturbance does not exist, one will be generated. Now it can be seen that the term on the right-hand side is doing work on the acoustic field and it is referred to as an acoustic "source" or "sink" depending

on whether the energy is gained or losed. As can be seen from equation 3.1 the physical mechanism that modifies the acoustic disturbance field is the acceleration of airflow, which is represented by the x-component of the term $\omega \times \mathbf{v}$. This term has a central role for understanding turbulent flow dynamics and it is related to the direct radiation of sound from the jets and to the aerodynamic forces induced on the tract wall by the jet[1]. [37] v The sound-pressure fluctuation caused by the interaction of the jet and the constriction is found by solving equation 3.1. The solution is found by convoluting the tailored Green's function with the source term $\partial(\omega) \times \mathbf{v}_x / \partial x$ (details passed) and the pressure becomes [37].

$$p'(x,t) = -\rho_\infty \frac{\text{sgn}(x-y)}{2A(1+M)} \int_A \int_y [\omega \times \mathbf{v} \cdot \mathbf{U}^*] \, dA dy, \qquad (3.3)$$

where sgn(x-y) is the signum function

$$\text{sgn}(x-y) = \left\{ \begin{array}{rl} 1 & \text{if } x-y > 0 \\ -1 & \text{if } x-y < 0 \end{array} \right.$$

The sign of the pressure is changed across the source. This is consistent with the dipole source(see fig. 3.5). A is the cross-sectional area and M is the steady flow Mach number at the receiver location x. The integrand is written inside brackets. It means that it is a function of both y and the retarded time $t - |\mathbf{x} - \mathbf{y}|/c(1+M)$ which is the time that the signal takes to travel from the source at $\mathbf{y}$ to the observer at $\mathbf{x}$. The $\mathbf{U}^*$ is the ideal flow velocity field which would exist if the duct contained steady airflow with a unit speed. This ideal flow does not exist in the tract and so $\mathbf{U}^*$ reflects the effects of source motion and diffraction around the change of the shape of the vocal-tract instead. Therefore, $\mathbf{U}^*$ can be thought of as a property of the tract shape. The integral solution of equation 3.13 shows that the sound is produced whenever the vorticity moves across the streamlines of $\mathbf{U}^*$. The streamlines are simply tangentials of the velocity field. Sound production is maximum when $\mathbf{v}$ is perpendicular to $\mathbf{U}^*$ and zero when $\mathbf{v}$ and $\mathbf{U}^*$ are parallel. The vorticity acceleration $(\omega) \times \mathbf{v}$ is difficult to approximate when compared to $\mathbf{U}^*$, but the dot product in equation 3.13 simplifies the situation. Only those components of the vorticity field that are normal to the streamlines of $\mathbf{U}^*$ will contribute to sound production. This also means that the sound production process may be thought to "filter" the vorticity field while the shape of the filter is determined by the shape of the vocal tract ($\mathbf{U}^*$). [37]

As presented above, the condition for the presence of the aeroacoustic sound source is that the vorticity moves across the streamlines of the ideal flow velocity

---

[1]These aerodynamic forces also cause tracheal wall vibrations that can be observed with transducers at the surface of the tissue

field. This means that the flow must be separated from a tract wall somewhere in the tract in order for sound to be produced. Because the tract wall lies on the streamline of $\mathbf{U}^*$, the vorticity in the streamline will not contribute to sound production. According to the author, this means that the unsteady axial force (drag) resulting from the sum of the axial pressure force contribution on the tract walls is zero unless flow separation occurs (a phenomenon which is known as D'Alambert's paradox). In addition, it is worth remembering that the jet does not have to collide with or strike on an obstacle in order to produce sound, the change in the cross-sectional area is in itself enough. [37]

The structure of the vortex ring is shown in figure 3.4. The ring is made of the

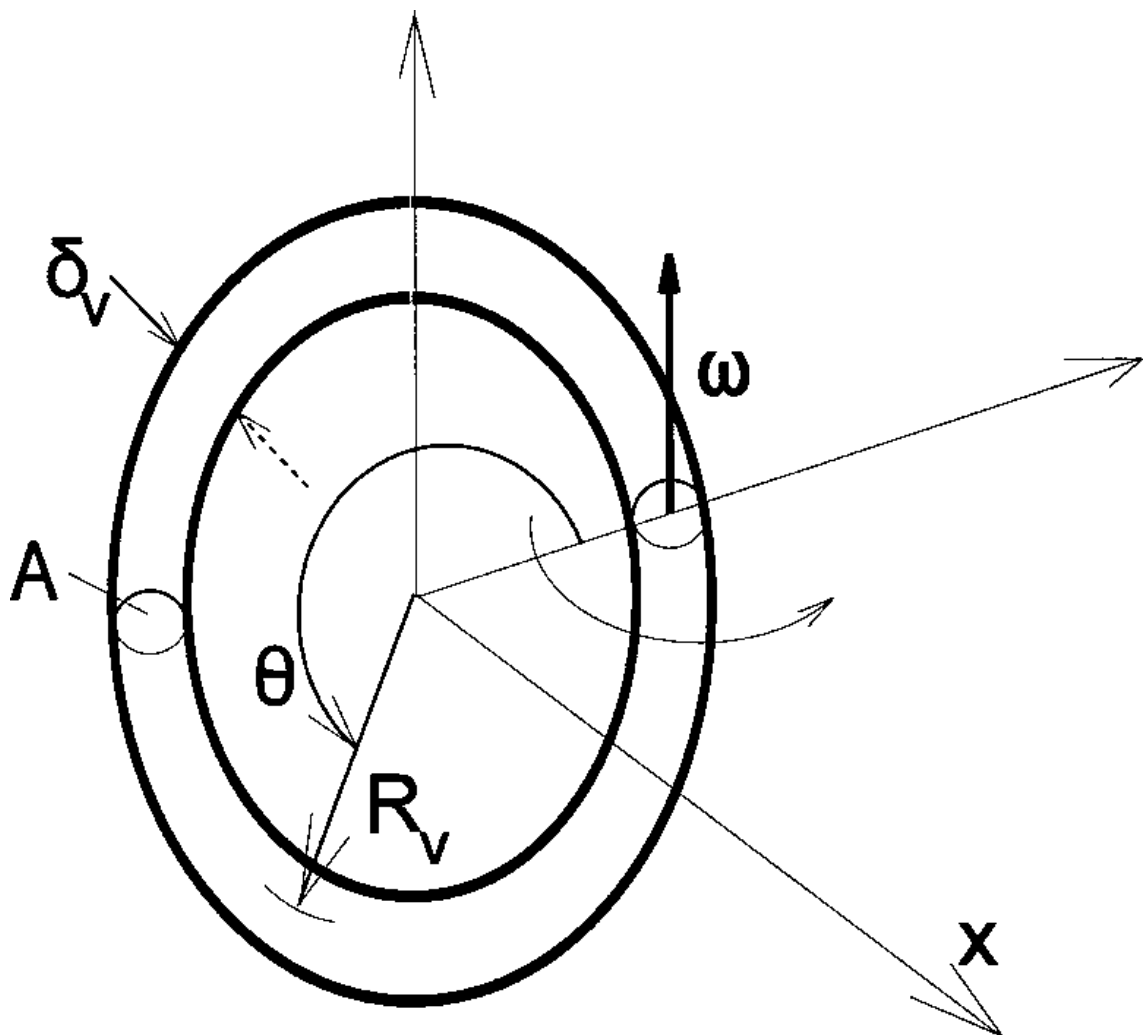Figure 3.4: The structure of the vortex ring. The radius of the ring is R_v, the core radius is δ_v, A_v is the cross-sectional area of the core and θ is the azimuthal coordinate. Adopted from [37].

circular tube of the vorticity. The vorticity vector points to an azimuthal direction θ. The tube has a cross-sectional diameter δ_v, and the ring has a radius R_v. The

ring moves from left to right, i.e. along the z-axis in fig 3.4. The volume integral of this vorticity can be expressed with a definition of circulation Γ [37]

$$2\pi R_v \Gamma = \int_0^{2\pi} \int_{A_v} \omega \cdot d\mathbf{A} d\theta, \tag{3.4}$$

where $A_v = \pi \delta_v^2$ is the area of the vortex ring normal to ω, and the circulation $\Gamma = \omega(\pi \delta_v^2) = \omega A$. The pressure fluctuation $p' \approx \rho_\infty B'$ in the region far away from the source is [37]

$$p' = -\frac{\rho_\infty \pi \mathrm{sgn}(x-y) R_v}{A(1+M)} \left[\Gamma(\mathbf{e}_\theta \times \mathbf{v} \cdot \mathbf{U}*) \right], \tag{3.5}$$

where e_θ is a unit vector in the vorticity direction. If assumed that the vortex takes a radial path then $\omega \times \mathbf{v}$ always points radially outwards and the vortex produces sound if the radial component of $\mathbf{U}^*$ is nonzero.

Figure 3.5(a) presents a vortex ring convecting through the constriction. The amplitude and the spectrum of the sound are presented in figures 3.5(b) and 3.5(c), respectively. The radial component of $\mathbf{U}^*$ points initially inward, upstream of the constriction and outward, downstream of the constriction. The source strength is a single period sinusoidal wave in figure 3.5(b). The change of sign in the middle of the constriction is consistent with the dipole source, as previously mentioned. The source duration can be approximated as l/U_c, where U_c is the convection speed of the vortex. l is chosen to be the axial length of the constriction H or the radius of the vortex core δ_v. If $\delta_v/H << 1$ then H is chosen and in the opposite case δ_v is chosen. The spectrum in figure 3.5c has a broad spectral peak centered at $f = U_c/l$ and the width of the peak is approximately U_c/l. When the scalar product in equation 3.4 is written out and when using $U^* = U(x)/U = A(x)/A$ the following is obtained: [37]

$$p' = \rho_\infty \frac{\pi}{A} R_v \Gamma_v \frac{A}{A(x)} sin\alpha(t), \tag{3.6}$$

where α is the angle between the pipe axis and $\mathbf{U}^*$. When estimated that $\omega \sim U_c/\delta_v$, then $\Gamma \sim \omega \delta_v^2 \sim U_c \delta_v$. $A(x)$ may be estimated by using its minimum, $A_{min} = \pi R_{min}^2$ and $sin\alpha \sim (R_P - R_{min})/(H^2 + (R_P - R_{min})^2)^{1/2}$, where R_p is the radius of the pipe away from the constriction and R_{min} is the minimum radius of the constriction. When these approximations are applied to equation 3.6, the acoustic pressure is proportional to [37].

$$p' \sim \rho_\infty \frac{R_v \delta_v}{R_{min}^2} U_c^2 \frac{R_p - R_{min}}{((R_p - R_{min})^2 + H^2)^{1/2}}. \tag{3.7}$$

From the previous equation it can be seen that the larger the vortex rings (in terms
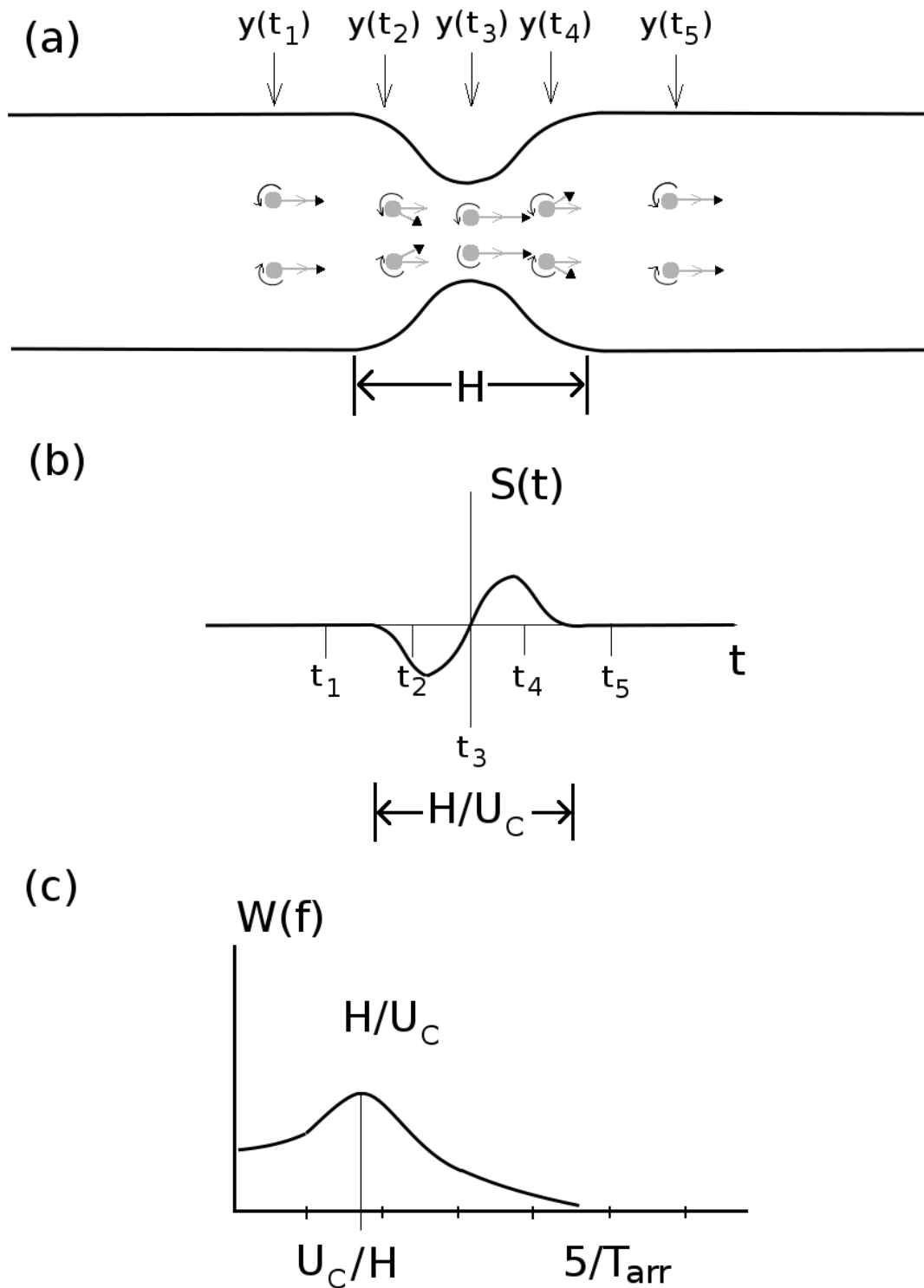
Figure 3.5: (a) Location of the vortex at five instants t_1,t_2, t_3, t_4,t_5. (b) Sound strength at five instants corresponding to figure (a). (c) T he spectrum of the sound pressure, when a single vortex is convected through the constriction. Modified from [37].

of either core size or radius), the faster the convection speed is, the larger the radial changes of the tract shape are and the greater the sound production will be.

The scaling law for the acoustic source energy as a function of the vortex ring convection speed is [37]

$$\frac{p'^2}{\rho_\infty c^2} \sim \rho_\infty U_c^2 M^2. \tag{3.8}$$

It can be seen that the sound power is proportional to the second power of the convection speed and the Mach number. The Mach number is based on convection speed which means that the power is proportional to the fourth power of convection speed. In freefield the power of the acoustic dipole is proportional to the sixth power of convection speed. The difference occurs because in these models the acoustic wave is a one dimensional plane wave and wave in free space is spherical. The U_c^2 scaling is in accordance with the results of Flanagan and Cherry (1969)[38]. They used correlation developed by Meyer-Eppler (1953), reviewed in [37], who studied the sounds produced by flows in constricted pipes. In Mayer-Eppler's scaling the source impedance has to be modelled. In this theoretical work it was done by inverting the wave equation with Green's function. [37]

With low velocity the flow is laminar, but when the velocity increases the turbulence also increases until the flow becomes fully turbulent in high velocity. The amount of turbulence in the flow can be expressed with the help of the Reynolds number: [36], [34]

$$Re = \frac{\rho u d}{\eta}, \tag{3.9}$$

where u is the average velocity through the cross-section of the tube, d is the diameter of the tube, ρ is the density of the air and η is the viscosity of the air. Mayer-Eppler has shown with a plastic tube model of the vocal tract that turbulence was generated and it was associated with sound. The turbulence was created and the sound observed when the critical Reynolds number was achieved due to an increase in flow velocity. [34]

The effects of a finite-length tract Krane has modelled the vocal tract as a closed-open tube, with the closed end at the constriction where the jet is formed and the open end is at the lips. The constriction behaves like a closed end, but the lungs behave like an open end, because of the large cross-sectional area of alveolar tissue. Green's function may be constructed for a finite-length closed-open pipe(not presented in this work, see Krane [37]). The reflections of the tube ends are represented as virtual sources outside the pipe. This is expressed schematically in figure [37]. 3.6. The sound at the open end is [37]
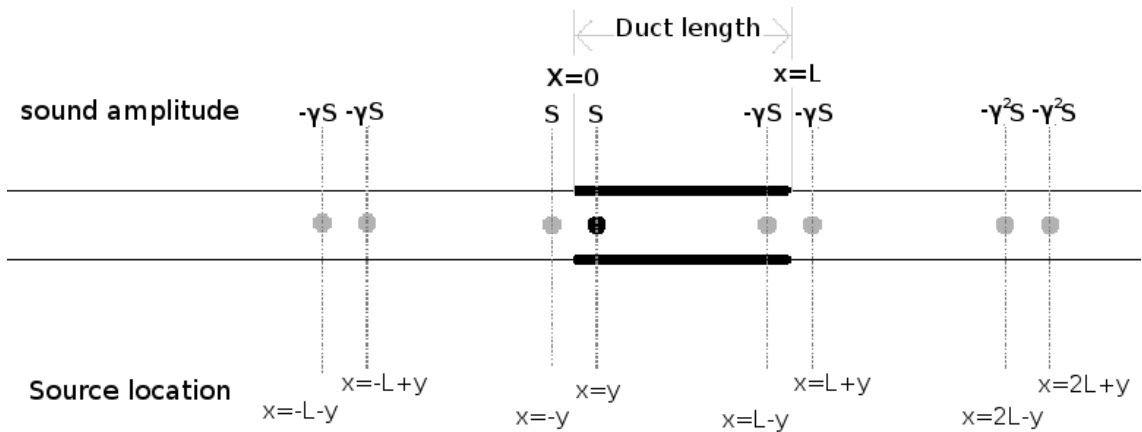
Figure 3.6: The effects of the finity tract. The real sound source S is located at $x = y$ inside the tract. The virtual sources which are images of the real source are distributed into an infinite pipe outside a real tract. The total acoustic effect is the superposition of these virtual sources. In this figure there are only a few virtual sources, but as in equation 3.10 the amount of virtual sources is infinite. Modified from [37].

$$p'(x,t) = -\frac{\rho_\infty}{2A(1+M)} \sum_{n=-\infty}^{\infty} \gamma^n \times \int_{-\infty}^{\infty} \int_A [\omega \times \mathbf{v} \cdot \mathbf{U}*]_n \, dA dy, \qquad (3.10)$$

where the expression in square brackets is evaluated at location $x = x \pm 2nL_f$ and the restarted time is $t - (x \pm 2nL_f)/(c(1+M))$. Even if the source space y is defined over infinity, the observer space x is defined only in the range $0 < x < L$. As can be seen, the effect of the finite tract length is equal to a situation where time synchronous virtual sources are placed outside the tract domain. These additional sources cause quarter wave odd resonance frequencies [37]

$$f_n = \frac{c(2n+1)}{4L_f}, \qquad (3.11)$$

where $n = 0, 1, 2, ...$ and L_f is the axial length of the front cavity i.e. the length from the constriction to the teeth in the case of a vocal tract. If the source is not located either in the inlet or outlet of the cavity the zeros appear in the spectrum at approximately [37]

$$f_z = \frac{nc}{2L_s}, \qquad (3.12)$$

where $n = 0, 1, 2, ...$ and L_s is the distance from the jet origin and the source location. The pressure reflection coefficient γ determines the width of the resonance peaks. The greater the γ the greater the amount of the acoustic energy that radiates out from the open end per round trip. This means that the resonance peaks become wider and the resonant levels become lower. [37]

The aerodynamic source of unvoiced speech sounds is produced by convection of

turbulent air through the change of the tract shape. The sharper the shape change the more noise is produced and the higher the frequency at which the sound is radiated. For studying the generic properties of the source spectrum, a tube with a radius R_p and the constriction with an axial length H and a minimum radius R_{min} is studied (see figure 3.3). If assumed that the acoustic excitation is concentrated in a point location, the sound pressure in an infinite tube becomes [37]

$$
\begin{aligned}
p'(x,t) &= -\frac{\rho_\infty \operatorname{sgn}(x-y)}{2A(1+M)} \int_{-\infty}^{+\infty} \int_A [\omega] \times \mathbf{v} \cdot \mathbf{U}* \, dA dy \\
&= \frac{\operatorname{sgn}(x-y)}{A} S\left(y, t - \frac{|x-y|}{c(1-M)}\right),
\end{aligned}
\tag{3.13}
$$

where S is the source strength located at $x = y$. If a single vortex pair is passing through the constriction the source strength has the same form as in Fig. 3.5(b). When a train of vortexes arrive in an arbitrary time the source function $S(y,t)$ can be expressed as a convolution of a waveform of a single vortex $W(t)$ with an arrival function $I(t)$. The waveform function is scaled by source amplitude for each vortex and the arrival function is a series of delta functions, the phase of which is adjusted to the arrival time of the vortex. Hence, the source function is [37]

$$
S(y,t) = W(t) * I(t) = \int_{-\infty}^{+\infty} W(\tau) I(t - \tau) d\tau.
\tag{3.14}
$$

The effect of convolution is demonstrated in figure 3.7. In the upper figure there is a

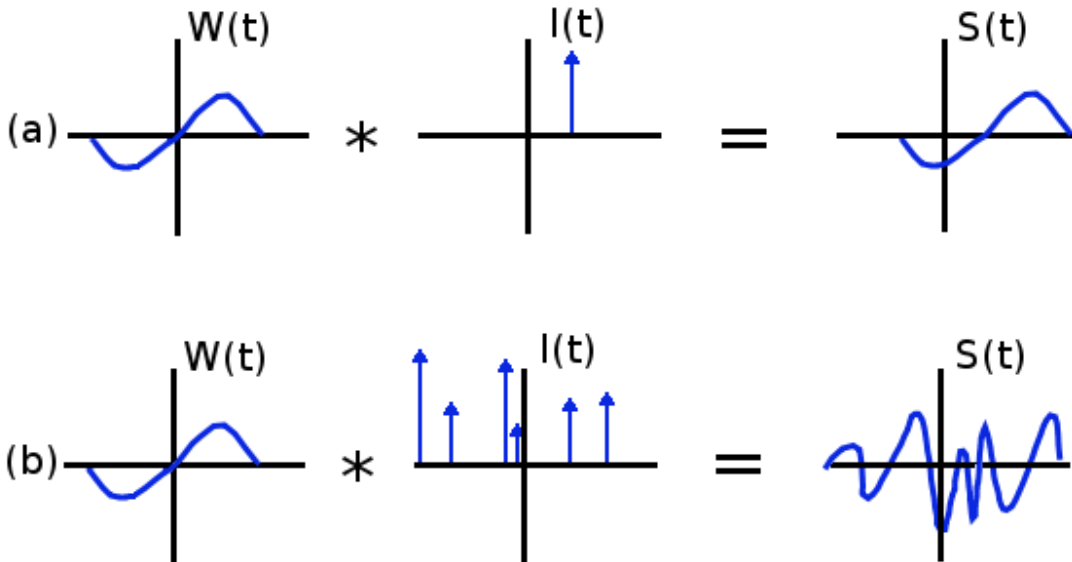
Figure 3.7: The aeroacoustic source spectrum expressed as a convolution of $W(t)$ and $I(t)$, where $W(t)$ is the unit-circulation vortex, reflecting the shape of the duct and the path of the vortex and $I(t)$ is the circulation-weighted arrival function of vortices. (a) The effects of a single vortex and (b) the effects of a train of vortices. Modified from [37].

single vortex ring and in the lower there are six vortex rings. The vortex arrival jitter may be considered as a random process of zero mean and a variance σ_ϕ which is a function of the jets Reynolds number and the distance from the jet origin. The larger the σ_ϕ the shorter the coherence time scale of the $I(t)$ time series. In figure 3.8 the
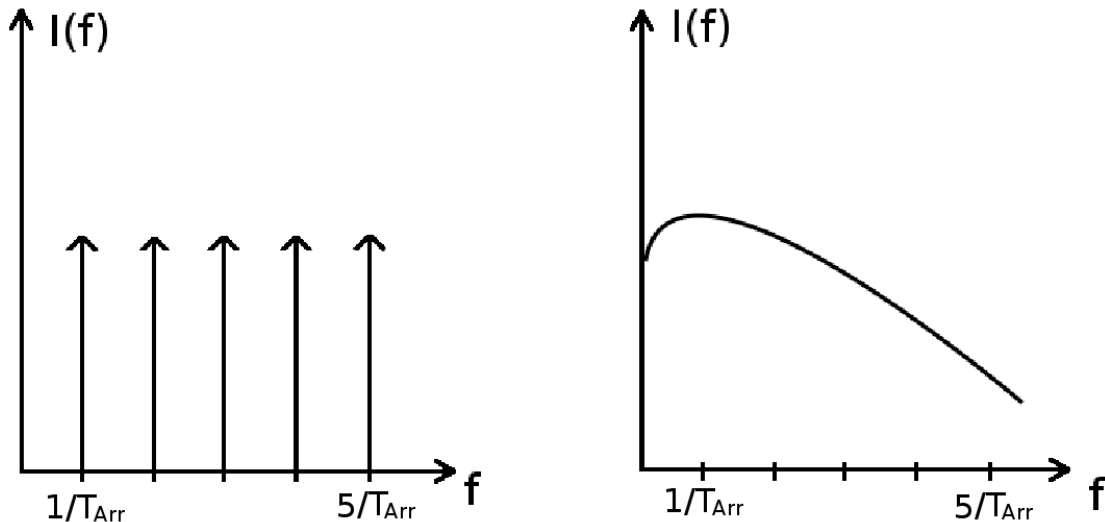

Figure 3.8: The extremes of the vortex arrival function $I(f)$. (a) Situation where the flow is highly periodic and $\sigma_\Phi << 1$ as found in whistling. The spectrum is composed of fundamental frequency and its harmonics. (b)Situation where the flow is highly turbulent and $\sigma_\Phi >> 1$. This is a typical situation of breath sounds. Modified from [37]

spectrum of $I(f)$ is presented in two cases. In figure 3.8(a) shows the case $\sigma_\phi = 0$, meaning that $I(f)$ is periodic and the spectrum is a series of sharp peaks (Dirac delta function) occuring at harmonics of the vortex arrival frequency $f_{arr} = 1/T_{arr}$, where T_{arr} is the mean period of the vortex arrival in the source region. Figure 3.8(b) shows the case where σ_ϕ increases and the spectral peaks become wider until the spectrum becomes broadband with a peak at the mean vortex passage frequency. Because the source function $S(t)$ is the convolution of W and I the spectrum of S is the product of $W(f)$ and $I(f)$ [37] in frequency space

$$S(f) = W(f)I(f) \tag{3.15}$$

It can be thought that the vortex arrival spectrum filters the contribution of a single vortex ring. In figure 3.9 the character of the source spectrum $S(f)$ is presented in two different cases. The source spectrum is dominated by $W(f)$ or $I(f)$ depending on which one has a narrower spectral peak. The unvoiced speech sounds correspond more closely to the case in figure 3.9(b). This is the case with breath sounds, when the variance of the vortex arrival is great. Figure 3.9(a) presents a case where strong flow-acoustic interaction such as whistling occurs. The vortex shedding and the resonant sound field become phase locked and the highly periodic jet vortex
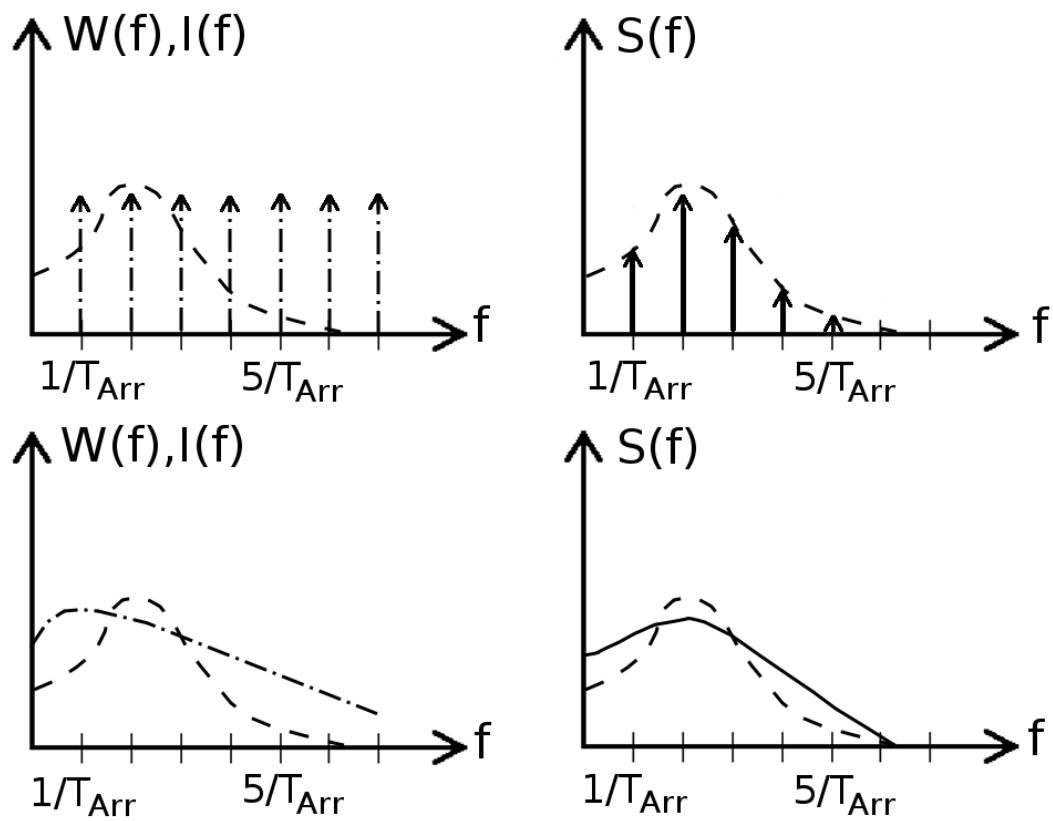
Figure 3.9: The effects of the vortex arrival spectrum $I(f)$ on the sound spectrum $S(f)$. The spectrum of the single vortex $W(f)$ is filtered by the arrival function $I(f)$ and therefore the arrival function determines the character of the source function $S(f)$. $S(f)$-----; $I(f) - \cdot - \cdot - \cdot -$; $W(f) - - - -$.(a)$\sigma_\Phi \ll 1$ (b)$\sigma_\Phi \gg 1$. Modified from [37]

structure is produced. [37]

3.1.2 Turbulent pressure fluctuations

Irregular pressure fluctuations in time and space are present in turbulent flow. Fluctuations propagate at the speed of flow, not at the speed of sound. The sound amplitude is proportional to the pressure fluctuations in the turbulent flow. The pressure fluctuations are related also to the mean pressure drop on the fluid. Only a part of the total energy of the turbulent flow is radiated as true sound propagating in the air. The pressure drop in the fluid is [36]

$$\Delta p \approx \frac{0.241 L \rho^{0.75} \mu^{0.25} F^{1.75}}{d^{4.75}}, \tag{3.16}$$

where Δp is the pressure drop, L the length of the tube, ρ the density of the fluid, μ the viscosity of the fluid, d the diameter of the tube and F the flow rate. The propagation speed of turbulent eddies along the trachea is almost the same as gas particles. For flows of 1-2.5l/s and a tracheal cross section of 5cm^2 the speed is approximately 200-500cm/s which is much slower than the speed of sound which is 34000cm/s. [36]

The sound detected on the surface of the trachea is a vibrational movement, not a real free field sound. The pressure fluctuation p' inside a trachea causes the walls of the trachea to vibrate and this vibration is transferred to the surface. The pressure fluctuation has two components. The dynamic pressure change p'_{turb} due to turbulent eddies on the wall and the acoustic pressure fluctuations p'_{ac} that propagate from near and far (acoustic sound wave). The pressure is a scalar quantity and therefore the following applies [36]

$$(p')^2 = (p'_{turb})^2 + (p'_{ac})^2. \tag{3.17}$$

The equation 3.21 determines the fundamental frequency. The spectral pattern of p'_{ac} is as follows: [36]

$$p'_{ac} = \sum_i \frac{p_0/m}{\sqrt{(1/\tau)^2 + (2\pi f)^2[(if_0/f)^2 - 1]^2}}, \tag{3.18}$$

where p_0 is the source amplitude, m the mass of the column of air in the trachea, τ the time constant of oscillation decay and f_0 the fundamental frequency. The mass can be presented as $m = \pi r^2 L \rho$, where r is the radius and L is the length of the trachea. ρ is gas density. The time constant can be presented as $\tau = m/R$, where R is the acoustic resistance of the trachea. The spectral pattern of turbulent eddies have previously been shown to be flat over a wide range of frequencies. Turbulent eddies cause pressure fluctuations in the pipe and these fluctuations are proportional to

flow-induced pressure drops Δp which are defined in equation 3.16. The local eddies p'_{turb} can be expressed as a function of pressure drop $p'_{turb} = k\Delta p$ and by substituting in equation 3.17 [36]

$$(p')^2 = (k\Delta p)^2 + (p'_{ac})^2. \tag{3.19}$$

The effect of p'_{ac} is much more difficult to estimate. All the sound sources inside a trachea and the upper airways may contribute. These sources of sounds are caused by turbulent eddies, jet noise and boundary layer noise. Jet noise and boundary layer noise correlate with flow to the third power and peak frequency of the jet noise increases with higher flow velocity. Therefore these cannot be the primary sources of the externally measured TBS. However, they dominate if intra-airway measurements are done. [36]

It has been suggested that there is a relation between flow rate and sound amplitude [36, 39, 40]

$$A = \kappa F^\alpha, \tag{3.20}$$

where A is the sound amplitude, F is the flow rate and κ and α are constants.

3.2 Measurements of tracheal breath sounds

In 2005 Raphael Beck[36] and coworkers wrote an article where normal TBS were measured and compared to a theoretical model. The spectrum of the normal tracheal breath sounds has characteristics of broad-band noise and several shallow resonance peaks. The acoustic parameters of the TBS can be obtained from the spectrum and from the overall amplitude of sounds. The spectrum shows the number and position of the resonance peaks and the low-pass cutoff frequency. The source of the tracheal breath sound is a flow in the trachea. A simplified physical model of the sound generation is a flow in a tube which simulates the trachea. The measurable sound field is a combination of several acoustic phenomena, of which the most important are the sounds of turbulent flow, the longitudinal standing waves in the trachea and the sound propagation trough the tissue. [36]

Sound amplitude and flow rate   Beck et al. [36] measured eight subjects with piezoelectric contact microphones. A sensor was attached to the anterior cervical triangle with an elastic rubber belt. The signal was bandpass filtered in the range of 75-2000Hz and the sampling rate of the signal was 4800Hz. The spectrum was calculated using the FFT method and the Hanning window. The subjects breathed through a pneumotachometer. The effects of breathing gas were also studied with normal air and with a Heliox gas mixture. The Heliox is an 80-20% mixture of He-O_2 and it was administered through the Rudolp valve and care was taken to equilibrate the lungs with test gas prior to data collection. Data was collected with the flow

rates 1.0l/s, 1.5l/s, 2.0l/s and 2.5l/s. Background noise was measured during the breath hold. If the flow rate exceeded the target value by 15% or more the data was rejected. The data from expiration and inspiration were analysed separately. The area below the spectral curve was calculated by Beck for each flow rate. These values were normalized using the mean spectral amplitude of the four flow rates. The obtained values are presented in figure 3.10 where the effect of the flow is clearly visible. [36] The curve defined by equation 3.20 was fitted to the data. The parameters for inspiration were $\kappa = 0.29$ and $\alpha = 1.89$ and for expiration they were $\kappa = 0.36$ and $\alpha = 1.59$. [36] For comparison Olson et al. have made a study with a physical model of the trachea with the glottis, showing that $\alpha = 3.0 \pm 0.2$ when κ was kept constant. [36] At the lowest flow rate 1.0l/s Beck estimated the Reynolds numbers to be 3130 and 1130 for air and Heliox, respectively. At the flow rate of 2.5l/s the Reynolds numbers were 7825 for air and 2825 for Heliox. This means that transitional turbulent flow conditions exist for all study conditions.[36] The effects of flow rate is presented in figure 3.11. The amplitude of the TBS increased throughout the frequency range as the flow increased. This was observed during inspiration and expiration with both air and Heliox(80-20%). The flow rate had no effect on the location of the spectral peak. [36]

Harper et al. [34] measured tracheal sounds from four subjects. A contact type transducer was attached at the level of the cricoid cartilage, slightly lateral to it. The flow rate was measured with a pneumotachograph. All of the measurements were done in a sound proof chamber. The target flow rates were 0.5l/s, 1.0l/s, 1.5l/s and 2.0l/s with a tolerance of 20% each. The sampling rate was 10.240kHz. The results from one of the subjects is shown in figure 3.12. A spectrum in the range of 200Hz to 2500Hz and sound power over the 300Hz to 600Hz octave band as a function of flow were presented. It was noted that with every subject the spectral power was greater than the background noise level even at flow levels as low as 0.5l/s. The spectral power increased with the flow and the tracheal sound power of expiration was greater than or equal to the power of inspiration (see fig. 3.12). [34]

Harper et al. [32] studied the breath sounds of four subjects. The Russian vowel sound /a/ was recorded from one inch over the subjects mouth using a Sony TCD-D7 DAT recorder. The sampling frequency was 11kHz and a 16bit A/D conversion was used. The formants were extracted from the data. The /a/ vowel was chosen because the vocal tract has the same acoustical properties when pronouncing it as during breathing. Next, a contact type transducer was attached to the suprasternal notch and the same /a/ vowel recordings than with formant extraction were done. Finally, the subjects breathed quietly and tracheal sounds were recorded from the suprasternal notch with the same PPG #01 contact sensor as used before. Resonance frequencies were determined from the measurements and they were compared
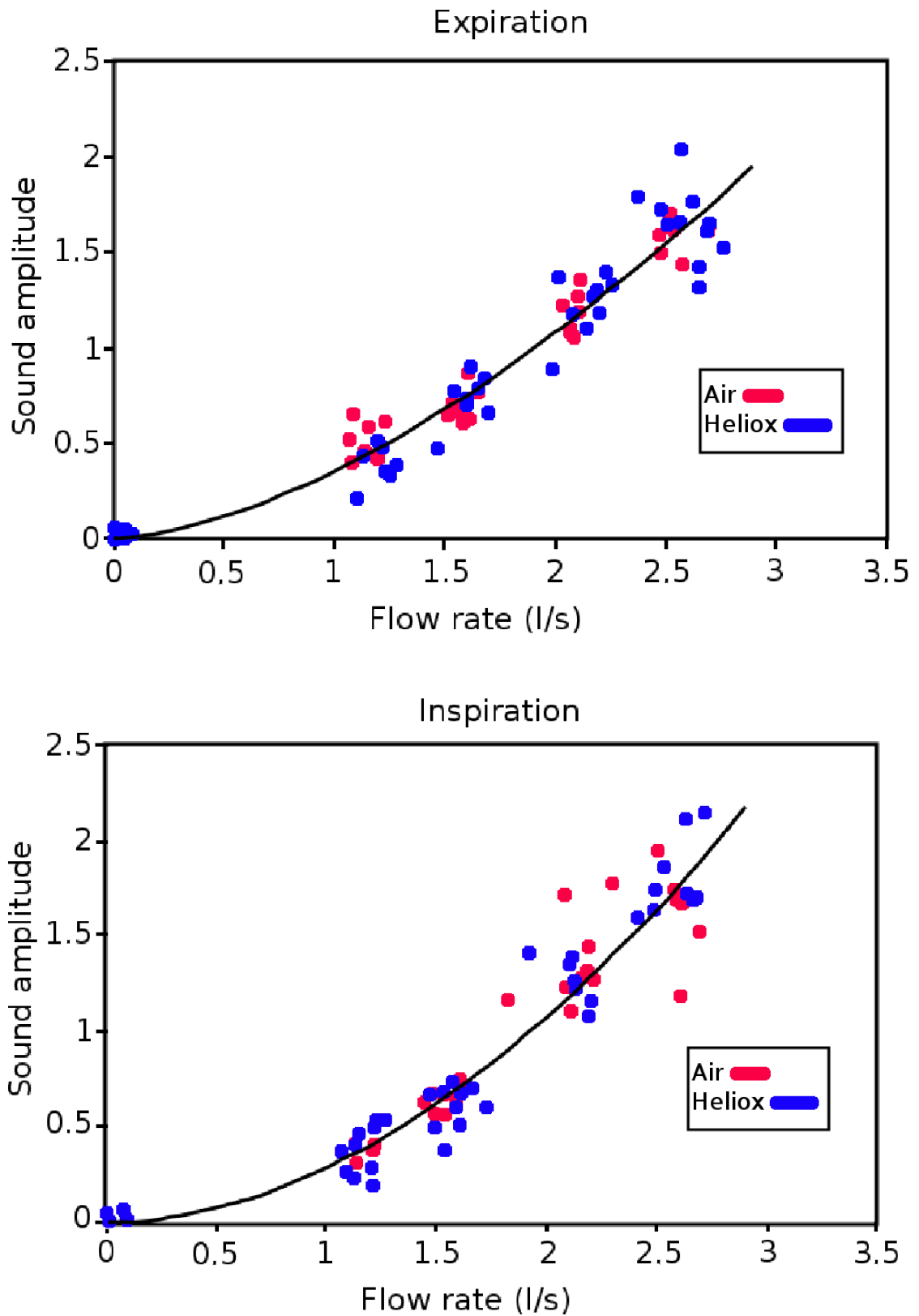
Figure 3.10: Tracheal sound amplitude as a function of flow rate. The amplitude was calculated from the area below the spectral curve. (A) presents the results from inspiration and (B) from expiration.
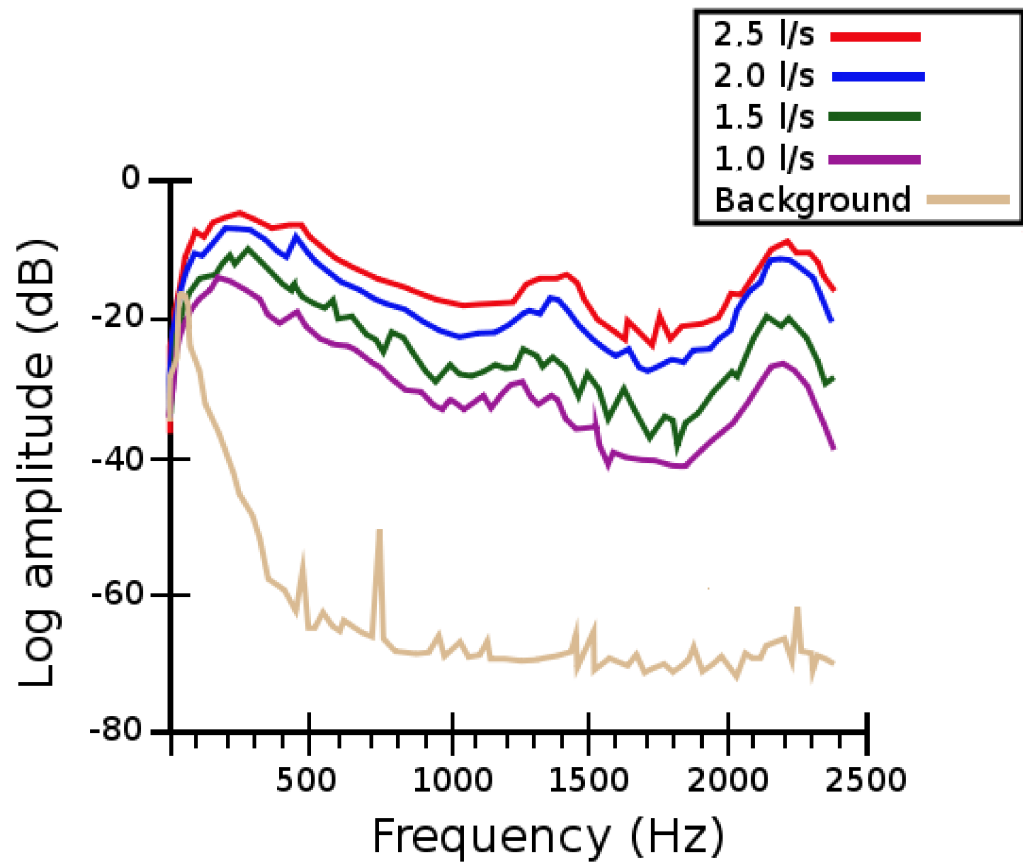
Figure 3.11: Inspiration of air. In addition, plots of inspiration and expiration were presented in the case when Heliox(80-20%) was breathed. The effects of four different flow rates on sound amplitude were observed in every spectrum. Modified from [36]
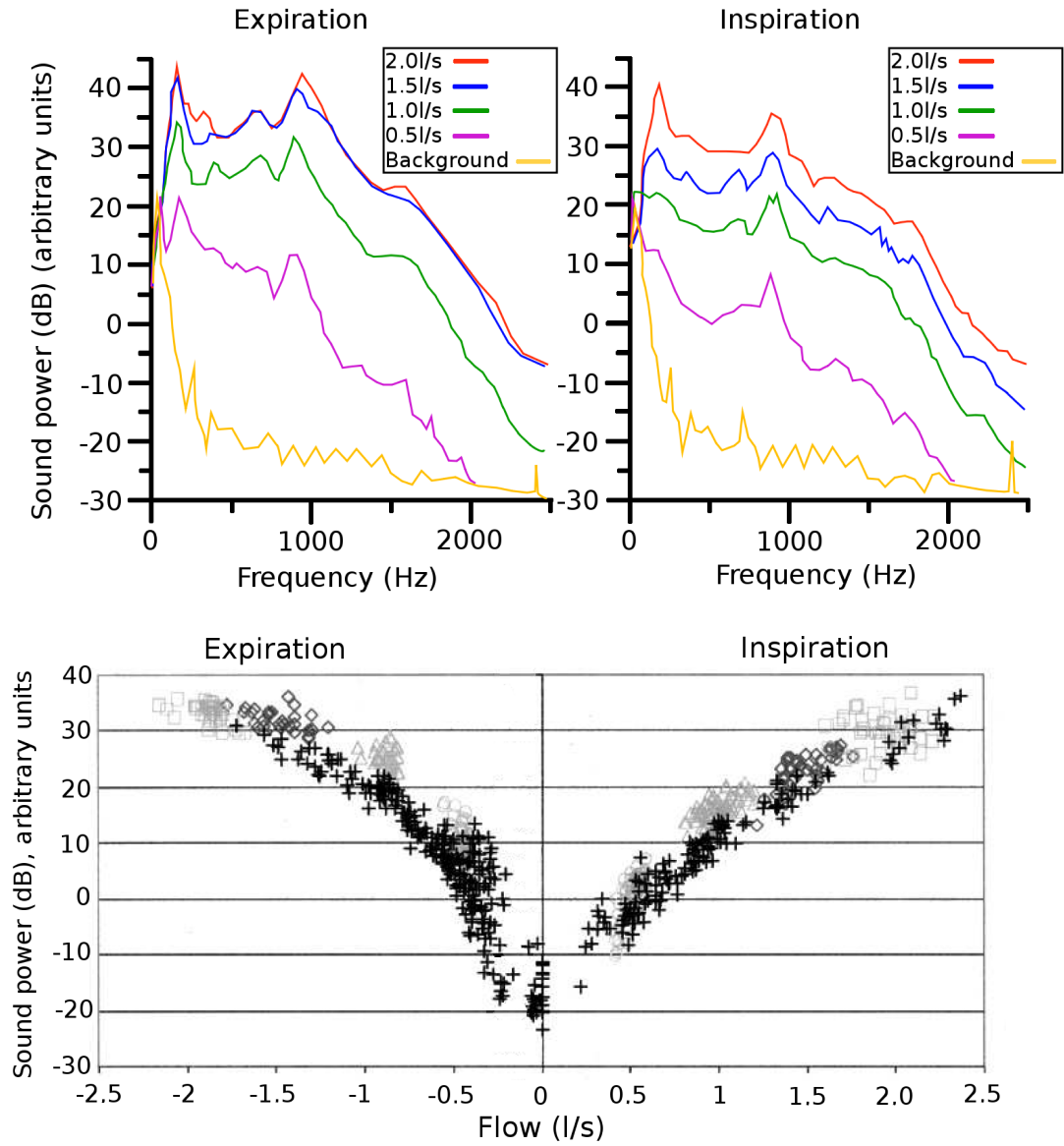
Figure 3.12: Spectra and amplitude of inspiration and expiration. Modified from [34]

with model predictions. The resonant frequencies of each measurement are presented later with simulations in section 3.3. [34]

Beck [36] presents that the increase of sound amplitude with higher flow rates is not supported by Harper et al. [34] Beck [36] states that there is a smaller incremental increase in power at higher flow rates and that Harper [34] explained that it was due to a decrease in the size of the glottal aperture. Beck claims that the model of turbulent flow explains 10dB of this variation and that Harper's model does not correlate with their measurements. On the contrary, the results of the measurements made by Beck and Harper do concord. Beck ignored the fact that Harper presented the sound power in decibel units whereas Beck used a linear scale (compare figures 3.10, 3.12, 3.24 and 3.25). Because the decibel is a logarithmic unit the y-axis of Harper is also logarithmic and so the curves present exponential growth in sound amplitude as the flow increases.

Effects of breathed gas The breathing of Heliox(80-20%) shifted the spectral peaks to higher frequencies as Pasterkamp et al. have demonstrated. [36, 31, 41] In figure 3.13 (adopted from Pasterkamp [41]), the amplitude difference and spectral shifting caused by Heliox(80-20%) are clearly demonstrated.

The sound amplitude was lower during Heliox(80-20%) breathing than during air breathing at all flow rates. [36, 41] The amplitude ratios with different flow rates during Heliox and air breathing are presented in table 3.1. [36]

Table 3.1: The average Heliox(80-20%)-air ratios with different flow rates. The effects of amplitude reduction with Heliox is clear with every flow rate. Modified from [36]

Flow rate (l/s)	Average Heliox-air ratio	Standard deviation
Expiration		
1	0.33	0.09
1.5	0.48	0.20
2	0.48	0.17
2.5	0.49	0.12
Average	0.44	0.16
Inspiration		
1	0.29	0.14
1.5	0.29	0.10
2	0.33	0.09
2.5	0.39	0.13
Average	0.33	0.12
Background		
0	1.21	0.54

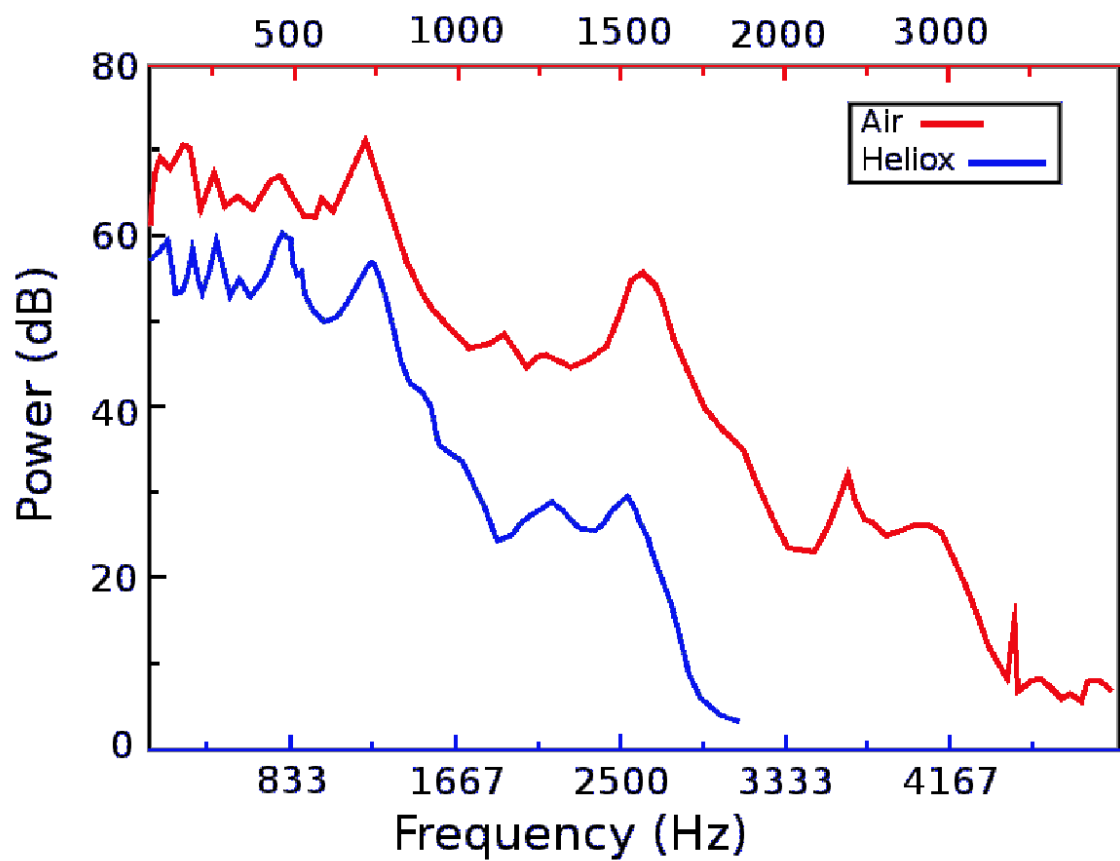The overall mean ratios were 0.33 ± 0.12 and 0.44 ± 0.16 during inspiration and

Figure 3.13: The effects of breathing gas on a sound power spectrum. The upper frequency axis is for breathing air and the lower one is for breathing Heliox(80-20%). Only inspiratory breaths were recorded. Modified from [41].

expiration, respectively. They are not statistically different from each other and they are similar to the density ratio of Heliox and air ($\rho_{heliox}/\rho_{air} = 0.385$). The spectral peaks were shifted from $1280 \pm 150Hz$ to $2250 \pm 150Hz$ when Heliox(80-20%) was breathed instead of air. The ratio of the frequencies was 1.76 ± 0.2 [36], which is slightly higher than 1.47 ± 0.1 observed by Pasterkamp et al. [31]. The increase in frequency can be calculated from the acoustic wave speed ratio. For inspiratory gas the ratio is $\sqrt{\frac{v_{air}}{v_{heliox}}} = 1.8$ and for expiratory gas the ratio is 1.76 [36] when concentrations of water vapour and carbon dioxide are taken into account. [36]

3.3   Modelling the respiratory tract

An acoustic model is needed for explaining physiological and ultimately pathophysiological phenomena observed in breathing sounds. This theoretical modelling of the respiratory tract has been made by Harper et al. [32, 34, 36] The models are used to predict the locations of the spectral peaks in the TBS spectrum. Harper et al. [32] hoped that their modelling could provide a foundation for more complex modelling approaches needed to explain flow-induced sound generation mechanisms in pathologies that interact with the acoustic resonance properties of the respiratory tract. [32]

3.3.1   The single tube model of the respiratory tract

Resonances of the respiratory tract   The resonance peaks of the TBS spectrum are caused by wave reflections from opposing surfaces and airway widenings of the tube. The fundamental frequency f_0 of an acoustic wave is affected by the length of the tube and by the speed of the sound. It can be expressed as [36]

$$f_0 = \frac{v}{2\pi L} = \frac{\sqrt{B/\rho}}{2\pi L} = \frac{\sqrt{(\gamma RT)/M}}{2\pi L}, \tag{3.21}$$

where γ is the adiabatic compression constant, R the ideal gas constant, T the absolute temperature and M the molecular mass. If assumed that an adult tracheal length is $L = 12cm$, the speed of the sound $v = 35800cm/s$ and gas density $\rho = 0.001114g/cm^3$ then equation 3.21 gives $f_0 = 474Hz$. [36]

Sanchez and Pasterkamp have demonstrated that the frequency of the first resonance peak in the tracheal sound spectrum is inversely related to the subject's height.[35, 41] They estimated the length of the subject's trachea from the subject's height and the resonance frequency was determined from the tracheal length. The results of Sanchez and Pasterkamp are presented in table 3.3.1 in columns 1-3. [36]

The resonance values calculated from equation 3.21 are in the fourth column of table 3.3.1. For the calculations, only one half of the wavelength oscillation was

Table 3.2: Data from Sanchez and Pasterkamp. Column 4 shows the predictions made by Beck [36]. Adopted from [36].

Height (cm)	Length (cm)	Measured f_0 by Sanchez[35] and Pasterkamp (Hz)	Predicted [36] $f_0{}^2$ (Hz)
176	12.6	727	872
166	11.5	844	959
149	10.3	986	1068
127	8.4	1158	1307

assumed to exist in the trachea. Equation 3.21 shows that the resonance frequencies are proportional to the speed of the sound in the trachea and inversely proportional to the gas density and to the tracheal length. Beck et al. state that their measurements demonstrated the density dependence of resonance peaks and that the equation 3.17 is a valid model because it predicts the resonances of the work of Sanchez and Pasterkamp. [36, 35]

The predicted spectrum of the equation 3.18 is presented in figure 3.14. [36] For
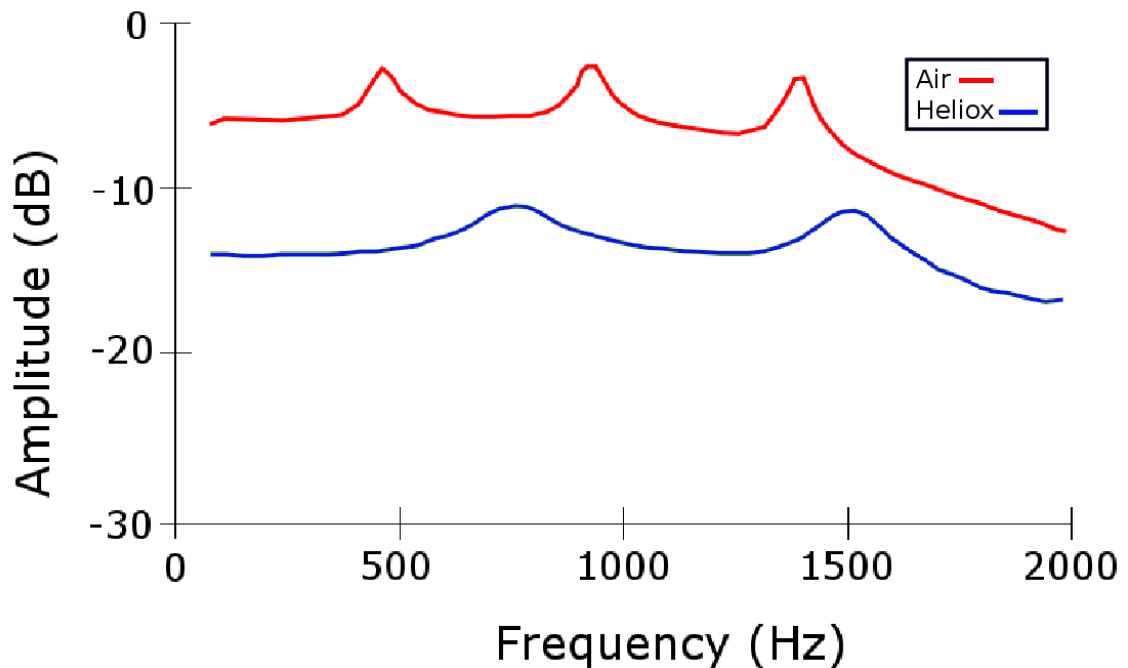
Figure 3.14: The predicted spectrum of tracheal breath sounds. Figure modified from [36]

the calculations the following assumptions were made: the tracheal length was 12cm, the diameter of the trachea 1.5cm, the gas densities 0.00043g/ml and 0.00114g/ml for Heliox(80-20%) and air, respectively. The time constant for Heliox(80-20%) was 0.83ms and for air 2.42ms. The transmission characteristics have not been addressed in this study either. The measured TBS are influenced by transmission parameters

of the tracheal wall and by overlying soft tissue and transducer coupling to the skin. The TBS generation has several parameters that have not yet been tested: the effects of ambient pressure, the changes of the area-distance curve, e.g. flexion or extension of the neck and simultaneous intra-airway and neck surface measurements.[36]

3.3.2 The lumped acoustic element model of the respiratory tract

The lumped acoustic element model is analogous to the electrical circuit model and this same approach has been used in speech research when acoustical processes in the vocal tract have been modelled. Acoustical transmission and reflections of the branching structures have been included in this model by Harper et al. [32]. The model combines two different parts of the respiratory tract; the supraglottal and subglottal parts.

In the modelling of the supraglottal or vocal tract portion, the previous models used in speech research were used and these predicted natural resonances were compared with measurements of speech made at the mouth. The subglottal portion was simulated with a branching tract and the simulated results were compared with speech measurements from the tracheal site. The final model coupled both the subglottal and supraglottal portions. The analogous electric circuit model uses the lumped acoustic element approach for the T-section. This method is introduced in chapter 2 where breath sound sensors were modelled. When the acoustic parameters of a tube-like environment have been simulated, the studied frequency range must be fixed. When the highest frequency is much greater than the largest cross sectional dimension of the tube-like structure of the modelled system, the sound will propagate only as plane waves in the tube. The first non-planar propagation mode is observed at the "cuton" frequency. For the acoustic wave in the rigid cylindrical tube this cuton frequency can be expressed as [32]

$$f_{1,1} = \frac{1}{2\pi} \frac{1.84v}{r},$$
(3.22)

where v is the speed of sound and r is the radius of the tube. If assumed that the speed of sound in moist air is 35400cm/s and the maximum radius of the tube is 2.0cm, then only the plane waves exist up to 5000Hz. In the lumped acoustic element approach the length of a single element must be determined. Normally it is chosen to be less than 1/8th of the shortest wavelength and in this model it was chosen to be 1.5cm[32]. The respiratory tract wall yields in response to applied acoustic pressure. Therefore, the model of the rigid wall tube will not give a correct result for the respiratory tract. [32]

Acoustic circuit   The lumped acoustic circuit element for a tube with a rigid wall is shown in figure 3.15. [32] The nonrigid wall case can be modelled by adding a
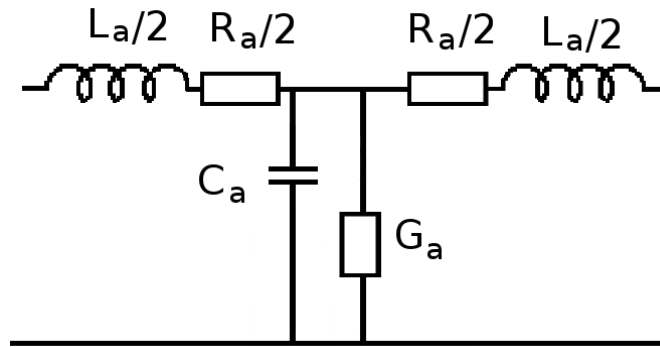

Figure 3.15: A circuit of lumped acoustic elements. The circuit is a lossy tube section with rigid walls in a T network. Modified from [32]

series combination of a resistance (R_w), an inertance (L_w) and a compliance (C_w) to the section with a rigid wall.The non-rigid tube circuit is presented in figure 3.20. The respiratory tract wall has a spatial inhomogeneity because the tracheal wall consists of soft tissue interspersed with cartilage rings. Therefore, both tissue types have their own lumped elements in figure 3.20 marked with the subscript letters c (cartilage) and s (soft tissue). The parameters of the components in the lumped circuit for rigid and non-rigid models are presented in tables 3.3.2 and 3.3.2 respectively.[32]

Table 3.3: Parameters for a lossy rigid wall segment. $r=$ tube radius, $l=$segment length, $\omega=$radian frequency,$\rho=$density of medium, $\eta=$shear viscosity, $A=$cross-sectional area, $c=$speed of sound, $v=$ratio of specific heats, $\kappa=$heat conduction coefficient and $c_p=$specific heat at constant pressure. Adopted from [32].

Parameter	Value	Units
Resistance	$R_a = \frac{2l}{\pi r^3}\sqrt{\frac{\omega\rho_0\eta}{2}}$	$\frac{\text{dyne·s}}{\text{cm}^5}$
Inertance	$L_a = \frac{\rho_0 l}{A}$	$\frac{\text{dyne·s}}{\text{cm}^5}$
Compliance	$C_a = \frac{Al}{\rho_0 c^2}$	$\frac{\text{cm}^5}{\text{dyne}}$
Conductance	$G_a = 2\pi rl\frac{v-1}{\rho_0 c^2}\sqrt{\frac{\kappa\omega}{2c_p\rho_0}}$	$\frac{\text{cm}^5}{\text{dyne·s}}$

The subglottal airways were divided into segments. The segmentation is made in such a way that segments form a tree-like branching network. Terminology from graph theory is adopted by Harper [32] to describe the architecture of the network. The *branch* is an airway segment connecting two nodes, the *node* is a point of intersection of two or more branches and the *depth* of the node is determined by counting the branches of the root nodes. The trachea is the branch with zero depth and the left and right main bronchi are branches with depth one. The maximum

Table 3.4: Additional parameters for a lossy non-rigid wall segment. The subscript x can be replaced with s (=soft tissue) or c (=cartilage). r=tube radius, ω=radian frequency, l= segment length, h=wall thickness, ρ_{wx}=tissue density, η_{wx}=tissue shear viscosity and E_{wx}=tissue elasticity. Adopted from [32].

Parameter	Value	Units
Resistance	$R_{wxt}(\omega) = \frac{\eta_{wx}(\omega)h}{2\pi r^3 l}$	$\frac{\text{dyne·s}}{\text{cm}^5}$
Inertance	$L_{wxt}(\omega) = \frac{\rho_{wx}h}{2\pi r l}$	$\frac{\text{dyne·s}}{\text{cm}^5}$
Compliance	$C_{wx}(\omega) = \frac{2\pi r^3 l}{E_{wx}(\omega)h}$	$\frac{\text{cm}^5}{\text{dyne}}$

node depth was chosen to be 10 because with greater depths the total cross-sectional area is very large and it acts as an open-tube termination. The branching was assumed to be symmetrical in order to reduce the computational complexity of the model. However, the authors note that in previous articles the asymmetries are shown to have an important effect when the frequency increases. The properties of segments are shown in table 3.3.2. [32] The amount of cartilage and soft tissue

Table 3.5: Parameters for subglottal airway segments. Adopted from [32]

Depth	Tube length, l [cm]	Tube radius, r [cm]	Wall thickness, h [cm]	Fraction of cartilage, c_{frac}
0	10.0	0.80	0.3724	0.67
1	5.0	0.6	0.1735	0.500
2	2.2	0.55	0.1348	0.500
3	1.1	0.40	0.0528	0.3300
⋮	⋮	⋮	⋮	⋮
10	0.860	0.175	0.0114	0.0525

changes depending on the anatomical location of the airway segment. The fraction of cartilage that each segment contains is c_{frac} and the fraction of soft tissue is $1 - c_{frac}$. Each circuit parameter is scaled as follows:[32]

$$
\begin{aligned}
R_{wc} &= R_{wct}/c_{frac} \\
R_{ws} &= R_{wst}/(1 - c_{frac}) \\
L_{wc} &= L_{wct}/c_{frac} \\
L_{ws} &= L_{wst}/(1 - c_{frac}) \\
C_{wc} &= C_{wct} \times c_{frac} \\
C_{ws} &= C_{wst} \times (1 - c_{frac})
\end{aligned}
\tag{3.23}
$$

The single volume velocity source was placed between every T-segment for the series simulations. The source was positioned for model excitation, but the locations of the sources were not determined by a real distribution of sound sources in the human body. The purpose of the sources was to excite the resonance properties of the

simulated tract. The simulated measurement location was situated away from the locations of the nodes and the antinodes of standing waves. By doing this the authors avoided nulls in acoustic pressure.[32]

Vocal tract modelling (Supraglottal region)   The authors first assumed that the human subject breathes with the mouth open and that the soft palate blocks the whole nasal cavity. The vocal tract wall was set to be rigid and the cross-sectional area was uniform. By these definitions the authors modelled the tract as a 17-cm-long non-uniform tube that has an open proximal end and a nearly closed distal end due to vocal folds. For that kind of open-closed tube systems the odd quarter-wave harmonic resonances exist according to equation [32]

$$f_n = \frac{(2n+1)\,c}{4L},\tag{3.24}$$

where c is the velocity of the sound and L is the tube length and $n = 0, 1, 2, \ldots$. The results of the lumped acoustic element model with the assumption of a rigid wall and equation 3.24 are in table 3.3.2. It can be seen that the models give

Table 3.6: Predicted resonances of the simple tube model and the acoustic transmission line model. Adopted from [32]

Harmonics	Simple lossless tube model	Acoustic transmission line model
n	Resonances [Hz]	[Hz]
0	520	508
1	1562	1560
2	2603	2560

results with great similarity. The other simulation was performed with non-rigid walls and it was assumed that the subject uttered the Russian vowel sound /a/. It was also assumed that the glottis was closed like in all vocal tract simulations. The modelled spectrum is shown in figure 3.16. The vowel /a/ was chosen because Fant (reviewed in [32]) has measured the same spectrum and Portnoff (reviewed in [32]) has modelled it. The authors also measured the spectrum with a microphone that was one inch from the subject's mouth. The recording device was a Sony TCD-D7 DAT recorder and four subjects in total were measured. The formant frequencies of the authors' simulations and measurements, Portnoff's simulations and Fant's measurements are shown in table 3.3.2.

The authors' simulations were within ±6% of the means of the authors' measurements, within ±10% of Portnoff's simulations and within ±10% of Fant's measurements. The authors state that their predictions are closer to Fant's measurements
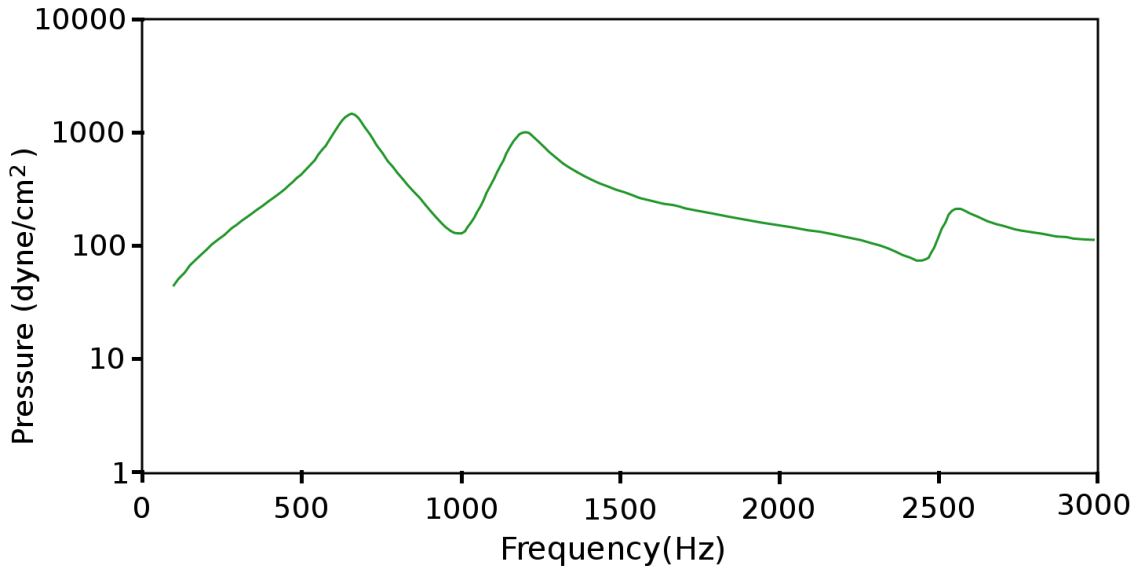
Figure 3.16: The predicted spectrum of the vocal tract. Modified from [32]

.

Table 3.7: Predicted and measured resonances of the vocal tract. Adopted from [32]

	F1[Hz]	F2[Hz]	F3[Hz]
Model	686	1180	2530
Portnoff's[32]	650	1075.7	2463.1
Subject no.			
1	632	1062	2461
2	711	1124	2779
3	660	1115	2464
4	609	1165	1893
mean	653.1	1116.5	2399.2
Fant's[32]	700	1080	2600

than to Portnoff's model.[32] However, Porthnoff's model predicts the authors' measurements better than the authors' own model.

Subglottal modelling : The glottal was assumed to be yielding and the glottis was assumed to be closed like in the supraglottal model. The terminal branches were considered at the depth of ten. The authors suggest that the quarter-wave, an odd harmonic model of the equation 3.24 proved to be an inadequate model even when compared to simulations where only the trachea and the main bronchi were included. The more the branches were included in the simulation, the more the spectral peaks shifted downward in the frequency range compared to the simple open-closed tube model of equation 3.24. Figure 3.17 shows the final simulated spectrum where a branch depth of ten is used. The first spectral peak is at 569 Hz and the second one is at 1360Hz. It is clear that they are not harmonically related.
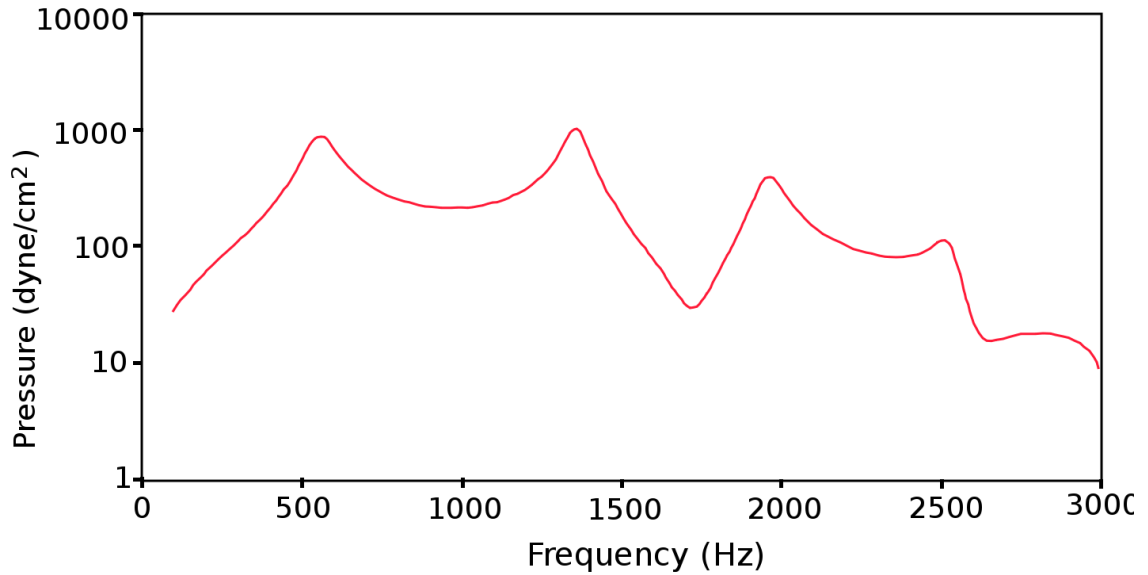
Figure 3.17: Predicted spectrum of the subglottal tract. Modified from [32]

The authors note that this clearly shows the fact that the branching and yielding of the subglottal airways deviate significantly from that of a uniform and rigid tube. The formants of the sound /a/ was modelled and compared to Ishizaka's model [42] and to measurements made by the authors. The sound was recorded with a contact sensor attached to the suprasternal notch. The modelled and measured formant frequencies are presented in table 3.3.2. [32]

.

Table 3.8: Predicted and measured resonances of the subglottal tract.[32]

	F1 [Hz]	F2[Hz]	F3[Hz]	F4[Hz]
Model	569	1360	1980	2500
Ishizaka's[42]	615	1335	2110	2379
Subject no.				
1	606	1279	2256	2919
2	529	1092	1729	2343
3	753	1191	2024	2617
4	699	1318	2228	2748
mean	644.4	1220.0	2656.8	

The simulation results were within $\pm 8\%$ of Ishizaka's predictions and within $\pm 12\%$ of the means of resonances of the authors' measurements.

Respiratory tract modelling   Supraglottal and subglottal models were coupled to create a model of an entire respiratory tract. The degree of coupling was determined by the size of the glottal opening. The modelling parameters were the same as in the modelling of the vocal tract and the subglottal portions. For coupling,

the cross-sectional area was assumed to be 1.6cm^2. The volume velocity source was located 6.0cm above the glottis and the pressure output was 1.5cm below the glottis where no apparent nulls resulted as resonances. The modelled spectrum of the entire respiratory tract is presented in figure 3.18. [32] It can be observed that there
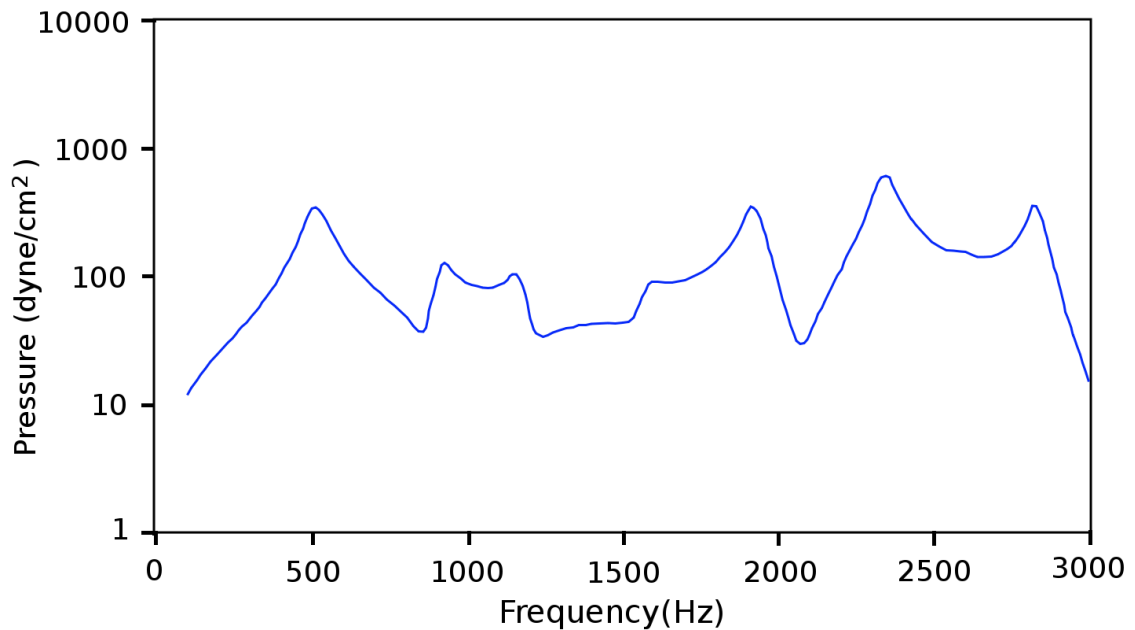

Figure 3.18: Predicted spectrum of the respiratory tract. Modified from [32]

are more spectral peaks compared to either subglottal or supraglottal simulations. A clear harmonic behaviour is not observed. Complex spectral behaviour is a result of many factors such as the variable vocal tract cross-sectional area, non-rigid walls and subglottal airway branching. The respiratory tract model was compared with tracheal breath sound measurements made with contact sensors attached to the suprasternal notch. It was assumed that during quiet breathing the maximum size of the glottal opening corresponds to the modelling conditions. Only expirational sounds were analyzed because it was believed that sound sources are more centrally located compared to inspiration. The results of the measurements and the modelling are presented in table 3.3.2. [32]

The first five predicted resonances were within 11%, 4%, 2% and 7% respectively of the average measured values. The model has a quite high predictive capability as noticed also by the authors. The authors did not discuss the reasons behind the fact that their models had larger errors in subglottal and supraglottal simulations compared to the model of the respiratory tract The authors state that the purpose of the modelling was to model the spectral peak locations. The heterogeneous anatomy of the respiratory tract makes it difficult to relate specific spectral peak locations to individual anatomical structures. The authors state that the anatomical parameters

.

Table 3.9: Predicted and measured resonances of the respiratory tract. Adopted from [32].

	R1 [Hz]	R2[Hz]	R3[Hz]	R4[Hz]	R5[Hz]
Model	510	920	1160	1620	1950
Subject no.					
1	530	848	1232	1497	2080
2	596	927	1298	1682	
3	596	933	1298	1669	
4	568	848	1285	1550	2080
mean	573	889	1279	1600	2080

have a strong impact on the spectral peak location while the sound source location and type are not so important. However, they note that such sources have a key role in the determination of the overall spectral features and amplitude of TBS.[32]

3.3.3 Effects of asymmetric airways and sound source distribution

In a previous study, Harper et al. [32], demonstrated that the computer model of the respiratory tract (3.19) can predict the resonance frequencies accurately. In
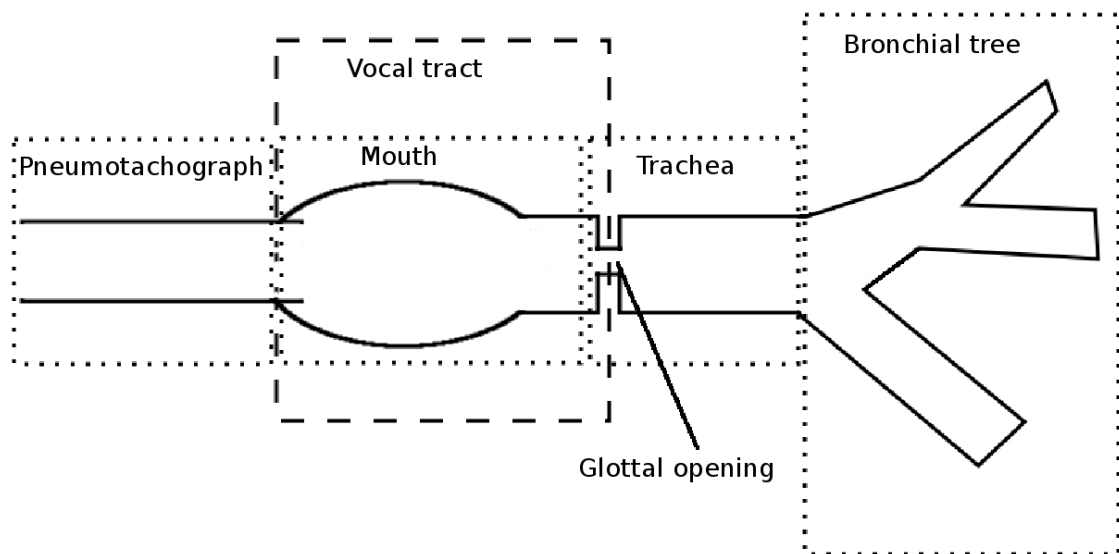
Figure 3.19: A model of the respiratory tract from the mouth to the bronchi including the pneumotachograph. Modified from [34].

this model the average values for the airway geometry were used and the acoustic point sources were included to excite the resonances. The locations of the point sources were not based on the true locations of flow-generated sound sources. In this work the previous model is improved by including flow-generated sound sources

into the respiratory tract. Measurements from healthy subjects are compared to with the model. A symmetrical treatment of the subglottal portion is replaced with the asymmetric model (see Pastekamp et al. [43]. This means that when moving
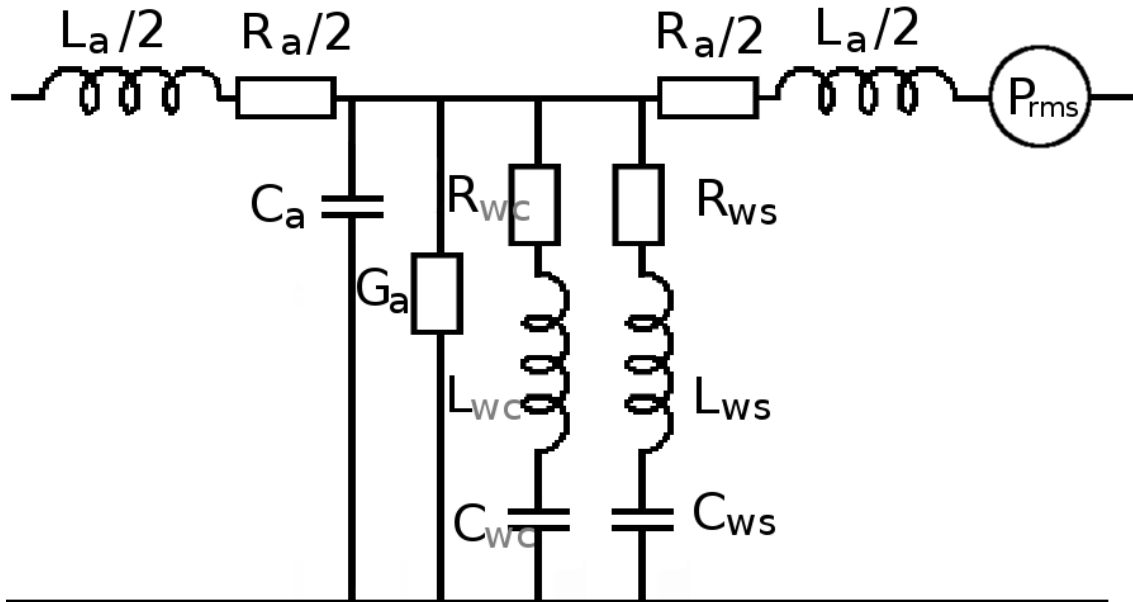

Figure 3.20: A lumped acoustic element model of the T-section with yielding walls and a distributed sound source. Modified from [34].

from the parent branches to the daughter branches, the depth of each daughter increases, but the increase is not symmetrical. This was done so that an airway of order will bifurcate into two daughters; one of the order n+1 and the other of n+1+Δn. This Δn is a recursion index and the values for each parent branch is shown in table 3.3.3. [34] The asymmetric branches down to level 4 are showed in figure 3. The modelling of the turbulent flow-induced acoustic sources is difficult. Therefore, the authors made some simplifications. The two key parameters for the respiratory tract are the cross-sectional area and the volumetric flow rate. The flow rate can be measured from the mouth and its value can be used for every respiratory tract segment before the first bifurcation. At the point of bifurcation the flow rate is divided into daughter branches proportional to their cross-sectional area and in a way that the sum of the flow rates in the daughter branches is the flow rate at the parental branch. [34]

Pressure of the Acoustic source modelling : The root-mean-square (rms) pressure of the sound can be expressed with the Reynold number as follows[34]:

$$P_{rms} = K(Re^2 - Re^2_{crit}), \tag{3.25}$$

Table 3.10: Parameters for subglottal airway segments. Modified from [34].

Depth	Tube length, l[cm]	Tube radius, r[cm]	Wall thickness, h[cm]	Fraction of cartilage, c_{frac}	Recursion Index Δn
0	10.0	0.80	0.3724	0.67	1
1	5.0	0.6	0.1735	0.500	2
2	2.2	0.55	0.1348	0.500	3
3	1.1	0.40	0.0528	0.3300	3
⋮	⋮	⋮	⋮	⋮	
21	0.420	0.070	0.0055	0.000	3
22	0.360	0.055	0.0047	0.000	2
23	0.310	0.048	0.0043	0.000	2
24	0.250	0.038	0.0038	0.000	1
25	0.110	0.032	0.0034	0.000	0
⋮	⋮	⋮	⋮	⋮	
35	0.048	0.040	0.0039	0.000	0
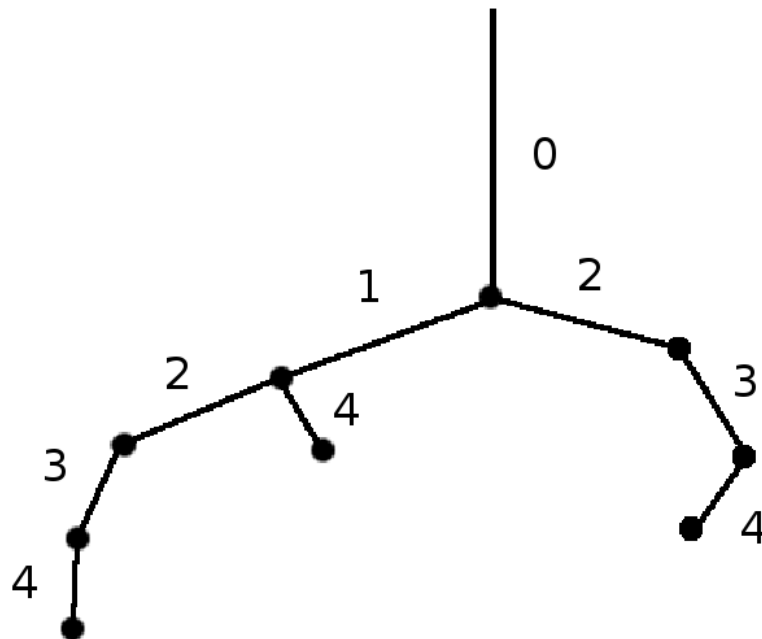
Figure 3.21: A model of asymmetric branching of the subglottal tract down to level four. Modified from [34].

where K is the constant and Re_{crit} is the critical Reynolds number. It is accepted that the critical Reynolds number for the respiratory tract is in the range of 1800 to 2700.[36] Below this range there is no significant sound generation. In the previous modelling of Harper[32] the authors noted that the coupling of the acoustic energy was affected by the location and the type of the sound source. This becomes important if also the amplitude of the sound is modelled and not just the frequencies of the

resonances. Equation 3.25 was used to estimate sound pressure in the respiratory tract, but the sound sources must be spatially distributed if more accuracy to the model is needed. The spatial distribution was applied as shown in figure [34] 3.22. The pressure of equation 3.25 was divided for each T-section so that the constriction

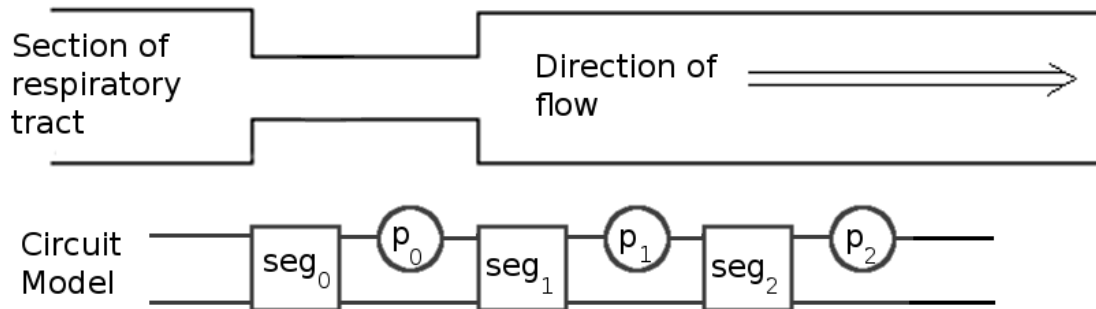Figure 3.22: Distribution of acoustic sources. The value of p_0 is 20% of the pressure calculated by 3.25. The pressure relationship between different sources is p_0:p_1:p_2:p_3:... is 20%:40%:20%:10%:... Modified from [34]

section had 20%, the next section downstream had 40%, then 20%, 10%, 5% and 2.5%, respectively, at each subsequent T-section. [34]

If the spectral parameters of the sound sources are defined with measurements over the trachea, the transfer function from the source to the transducer must be known. The sound is transmitted from a source trough the tracheal wall and the neck tissue. Above a few hundred Hz, the impedance of both tissue types is great and because the frequency response of the transducer used is flat, it can be said that the measured acceleration in the range of 100 to 2000Hz is approximately proportional to the pressure within the trachea.[34] There is no definitive way to model the envelope of the spectrum of the sound sources inside a respiratory tract as a function of cross-sectional area and flow rate. Therefore, a flat white noise source in the range of 100 to 1000Hz was used.When measurements and simulations were compared by the authors, the frequency range was restricted even more from 300Hz to 600Hz, because in that range the spectral distribution is relatively flat.[34]

In figure 3.23 the family of simulated tracheal sound spectra were presented. Average parameters of a male respiratory tract were used in the modelling process. The effects of nasal cavities were excluded, because the subjects had a nose clip in place during the recordings. A pneumotachograph was included in the model and the vocal tract had the Russian /a/ profile which is close to a relaxed vocal tract shape. Subglottal airways were included as anasymmetric branching network with a depth of 14. Tracheal sound spectra were estimated to be between 100Hz and 1000Hz while the sound power was estimated to be between 300Hz and 600Hz. The target flow inspiration measurements are compared with model predictions in
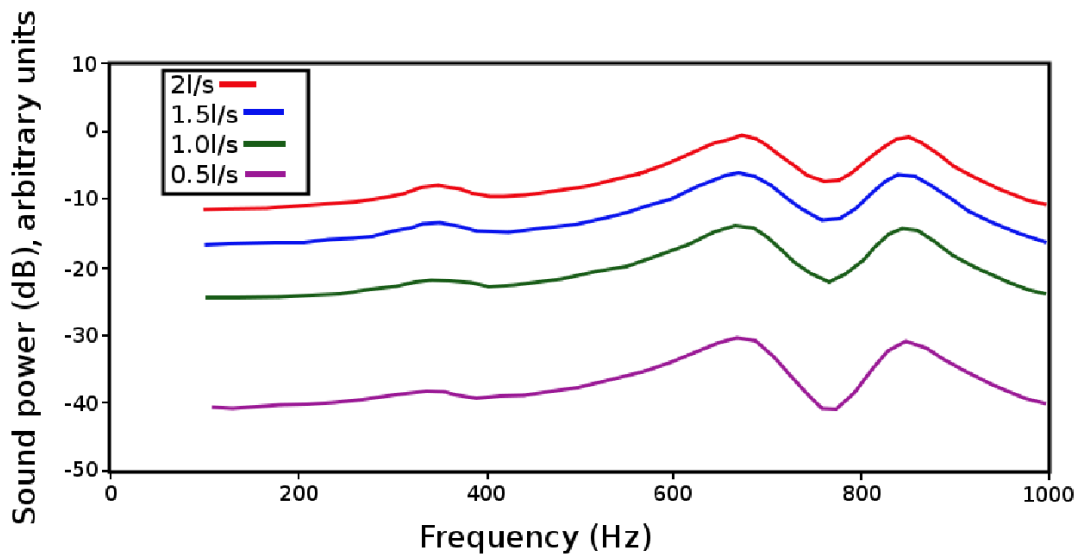
Figure 3.23: The family of simulated spectra with different flow rates. The Russian /a/ vocal profile for the vocal tract, the pneumotachograph extension model, a glottal radius of 0.55cm, spectral shaping of white noise source and the asymmetric subglottal system were applied in the model. Modified from [34].

figure 3.24.[34] Similar results were presented with expiration as well, but with the
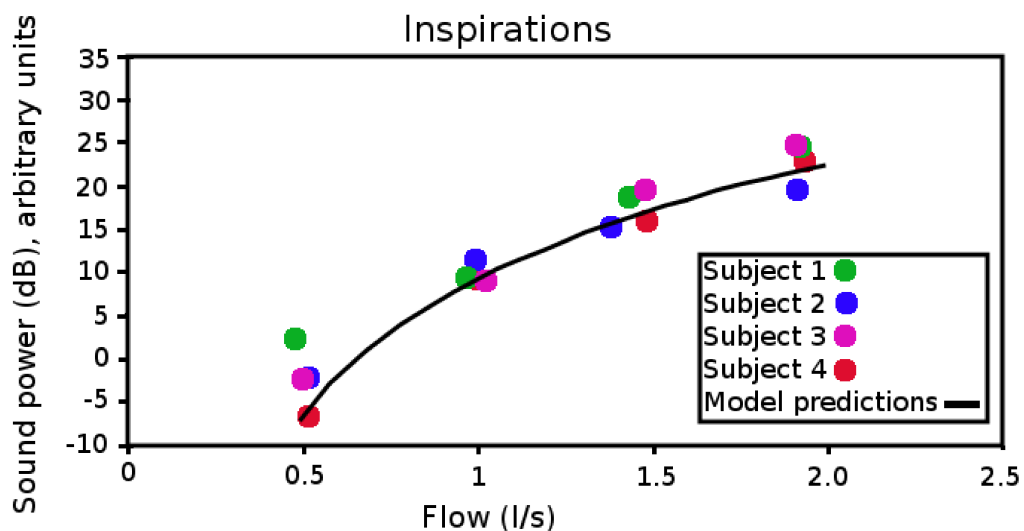
Figure 3.24: A measured and fitted model of sound power during inspiratory target flows. Modified from [34].

high flow rates the difference between the predicted sound power and the measured sound power grew. When sound power simulations were done the glottal radius was fixed to 0.55cm. The data points represent the average sound power over the

300Hz to 600Hz band. The model predictions were scaled to have the same values as was measured from subject one with flows of 1.0l/s and -1.0l/s for inspiration and expiration, respectively. The scaling factor κ in inspiration was 0.022 and in expiration the factor was 0.032. The same scaling was done for individual datapoints. In figure 3.24 the scaling factors for subjects 1-4 were, 0, -6.6dB, 3.0dB and -2.6dB, respectively. For expiration the factors were 0dB, -15.6dB, -2.0dB and -11.9dB, for subjects 1-4, respectively. The maximum deviation of the predictions from the measurement data points were approximately 10dB. Brancatisano et al.[44] observed that during relaxed respiration and with flows less than or equal to 0.5l/s, the glottal aperture varied. The minimum radius of the glottal was achieved during expiration and it was 0.45cm (area 0.64cm^2). The maximum radius of 0.65cm (1.4cm^2) was achieved during inspiration . [34] The effect of aperture change was also simulated
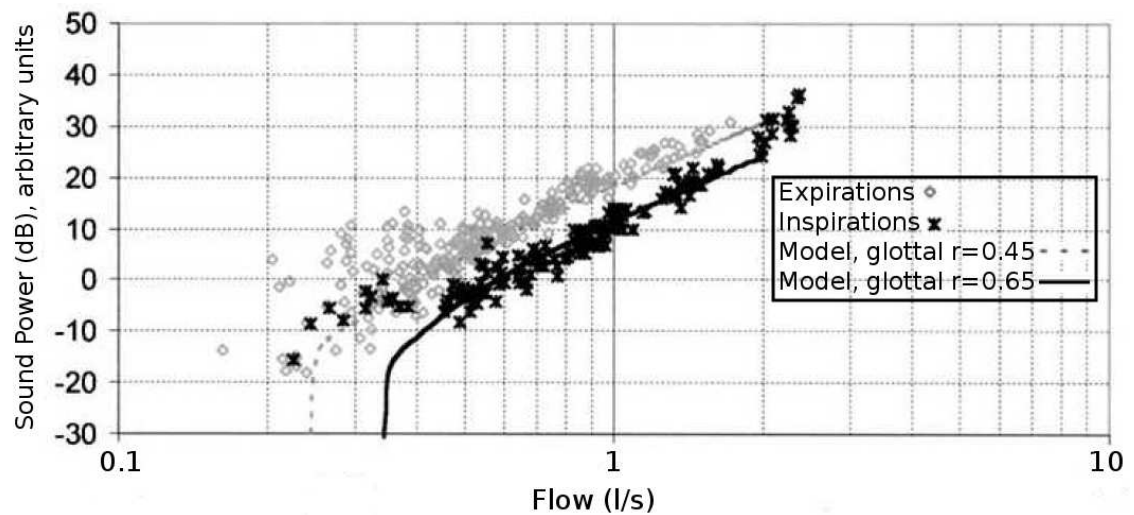
Figure 3.25: Sound power as a function of flow rate. The nontargeted breaths and the fitted model are presented. The minimum and maximum glottal sizes are represented by inspiration and expiration respectively

and figure 3.25 shows the simulation comparison of measured values with minimum and maximum aperture size during inspiration and expiration, respectively. The frequency range used in the simulations was from 300Hz to 600Hz and the curve was scaled with the measurements in the same way as in figure 3.24. The authors note that this scaling does not affect the results as long as it is fixed for every given subject [34]. The scaling was done in order to enable a comparison between the slope of the curves and the relative change in the sound power as a function of flow and glottal opening size. The specific scale factor for a subject was chosen so that the curve was placed in the mid range of the expiration sound power while the glottal opening was 0.45cm. The factors for subjects 1-4 were -8dB,-7dB,-1dB and -4dB, respectively.[34]

The authors state that the spectral power increased during targeted and non-targeted breathing when the flow was increased. The largest increase occurs between 0.5l/s and 1.0l/s. In the simulated spectrum the power increases about 30dB when the flow increases from 0.5l/s to 2.0l/s. In the measurements of the inspiration noises the increases over the 300 to 600Hz octave band were 28dB, 23dB, 30dB and 30dB for four subjects. The measured expiration noises were 23dB, 18dB, 33dB, and 23dB. The most limited spectrum was observed with patient 3 when the flow rate was 0.5l/s. The upper frequency limit was 1500Hz. In general, the spectra are limited by the sensitivity of the transducer and by the spectral roll-off which is characteristic of the acoustic sources in the trachea. Sound power as a function of flow rate is also presented in figure 3.12. The targeted flows appear as clusters of points whereas the nontargeted data points are scattered between and onto the clustered points. An offset between targeted and non-targeted breathing was observed. The offset patterns are similar to each other when measuring the same subject. However they vary between different subjects. The sound variability for a given subject was 20dB, which is quite much. The authors note that for subject 1 the breath sound power with targeted flows correlate with free sound power while with subject 2 and 3 the breath sound power is louder at nontargeted flows. The opposite happens with subject 4 where the louder targeted flow sounds are observed. The authors' hypothesis is that this variability and these offsets are the result of glottal size variation. The expiration sounds were greater than the inspiration sounds for all four subjects. This supports the hypothesis that the glottal opening is smaller during expiration than during inspiration. [34]

At the lowest flows, below 0.2l/s, only background noise was observed. In simulations the "cut-on" point where the Reynolds number exceeds the R_{crit} is approximately 0.25l/s for expiration and 0.35l/s for inspiration, respectively. This implies that for increasing flows the expiration sounds should be observed first. This is the case of subjects 1, 2 and 3. In the middle range from 0.2l/s to 0.5l/s it was assumed that the simulation curves would bracket the sound power data points in figure 3.25. This was true only for subject 2, for which the predictions bound much of the sound power variation over this range. The simulation curves followed the mean difference of expiration and inspiration power flow curves in other cases. The large flows of over 0.5l/s showed that the measured sound powers begin to fall below the predictions. According to the authors, this could be explained with the experiment of Brancatisano[44] who has noticed that an increase of 30% in the glottal size has been observed for flows between 0.5 and 1.2l/s. The increase in the glottal size would reduce the turbulent flow and sound being created. [34]

From the model it can be assumed that the only origin of sound is the turbulent flow. More complex sound source candidates for modelling would be air collisions

with obstacles or with the walls of the tract. However, these parameters cannot be included in the model easily because a detailed description of flow patterns and geometry of the airways of the modelled subject is needed. [34]

3.4 Effects of body position and nasal cavity on tracheal sounds

Nasal tract   The effects of nasal breathing have not been investigated in many research publications. In 1998 Kraman et. al [45] studied the effects of breathing pathways on tracheal sound spectral features. They showed that nasal breathing altered the spectra significantly. The difference between spectral features during nasal and mouth breathing can be observed. There was intersubject variability, but in general it can be said that the amplitudes of the power spectrum were greater during nasal breathing than during breathing through the mouth. The breathing was done through a mask, which generated some additional peaks to the spectrum. Therefore, a more detailed comparison of nasal and mouth breathing spectra is difficult. However, the observed difference between nasal and mouth breathing indicates that the vocal tract (i.e. nasal and oral cavities) has a role in the formation of tracheal breath sounds. The effects of the nasal tract is studied theoretically by Flanagan [46], The equivalent acoustic circuit for nasal and oral tracts are presented as well.Unfortunately, the modelling is done from the point of view of speech science, and therefore only voiced sounds are studied. In breath sound research the effects of nasal cavities have not been modelled. One reason is that usually subjects breath through the pneumotachograph while wearing a noseclip [36, 47]. In those situations, with targeted flows the effects of nose cavities are cancelled, but in long term monitorings (as in sleep studies) the subjects breathe freely and the nasal tract may affect the measurements. The nasal sounds measured from the nostrils were studied by Seren [48, 49] and Tahamiller et al. [50, 51, 52] These are reviewed in chapter 4.

Body position   The effects of the body position on normal tracheal breath sound is studied by few. The pharyngeal cross-sectional area of healthy humans decreases in the supine position, compared to the sitting position [53, 54]. The postural effects on upper airway geometry are well recognized [54]. Because the airway geometry affects the generation of respiratory sounds, the body position may affect the sounds [53]. In polysomnography recordings the situation is complicated even more, because the patients are sleeping and the neuromuscular activation affects the properties of the upper airways[53]. Fiz et. al [53] have shown in their article in 2008 that when the subject is sitting, the tracheal sound amplitude versus the flow rate is similar to when the subject is in the prone position. No differences were found in the frequency range

from 150Hz to 1200Hz. The main target of the article was the lung sound position dependence. It was noted that a significant change in the lung sound amplitude was found only in the lateral decubitus positions. In 1996 Pasterkamp et al. [54] wrote an article: "Posture-Dependent Change of Tracheal Sounds at Standardized Flows in Patients With Obstructive Sleep Apnea". Tracheal sounds were recorded when the subjects were sitting and supine. A control group included obese subjects and snorers. They discovered that with the subjects in the control group some positions of the spectral peaks changed with posture. The tracheal sound amplitude increased when the control subjects changed their position from sitting to supine. At low frequencies (0.2 to 1 kHz) the mean increase was approximately 1.7dB. It was 1.3dB at medium frequencies(1 to 2 kHz) and 5.6dB at high frequencies (2 to 3 kHz). The results of the obstructive subjects are discussed in chapter 4.

In 2002 Que et al.[55] wrote an article in which phonospirometry was studied. The goal of the study was to estimate the tidal volume as a function of time from tracheal sounds. The other spirometrical parameters were then calculated from the tidal volume. They did the flow rate calibrations with the pneumotachography in the sitting position while the tracheal sounds were recorded. They discovered that rotating the head had only a small effect on the estimated volume. The error due to rotation and neck flexion was less than 16%. With neck extension the error was 28% at most. It was also noticed that the breaths taken during the transition of the head from one position to another were larger than the breaths taken before and after the transition of the head. The position of the body had no effects on the results when the position was changed from sitting to standing. The errors between measured and estimated values were largest during the supine position. The volumes were systematically overestimated and the greatest error observed was as large as 70%. However, it should be noted that this error of over 50% only happened with one subject. The errors for the other subjects were 11%, 14%, 20%, 25% and 38% in the supine position. So there was high variability in the errors between the different subjects, but this was not discussed by the authors. [55]

3.5 Respiratory flow rate estimation based on tracheal sounds

The method for evaluating the flow rate from tracheal sounds has been proposed by several authors [56, 57, 55, 58, 40, 59, 60, 61]. All these results from previous authors confirm that tracheal sound amplitude is related to flow rate. The advantages of flow rate estimation from tracheal sounds is that the sound recording is noninvasive, it can be used for long term monitoring and it has a relatively high position independency when compared to the pneumotachograph. The signal analysing methods which have been used to evaluate flow have been under debate and several approaches have been used. It can be said that there are two ways to handle the sound signal.

The first is to analyze the signal in a time domain for example by calculating the amplitude, signal variance etc. The other method is to convert the signal into a frequency domain with FFT and to study the spectral parameters. A common feature for all studies has been the calibration of tracheal sounds versus the flow rate. Usually this has been done with pneumotachography. Because the sound power is exponentially related to flow rate as expressed in equation 3.20, this relation can be presented as a linear line when a semilog diagram is used. This linear line can be expressed with equation $y = C_1 x + C2$, where y is the sound power and x is the flow rate. When the calibration is done the coefficients C_1 and C_2 are determined. A disadvantage of the method is that for calibrating purposes the subject must breathe some sample breaths with each flow rate [60]. There is also another method which is based on the entropy of the breath sound. The advantage of this method is that the calibration coefficients can be determined from one breath sound sample at a known medium flow rate [60]. For clinical purposes this would mean that the patient does not have to make calibration breaths with high flow rates. In many cases the patient may be unable to take deep enough breaths. It has been shown in literature that the sound amplitude is zero even if small flow rate exists. This has been explained previously in this chapter where the critical Reynolds number was introduced. Therefore, evaluating small flow rates is much more difficult than evaluating larger ones.[34]

Summary It can be concluded from the tracheal sound measurements that

1. The flow rate does not affect the pattern or position of spectral peaks.

2. The sound amplitude is proportional to the air flow.

3. The sound amplitude is proportional to the gas density.

4. The frequency peaks are shifted to higher frequencies during Heliox breathing. The scaling ratio for Heliox has varied from 1.47± to 1.76 which is approximately the square root of the air to Heliox wave speed [36].

5. The lumped element models can roughly predict the resonance frequencies of the respiratory tract. A detailed spectrum cannot be calculated because of unknown sound source distribution.

6. It may be possible to evaluate the flow rate from tracheal sounds, but more research is needed.

Sound generation by turbulent flow was presented. The shape of a TBS spectrum is mainly determined by the turbulent flow source, but the sensor design affects the

measured values. There are shallow peaks on the decreasing slope of the spectrum. An explanation is that they are caused by the resonances of the respiratory tract. This is supported by the fact that their locations can be predicted with mathematical models and that they shift when the density of the breathing gas is changed. The measured shift agreed well with the calculated theoretical shift. Effects of gas density also support the model of aeroacoustic sound production by turbulent flow.

4. ADVENTITIOUS TRACHEAL SOUNDS

Adventitious sounds in the trachea originate either from the upper or lower airways (lungs). In this thesis, only the effects of the upper airways on adventitious sounds are studied. Adventitious sounds are found with healthy people as well as those suffering from sleep disordered breathing (SDB). However, it has been shown that the sounds are different with healthy people and people suffering from obstructive sleep apnea syndrome (OSAS). The most familiar adventitious sound is snoring. It is also one of the most widely studied adventitious tracheal sound. It may be caused by pathological relaxation of the pharynx muscles or anatomical obstruction in the nasopharynx. Some of the anatomical obstructions can be treated surgically. However, nasopharynx surgery as a treatment for sleep related breathing disorders has been under debate. [62, 63]

The snore sound production mechanism is different from the one for normal breath sounds which are produced by turbulent flow. When the subject is snoring, pressure fluctuations make the tissues vibrate and this vibration characterizes the snoring sound. Most of the snoring sound is generated mainly in the pharynx section of the airways.

Other types of sounds are produced during flow limitation where breathing sound intensity increases as well, but the periodic nature is different when compared to snoring. The mechanism of the adventitious sounds, with all kinds of SDB, is usually obstruction in the upper airways.

4.1   The physics of collapsible tubes

The mechanisms of time dependent obstructions are different from permanent anatomical obstructions. The lumen of the airways becomes totally or partially closed in cycles during the breathing event. The reason for this can be explained with the model of the Starling resistor [62, 63]. This resistor model explains the relationship between pressure inside the lumen, outside the lumen and the state of the lumen (i.e. open, partially closed or totally closed). This model can be used to give a physical explanation for obstructions in apnea, hypopnea and flow limitation. However, the chance of permanent anatomic obstruction is excluded in this model, even though it could also lead to snoring or flow limitation. [62, 63]

In figure 4.1 the Sterling resistor model is presented. There are rigid-walled tubes

upstream and downstream from the box which represents the pharynx section. Inside the box, the tube has flexible walls. Pressure upstream from the box inside the tube is P_{US}. Inside the box, outside of the tube, the pressure is P_{Crit}. Pressure in the tube downstream from the box is P_{DS}. [62, 63]

During normal breathing, the upstream and downstream intraluminal pressures are greater than the critical pressure, $P_{US} > P_{US} > P_{Crit}$. As seen in figure 4.1 the tube is open all the way from the upstream segment to the downstream segment. [62, 63]

Figure 4.1 b presents a case where the intraluminal upstream pressure is greater than the critical pressure, but the downstream intraluminal pressure is lower than the critical pressure, $P_{US} > P_{Crit} > P_{US}$. The tube opens and the flow begins, but the section between the box and downstream stays collapsed or it flutters. An intuitive explanation of this partial collapse is the following: When the pressure opens the tube section between the box and the tube downstream, the flow starts. Because the pressure downstream is below the critical pressure, the flow inside the flexible part of the tube decreases below the critical pressure and the flexible tube collapses again. This is the case during hypopnea and flow limitation. [63]

Figure 4.1 c presents a case where the intraluminal pressures upstream and downstream are lower than the critical pressure, i.e. $P_{Crit} > P_{US} > P_{US}$. As a result, the flexible walls have collapsed against each other. The situation is typical for an apneic event. It has been shown that the critical pressure is much higher with subjects with OSAS than with healthy snorers [64] and that tissue properties affect the critical pressure [63].

4.1.1   Mechanisms of snoring

Gavriely et al.[65] have presented a mathematical model of snoring from the point of view of airway collapsibility. The top wall of the collapsible airway segment is modelled as a plate which is connected to the external tissue with a spring. During the snoring, the spring and the plate start to oscillate. This causes constriction in the airways every time the top wall moves downwards from its initial position. A familiar equation of motion is presented for the spring mass system. The forces acting on the plate are expressed by the means of pressures inside and outside of the airways. The model is presented in figure 4.1.1. The plate has a length L, a width W and a mass m. The spring has an elastic constant K. When the spring is at rest the height of the airway segment is b_0. During the vibration the height of the airway at instantaneous time t^* is b^* which is also the y-axis position of the plate. The cross-sectional area of the segment at a given time is $A_{CS} = b^* \times W$. The area of the plate is $A_P = L \times W$. There is flow $\dot{V} = \frac{dV}{dt}$ in the x-axis direction. The collapsible airway segment is connected to the upper airways which have an

A Open upper airways $\quad P_{Crit} \quad P_{US} > P_{DS} > P_{Crit}$

P_{US} $\qquad\qquad\qquad\qquad$ P_{DS}

Upstream $\qquad\qquad\qquad\qquad$ Downstream
Segment (US) $\qquad\qquad\qquad$ Segment (DS)

Collapsible Segment

B Upper airway collapse $\quad P_{Crit} \quad P_{US} > P_{Crit} > P_{DS}$
with reduced airflow

P_{US} $\qquad\qquad\qquad\qquad$ P_{DS}

Upstream $\qquad\qquad\qquad\qquad$ Downstream
Segment (US) $\qquad\qquad\qquad$ Segment (DS)

Collapsible Segment

C Complete occlusion $\quad P_{Crit} \quad P_{Crit} > P_{US} > P_{DS}$
of the upper airway,
no airflow

P_{US} $\qquad\qquad\qquad\qquad$ P_{DS}

Upstream $\qquad\qquad\qquad\qquad$ Downstream
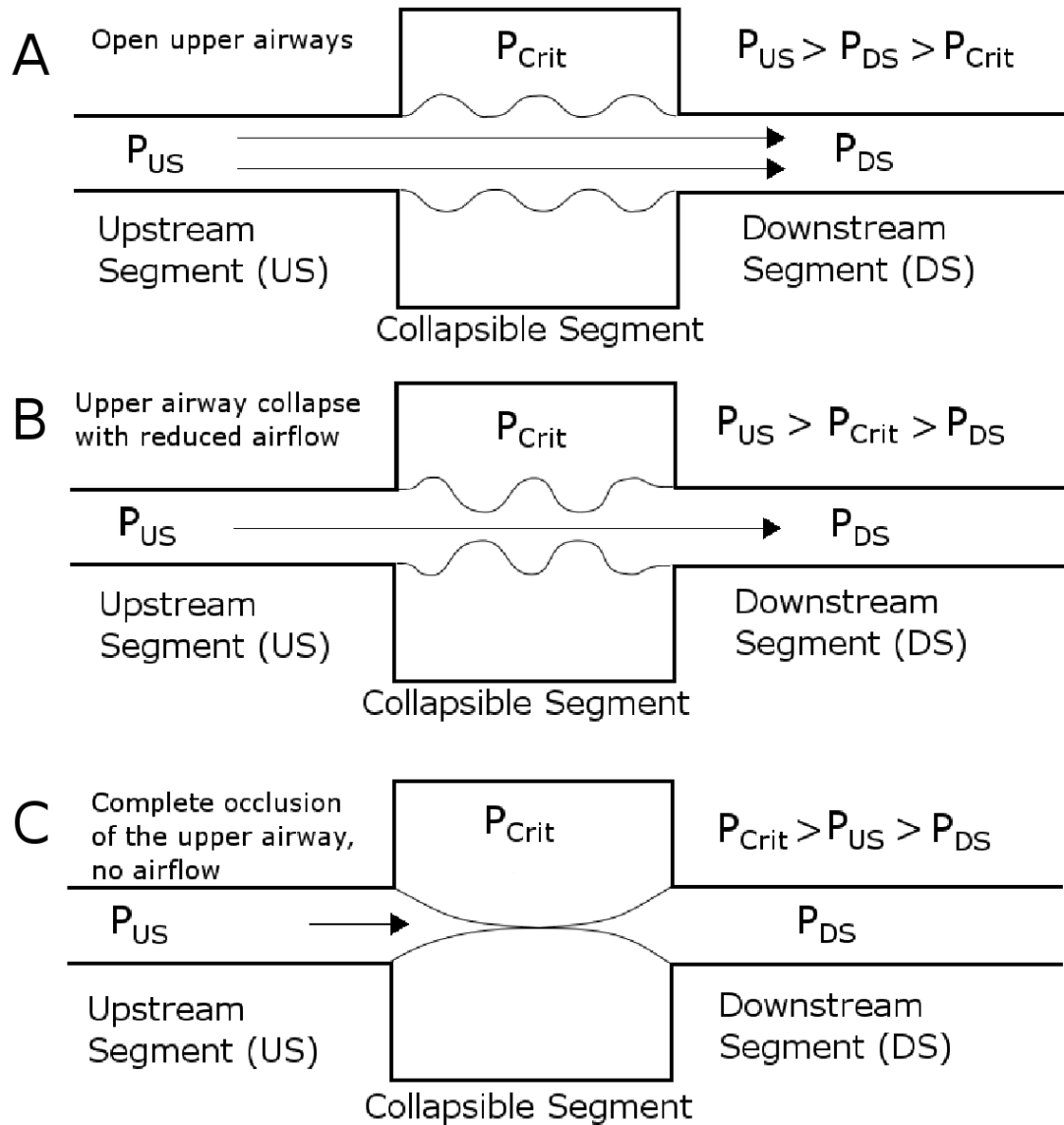Segment (US) $\qquad\qquad\qquad$ Segment (DS)

Collapsible Segment

Figure 4.1: The Sterling resistor model. Figure A presents a normal situation when the upper airways are open. Figure B presents a case of partial obstruction in the upper airways when flow is limited. Figure C presents total closure of the upper airways where flow cannot be detected. Modified from [63]
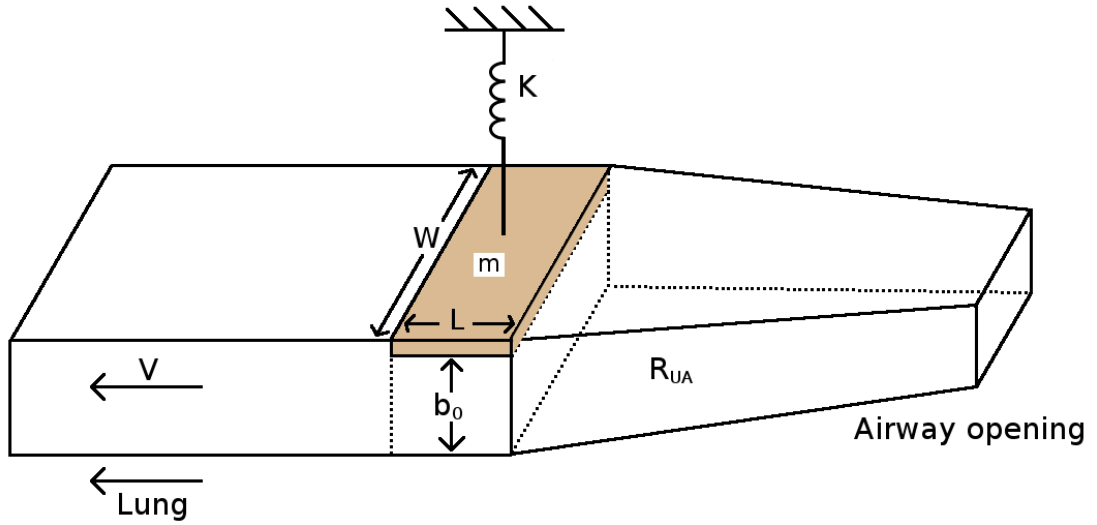
Figure 4.2: Resistor plate model of a collapsible tube. Modified from [65]

upper airway resistance R_{UA} to the flow. The intra-airway pressure drops across the collapsible segment. This happens because of two reasons; the effect of viscous resistance in the upper airways and the Bernoulli effect. The pressure drop due to viscous resistance is $P_R = R_{UA}\dot{V}$. Bernoulli's law states that the pressure drops in the system if the speed of the flow increases. In an airway segment this can be expressed as $P_B = \frac{1}{2}\rho U^2$, where ρ is the density of the gas and $U = \dot{V}/A_{CS}$ is the speed of the gas. The pressure drop inside a tube causes a force to the plate which tries to pull the movable wall (i.e. the plate) towards the opposing wall. A counter force is caused by the spring to the plate. The forces affecting the plate are products of the pressures and the surface area of the plate. The equation for the motion of the system can be expressed as [65]

$$ F = ma = m\frac{d^2 b^*}{d(t^*)^2} = F_E - F_R - F_B, \qquad (4.1) $$

where m is the mass of the system, a is the acceleration, $F_E = K(b_0 - b^*)$ is the elastic force caused by the spring, $F_R = P_R A_P = R_{UA}\dot{V}A_P$ is the force caused by resistive effects and $F_B = P_B A_P = \frac{1}{2}\rho U^2$ is the force originating from Bernoulli's effect. The minus signs in equation 4.1 indicate that the force is being directed towards the lumen of the airways. [65]

The right hand side of equation 4.1 determines the state of the segment. If the value is positive the plate accelerates outwards or decelerates inwards. With zero values no acceleration occurs and the plate is at rest or moving with constant velocity. If the value is negative the plate accelerates inwards. When the original state of the plate is at rest the right hand side of the equation must be nonzero in order to put the plate into motion. The maximal elastic force is obtained when $b^* = 0$ and the elastic force becomes $F_{E0} = Kb_0$. By substituting the expression of forces to equation 4.1 it follows [65]

$$m\frac{d^2b^2*}{d(t^*)^2} = K(b_0 - b^*) - R_{UA}\dot{V}A_P - \frac{1}{2}\rho U^2. \qquad (4.2)$$

The terms of the previous equation can be regrouped and expressed in a nondimensional form. The following definitions are done $b = \frac{b^*}{b_0}$, $t = t^*\sqrt{\frac{K}{M}}$, $\mu = \frac{F_{E0}F_{B0}}{F_R^2}$ and $q = \frac{F_R}{F_{E0}}$. By using these dimensionless variables, equation 4.2 can be expressed as: [65]

$$\frac{d^2b}{dt} = 1 - b - q - \frac{\mu q^2}{2b^2}. \qquad (4.3)$$

The solutions for equation 4.3 are not presented here, but the major outcomes are discussed briefly. The stability of the channel is determined by q and μ. The q is affected by the forces caused by Bernoulli's effect and by the maximal elastic force. In addition to these forces, μ is affected also by the force caused by resistive effects. There is a certain critical value of $q_C(\mu)$ which defines the stability condition of the system. If $q < q_C$ the system is stable, as long as b does not obtain very small ($b^* << b_0$) or very great ($b^* >> b_0$) values. When $q > q_C$, closure is inevitable. [65]

The global maximum of q_C is 1, when $\mu \to 0$. Therefore, closure happens every time the absolute value of $q > 1$. This means that the closure is independent of the mass of the wall of fluid, but if the upper airway resistance is great enough, the airway pressure during inspiration overcomes the supporting elastic force and closure takes place. [65] If $q < 1$ the instability depends on μ as well. This means that the fluid density of Bernoulli's effect gives its own contribution. As a summary it can be said that increasing μ or q causes destabilisation of the airway. If μ is small, an increase in μ causes the destabilisation more easily than an increase in q, but when μ is large the increase in q is a greater destabilator. [65]

In the previous paragraph the effects of nondimensional parameters were investigated. Here the effects of individual variables are studied while the other variables are kept constant.

Increase in the flow $\dot{V}$ causes q to increase, but keeps μ unchanged. As a result, the transition from a stable to an unstable state occurs and eventually the airway collapses. [65]

A decrease in the density of the gas ρ causes q to keep constant, but decreases μ. When $q < 1$, a transition from an unstable to a stable state may occur. [65]

The reduction of segment height b_0 causes both q and μ to increase and this causes a transition from a stable to an unstable state. [65]

As was previously suggested the increase of R_{UA} eventually induces closure. This is because q increases and μ decreases, but $q\mu^{1/2}$ remains constant. [65]

The wall plate dimensions also affect the state of the collapsible segment. When the wall length L increases, the area over which the pressure acts increases and eventually the channel will undergo collapse. The same result is achieved if K is decreased. It means that the wall becomes less rigid and more prone to close. From the point of view of q and μ, if the fraction rate of the spring elastic constant and the wall length K/L decreases, q increases or μ decreases, but $q\mu$ remains constant. [65]

The width W of the airway segment has a certain value range which supports the stable state of the segment. If the values increase or decrease from this value range the result is a collapse of the airway. When W increases the area over which the destabilizing pressure acts increases and when W decreases the cross-sectional area decreases and the flow speed increases, resulting in an increased destabilazing force. This is because q and W increase and μ decreases at the same time that $q\mu^{1/3}$ remains constant. [65]

The authors note that the damping effect can be added to the model, but it has no effect on the global stability results. They also note that this model does not address the reopening issue of the airways after the collapse. This reopening and oscillation are reviewed in the flutter section. They state that the model assumes that a complete airway collapse will happen if the unstable state has been achieved. They also assumed that the flow $\dot{V}$ remains constant independently of vibrational movement of the wall. This means that the model might not be adequate in the case of flutter when the wall vibration is rapid. [65]

4.1.2 Mechanisms of flutter

Flutter is a snoring-like phenomenon. The frequencies of flutter are much higher than the ones of snoring sounds. Flutter produces sound like wheezing which is most often associated with asthmatic breathing sounds. [66, 67] In the case of asthma the wheezing sound is produced in the lung tissue when the alveolar lumens are narrowed. Wheezing sounds can also be produced in the upper airways, for example in the case of polypical deformation in the nasal airways [68]. The power spectrum of wheezing sounds is characterized by a sharp peak at the nominal frequency of the wheeze [66]. Sometimes the harmonics of the nominal frequency have been seen as well [66]. In the following chapter flutter is studied as a review of two
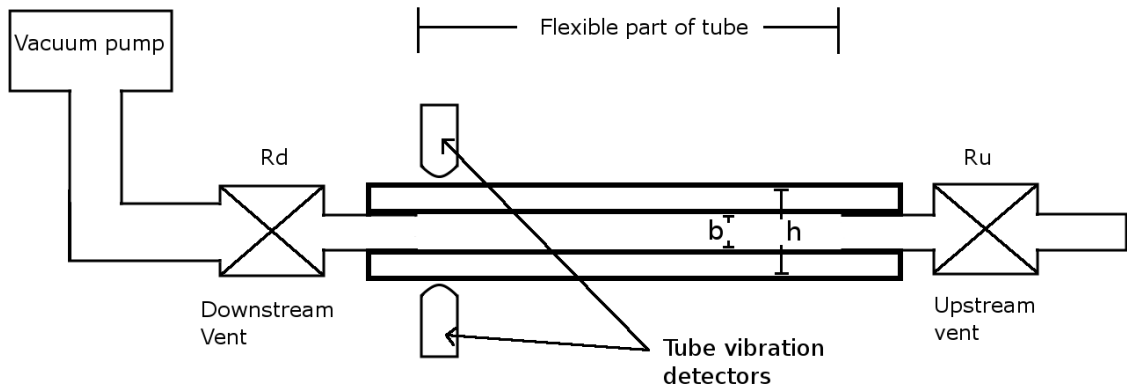
Figure 4.3: Experiment setup with a flexible tube. Modified from [67].

articles. Gavriely et al. [66] produced flutter artificially into a flexible-walled tube and made measurements in order to study the parameters affecting the creation of the flutter. They showed that in some cases the wheezing sound was produced when flutter occured. In the second article written by Grotberg et al.[69], theoretical models of wheezes are generated based on the measurements of Gavriely and his coworkers.[66, 69]

Before the measurements of Gavriely, thin-walled Penrose tubes had been studied by Conrad (reviewed in [67]). One of his findings was that the soft-walled tube collapsed when the liquid fluid (water) passed through the tube and the transmural pressure became negative. This Penrose tube model can be used to explain the blood flow in the cardiovascular system, but it does not describe the behaviour of air flow in the respiratory system that well. This is because the physical characteristics of these systems are different. The mass of flow is smaller and the mass of the self supporting airway wall is greater in the respiratory system. In fact the most effective parameter of the behaviour of the tube is the ratio of its wall mass to its flow mass [69]. Moreover, the differences in the flow exist because gaseous flow is compressible, unlike liquid flow. [69]

In the measurement setup made by Gavriely, air flow and a thick-walled tube was used. This increased the ratio of the wall mass to the flow mass. In particular, the mass ratio of the wall and the fluid is three orders of magnitude greater than that of the liquid flow in the Penrose tubes. The measurement setup is presented in figure 4.1.2

The soft-walled tube was connected from both ends to rigid fittings that were 60cm apart. The elastic properties of the tube wall were measured as well. Young's modulus Y is defined as the ratio of longitudinal stress to longitudinal strain and elasticity in the radial direction E_w. The external diameter $2h$ of the collapsible tube was measured with a micrometer at the location of constriction. The inner radius b was calculated by subtracting the thickness of the wall from the outer radius h of

the tube. A vacuum pump was used to generate low-pressure flow inside the tube. Two valves were positioned downstream and upstream from the tube. They were used to control the upstream and downstream resistances R_u and R_d. Differential pressure transducers were placed at both ends of the flexible tube. Proximity probes were used to monitor wall motion during oscillation. Two different sizes of tubes and two excised dog tracheas were used in the measurements. [67]

The flow inside the tube is related to the pressure difference $P_u - P_d$, where P_u is the upstream pressure and P_d is the downstream pressure. Gavriely et al. found out that there was a particular pressure difference for each tube and longitudinal tension when the circular shape of the tube was lost and the tube buckled. When the driving pressure was increased even more, the tube collapsed and the opposing walls collided with each other. [67] The flow pressure curves showed that the flow became limited at the same time , i.e. an increased $P_u - P_d$ pressure difference did not increase the flow in the tube. This tube buckling appeared at the downstream section of the tube. In addition, an onset of the tube oscillations was observed. However, certain pressure-flow conditions were found where flow was limited, but oscillations did not exist. Therefore, it can be said that flow limitation is a necessary condition for oscillations, but flow may be limited even if no oscillations are present. It was noticed that for a fixed flow the oscillation frequency increased and the distance between opposing tube walls decreased as the pressure difference was increased. Because the distance between the walls was reduced while the flow was constant, the flow speed U had to become higher as the cross-sectional area of the tube became smaller. It was discovered that the oscillation frequency increased as the inner radius of tube b decreased and when the flow speed U increased, i.e. the frequency $f \propto \frac{U}{b}$. [67] The longitudinal tension did not affect the oscillation frequency significantly. The external pressure applied to the tube had no effect on the relationship of driving pressure flow, but it decreased the upstream pressure P_u, at which the flow became limited. The authors noticed that transmural pressure dictates the tube geometry and pressure-flow behaviour instead of individual pressure inside or outside the tube. The amplitude spectrum of the wall oscillation showed a sharp peak and its harmonics were observed occasionally. [67] The theoretical investigation was based on the model of the Penrose tubes. The equation for flow velocity was $C = \sqrt{E/\rho_f}$, where E is the elastic property of the tube and ρ is the density of the fluid. In the model it is assumed that the tube wall is massless. The authors discovers that the measured values were smaller than the calculated ones. It was suggested that the difference was caused by an error in measuring the inner diameter of the tube during the collapse. As a result, this error would have affected the measured flow speed value U. Another, more probable reason was the oversimplified tube model; a thick-walled tube was used in the measurements and

in the calculations it was assumed that the wall had no mass at all.[67]

Grotberg et al. [69]have made a theoretical investigation based on the measurements of Gavriely et al. [67] The aim of the model is to predict the critical fluid speed at which oscillations begin and the frequencies of the oscillation. In the first condition it was required that the flow consist of an inviscid core region with viscous effects confined to a thin boundary layer near the wall. It was noted that the flow experiences a relatively rapid area contraction as it enters into a partially collapsed region in the fluttering tube. The second assumption was that the viscous effects of the wall are as small as those of the fluid. Grotberg calculated the Reynolds number based on the measurements of Gavriely and found it to be $Re = 1012$, which indicates that the flow was turbulent[69]. Gavriely pointed out that the flutter may be caused by fluid dynamic flutter and vortex induced oscillations. The theoretical analysis of the vortex shedding model was not available in any detailed form at that time. Gavriely was only able to estimate the natural frequency of the tube wall which pointed out to be lower than the measured ones. Gavriely does not discard the possibility of vortex induction; he stated that the calculations were "grossly" simplified. Grotberg used two different stability theories to explain fluid dynamic flutter. The linear stability theory shows that the system is stable as long as the dimensionless flow speed S is smaller than the critical dimensionless flowspeed S_0 and that it takes a large enough flow speed to cause flutter. The nonlinear stability theory also shows that the oscillations do not exist below the critical flow speed. It also indicates that the amplitude ϵ of the oscillations is proportional to the square root of the excess speed of the flow, i.e. $\epsilon \propto \sqrt{S - S_0}$. Based on the nonlinear theory the oscillation frequency can be expressed as [67]

$$\omega = \omega_0 + \frac{S - S_0}{S_2}\omega_2.$$
(4.4)

It can be seen from 4.4, that the frequency increases as the excess speed $S - S_0$ increases. If S increases 70% the frequency ω would increase by 70%. Grotberg pointed out that the nonlinear theory should always be used when fluttering occurs and the critical speed is exceeded. When discussing the parameters Grotberg states that the distance between the tube walls $2b$ has an effect on the critical velocity, but it has little effect on the oscillation frequency. Clinically this would mean that the wheezing frequency would not give much information about the airway size. Also helium-oxygen breathing has been shown to have little effect on the pitch of wheezing [70], which implicates that the frequency of wheezing is more dependent of the tube wall density than of the flow density. [69, 66]

4.2 Measured snoring sounds

SDB is a physiological state in which the subject has a frequent occurrence of pathological respiratory events while asleep. These pathological respiratory events which last at least 10 seconds are apneas where breathing is totally stopped and hypopneas where a partial collapse of the upper airway occurs. Snoring occurs during these pathological events and is different to "simple snoring" which occurs when apneas and hypopneas do not exist.[71]

Spontaneous snoring appears during natural or drug induced sleep, but the snoring sound can be also simulated while awake. Snoring appears usually during inspiratory breathing, but it may also exist during expiration. The sound of snoring is different during natural and induced sleep. Snoring is caused by the vibrational movement of the pharynx. The detailed characteristics of snoring are caused by oscillations of the soft palate, the pharyngeal walls, the epiglottis and the tongue [72, 73]. During sleep, the dilator muscles may reach a state of relaxation. As a consequence the lumens of the upper airways become narrowed and the flow resistance increases. The characteristics of snoring are not easily defined. The objective identification and quantification of snoring related events has proven to be difficult. The situation is complicated even more because the acoustics of snoring is not a homogeneous phenomenon. The snoring sound is dependent of the sleep period and it may change from night to night[71]. The sound is also affected by the path of the airway that the flow takes when it travels during breathing. Nasal, oral and oronasal breathing all produce different sounds from each other. The site of the narrowing may be at the palatal segment, the tongue base, the supraglottic space or the combination of these [73]. The site of the snoring affects the snoring sound. Variation of the snoring sound is also related to the sleep stage, the body position[74], naturally occured vs. induced sleep and the presence or absence of SDB.[71]

Snoring is often studied from the point of view of speech analysis in scientific articles. This is because physiological similarities between these two do exist. In speech the laryngeal area produces the sounds. The focal folds vibrate and produce a fundamental frequency F_0 and a harmonic series of frequencies for F_0. The upper airways (i.e. the vocal tract) act as a resonator which filter and amplify specific frequencies. The mechanism of snoring is roughly the same, but instead of vocal fold vibration the snoring sound is produced by vibrational activity of the pharyngeal structures. Vibration consist mainly of flutter in the soft palate combined with vibration in tonsils, the tongue base and the epiglottis[72]. Another major difference is that snoring occurs mainly during inspiration while speech is produced during expiration. Yadollahi [75] and Ng [76] have studied the formants of snoring and normal breathing. They used a tracheal microphone for recording and they observed

that the formant frequencies of snoring were lower than the formants of breathing sounds. Both the snoring and breathing sounds had frequency components up to 5kHz. However, the authors used the autoregressive (ARR) model for estimating the spectrum and linear predictive coding (LPC) for signal analysis. These methods might be more vulnerable to errors when estimating the spectrum than the FFT algorithm.[71]

Beck et al.[30] found two different dominant snoring patterns: "the simple-waveform" and the "complex-waveform". The simple-waveform is quasi-sinusoidal, with variants. The secondary internal oscillation is almost totally missing. The power spectrum contains only 1-3 peaks and the first one has the greatest amplitude. They have a higher frequency content than the complex-waveform snores and it has been thought that they are produced by the oscillation around the neutral position without actual closure of the lumen of the airways. The complex-waveform consists of repetitive, equally spaced trains of sound structures with large amplitudes followed by a decaying amplitude wave. The power spectrum of snoring is comb-like i.e. there are several equally spaced peaks in the spectrum. The complex-waveform may be explained as being a result of actual closure of the airway lumen and the collision of the airway walls. Endoscopic findings indicate that complex-waveform snoring occurs during palatal snoring and simple-waveform occurs when tongue base snoring occurs[77]. Endoscopic snoring was made under anaesthesia, so the characteristics of snoring is different than during natural snores. Snoring has been tried as a diagnostic tool for detecting obstructive sleep apneas (OSA). Another possible application of snoring sound analysis is in the area of ear-nose-throat medicine, where the place of sound generation is under interest. It has been suggested that surgical treatment of snoring is more effective if the soft palate is the main source of snoring [71]. However, more studying is needed. [65, 78, 73, 77]

Agrawal [73], has shown that different tissues have different vibrational frequencies. Snoring was studied with sleep nasendoscopy. The vibration frequencies of tissues were categorized in groups of palatal, tongue based, mixed, epiglottic, tonsillar, secretions and unknown snoring sounds. Palatal frequencies were up to 350Hz only, while snoring frequencies up to 3500Hz have been found. [71]

Abnormal sounds are produced mainly for two reasons: extra constrictions in the airways or increased flow velocity through the constriction. As a result, increased sound power and higher frequencies can be found [29]. Frequencies up to 10kHz have been found with children suffering from SDB when external microphones have been used [79]. Rembold et al. [80] have made a model which tries to explain the high frequency voices produced by children suffering from SDB. The mechanism of the production of high frequency noise is unknown. It has been suggested that flow with a high velocity produces sound when it travels through the constriction in the

upper airways. Rembold's mechanical model of upper airways include both open-open ended and closed-open ended plastic pipes. It was shown that open-ended pipes produce even harmonics whereas closed-ended pipes produce odd harmonics when a jet of air was injected into the pipes. [80]

As described in chapter 3, the power of tracheal breathing sounds is a strong function of flow. Therefore, when the breathing is stopped apneic events can be found easily from the tracheal signal. Yadollahi et al. [81] have developed a method for detecting the OSA acoustically. The OSA is traditionally detected from the respiratory flow signal and from the blood saturation level S_aO_2. With this method the tracheal sound is measured with a microphone instead of using respiratory flow measurement. Yadollahi has calculated the flow from the tracheal breath sound signal. An automatic method that uses only tracheal sounds and S_aO_2 signals was developed for apnea and hypopnea detection . The authors suggest that the method be used with ambulatory devices.

The posture dependent change of tracheal sounds in OSA patients have been studied by Pasterkamp et al. [31]. Tracheal sounds were measured from a group of healthy people and a group of OSA patients. The breath flow was standardized from 1.5l/s to 2.0l/s and the tracheal sounds were measured when the patients were sitting and when in the supine position. In both patient groups the intensity of the sounds was louder in the supine than in the sitting position. It was noticed that during the inspirations the sound intensity increased more with the OSA patients than with the normal patients throughout the frequency range. During the expirations the tracheal sound intensity was higher with OSA patients only at high frequencies (2kHz-3kHz) in the spectrum. The apnea-hypopnea index was directly related to tracheal sound intensity in both patient groups. The increased sound intensity with the OSA patients in the supine position was assumed to happen because of increased turbulence in the upper airways. This was explained to happen because OSA patients and heavy snorers may have anatomical abnormalities, like a narrow pharynx or a physiological dysfunction like a reduced ability to dilate the pharynx. The observation that during expiration the difference with patient groups occured only at high frequencies supports the idea of Tenhunen[29] and Tahamiller [50] that increased turbulence by obstruction increases the high frequency components of the sound spectrum.

Another type of abnormal breathing sounds may exist during flow limitation with patients that have SDB [29, 79]. The exact reason is unknown, but an explanation could be the difference in the airway geometry during sleep compared to healthy people. In obstruction the cross sectional area of the airway lumen is reduced. This constriction acts as an obstacle for turbulent flow and the sound spectrum is modified and extra sound is produced. These abnormal sounds have been studied much less

than pure snoring, but some observations have been published. The importance of these kinds of studies is that they suggest that apnea and hypopnea related snoring or wheezing may not be the only sound related phenomena that occur during sleep related breathing disorders.

Partial obstructions are much more difficult to find. It is a known fact that tracheal sound spectral power is increased and the frequency band is shifted upwards with patients with a lung disease[28], [82]. This kind of behaviour is also found in a work of Tenhunen et. al. [29]. They showed that the maximum frequency was slightly shifted upward with patients with flow limited breathing and obstructive apnea. Details of this study are refered to later in this chapter.

Nasal airways provide one half of airflow resistance of the whole respiratory system. This resistance is coupled to the cross sectional area of the nasal cavity which can be investigated with the method of acoustic rhinometry (AR)[52]. In AR the external sound is transmitted to the nasal airways and the cross sectional area is estimated from the reflections of the transmitted sound.[52] The attempts to simplify the AR measurement system has led the researchers to investigate the expiratory and inspiratory nasal sounds, recorded from the nostrils. It has been discovered that increasing the flow rate increases the mean amplitude and the mean frequency [48], [49]. The nasal sound generation mechanism is suggested to be a turbulent flow passing through a constriction.[50, 48]. More interestingly, it has been shown that patients with nasal obstructions produce higher frequency sounds than healthy patients. [49]. These nasal sounds support the idea that during flow limitation, higher frequencies can be heard in breathing as Tenhunen et al. have noticed [50].

Tenhunen et al.[29] have studied polysomnography patients with a HeLSA-microphone at the Sleep Laboratory of Pirkanmaa hospital district. Some of the subjects suffered from OSA and prolonged flow limitation and some of the patients were healthy. Tracheal sounds were recorded and analyzed and it was observed that in addition to periodic apneas and hypopneas a prolonged flow limitation was visually observed [29]. For these flow limitation periods it was typical that during every inspiration the tracheal sound was loud [29]. The spectrum and sound amplitude during the flow limitation are presented in figure 4.5.

Ten minute periods of tracheal sounds during flow-limitation, normal breathing and periodic apnea-hypopnea breathing were analyzed from 30 patients. Tenhunen et al. [29] hypothesized that flow limitation represented prolonged obstructive breathing. The amplitude spectra of tracheal sounds were calculated and it was observed that in general, the sound intensity was greater during flow limitation and periodic apnea-hypopnea breathing than during normal breathing (see figure 4.5).

The hypothesis was proven to be correct since it can be seen from figure 4.5 that at frequencies above 1000Hz the possible amplitudes of normal breath sounds have
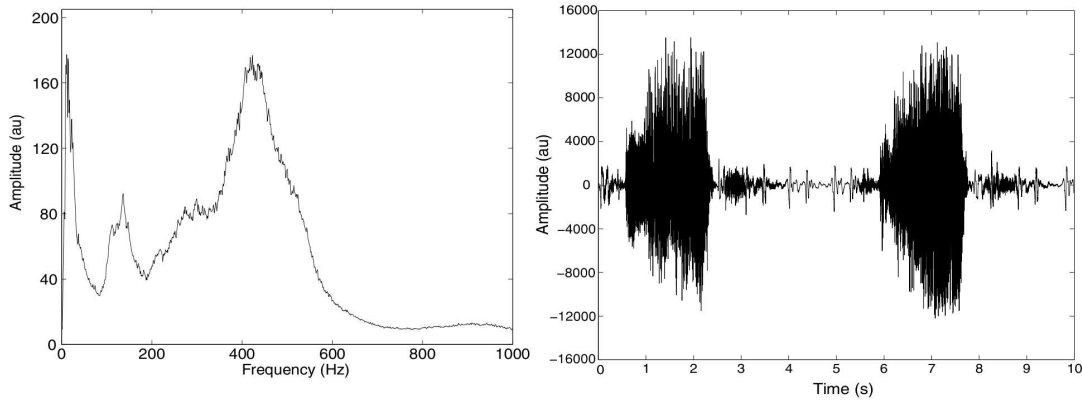
Figure 4.4: The spectrum and sound signal during flow limitation. The difference in the amplitude of the sound is obvious in the figure on the right-hand side. The amplitude during the inspiration is more than three times the amplitude during the expiration.[29]
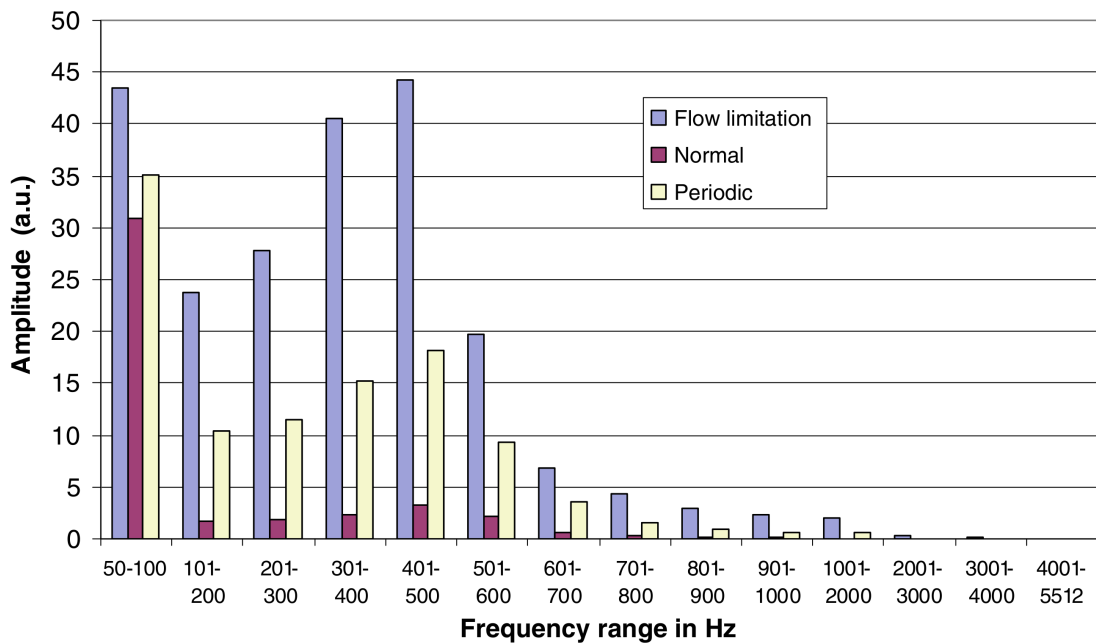

Figure 4.5: The distribution of frequencies in the power spectrum. It can be seen that the frequencies over 1000Hz are observed only in the case of apnea-hypopnea and in flow limitation.

vanished into the background and only the amplitudes of flow limitation and periodic apnea-hypopnea breathing are observed.[29] It would have been interesting to see differences in the amplitude spectrum between expirations and inspirations as well as the background spectrum during respiration pause in the cases of normal, flow limited and apnea-hypopnea breathing. The maximum observable breath sound frequency would have been possible to estimate as well, if a logarithmic representation of the spectrum amplitude scale would have been used.

Summary   Adventitious sounds were explained with the Sterling resistor model which describes the behaviour of a flexible tube wall when pressures of the tube are altered. The critical pressure which causes the tube to collapse was explained to be caused by the surrounding tissue. The effect of tube geometry on the collapse of the tube was studied with the resistor plate model. Several parameters affected the conditions of the collapse. Measured snoring sounds were compared and it was pointed out that snoring can be associated with normal or pathological conditions. It was also pointed out that normal TBS were different with healthy patients and OSA patients when sitting and supine positions were compared. The higher frequency components of sounds were reported also in the cases where obstructions were present. Some of the adventitious sounds are not well known. It was indicated that with flow limitations there are other than snoring sounds, which might be caused by flutter.

5. AUDIO SYSTEM MEASUREMENTS FOR THE HELSA-SYSTEM

The Helsinki lung sound analyzer (HeLSA)-system includes a microphone, a microphone preamplifier, a sound card and a registering computer. The microphone is an air coupled Panasonic WM-60A (Matsushita Electric Industrial Co, Ltd, Osaka, Japan) mounted into a plastic cylinder. The coupler is attached onto the skin with double sided adhesive tape. A custom made amplifier (Pulmer, Helsinki, Finland) is permanently attached to the microphone. The electronics of the amplifier is presented in appendix ??. The amplified signal is fetched to the line in the connector of the Sound Blaster Audigy 2NX sound card. The sound card is connected to the computer with a USB connection. The measurement chain is presented in figure 5.1.
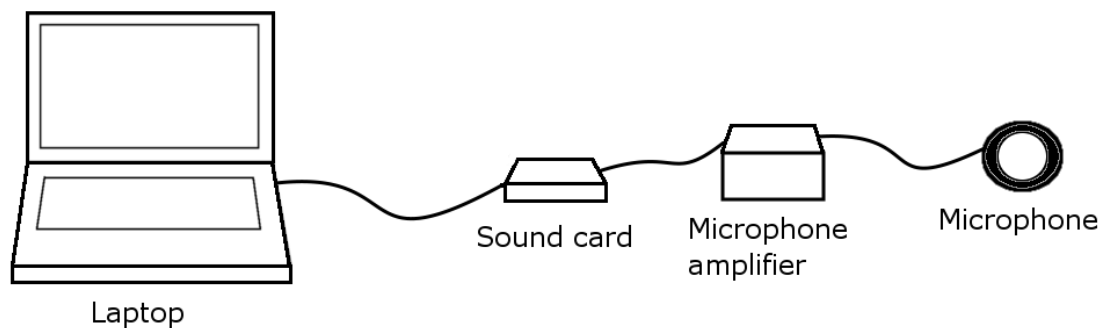
Figure 5.1: The measurement chain of the HeLSA-microphone

Audio system measurements, sometimes referred to as audio quality measurements, can be used to define the specifications of the system or to test if the system is working as defined in the specifications. These measurement are usually done by the designer of the system or the maintenance engineers.

In our case the audio system measurements were done, because the researchers wanted to know the quality of the sounds recorded with the HeLSA and in which way the HeLSA might distort the recorded signal.

It was observed that the two HeLSA microphones measured the sound signal

differently when the same sound source was measured. The signal of microphone 2 was much more noisy and the overall amplitude was greater than with microphone 1 (see fig 5.2 and 5.3). Because the HeLSA is an air coupled microphone, the technical weakness is the front surface of the microphone which is practically unprotected against environmental stress. As a result, both HeLSA microphones were sent to the manufacturer to be repaired. It was discovered that a simple and sophisticated quality measurement setup is needed. Moreover, quality measurements improve the knowledge of the way in which the tracheal sound recording instrument HeLSA works.

The quality changes may not be so critical in clinical work, because the methods used are quite robust, i.e. detection of apneas (no sound) or the patient is snoring (loud sound). What comes to scientific works made in the Sleep Laboratory, the quality of the microphone is essential and it is related to the reliability of the results achieved. For example the effects of an increased noise level in the microphone can be understood when figures 5.28, 5.29 and 5.30 of a tracheal sound spectrum is studied. The level of the sound signal and the maximum observable sound frequency decrease. The shallow peaks may disappear and the attenuation slope can be also distorted as the sensitivity of the microphone is reduced.

5.0.1 Sound card quality measurements

Sound card measurements were done at the Sleep Laboratory of the Pirkanmaa Hospital District (PSHP) and at the laboratory of electrical engineering of Tampere University of Technology (TUT). All the measurements were done according to the documents of PC Audio Quality Measurements [83] 9/1999 by Cirrus Audio Division and Audio Quality Measurement Primer by Intersil [84]. Both documents are based on AES17-1991 and EIAJ CP-307 CD measurement standards. Two different sound card models manufactured by Creative Labs were measured. The Sound Blaster Audigy 2 NX (Singapore) is clinically used at the PSHP and the Sound Blaster X-Fi (Singapore) is a candidate for replacing the model Audigy 2NX. Both sound cards are USB-connected to the computer. The sampling frequencies of 11 025Hz and 44 100Hz were used at the Sleep Laboratory of the PSHP, but 44 100Hz was used only at the Laboratory of Electrical Engineering at TUT. This was the case because the drivers of the sound cards could not be properly installed to the computer at TUT.

Sound card measurements at TUT The RightMark Audio Analyzer (RMAA) testing software was used for all quality measurements of the sound cards. The entire audio chain was measured. Therefore, the line out was connected to the line in and the external analog loopback was created. The weakness of the method is that the location of distortion or artefact cannot be determined exactly, because it
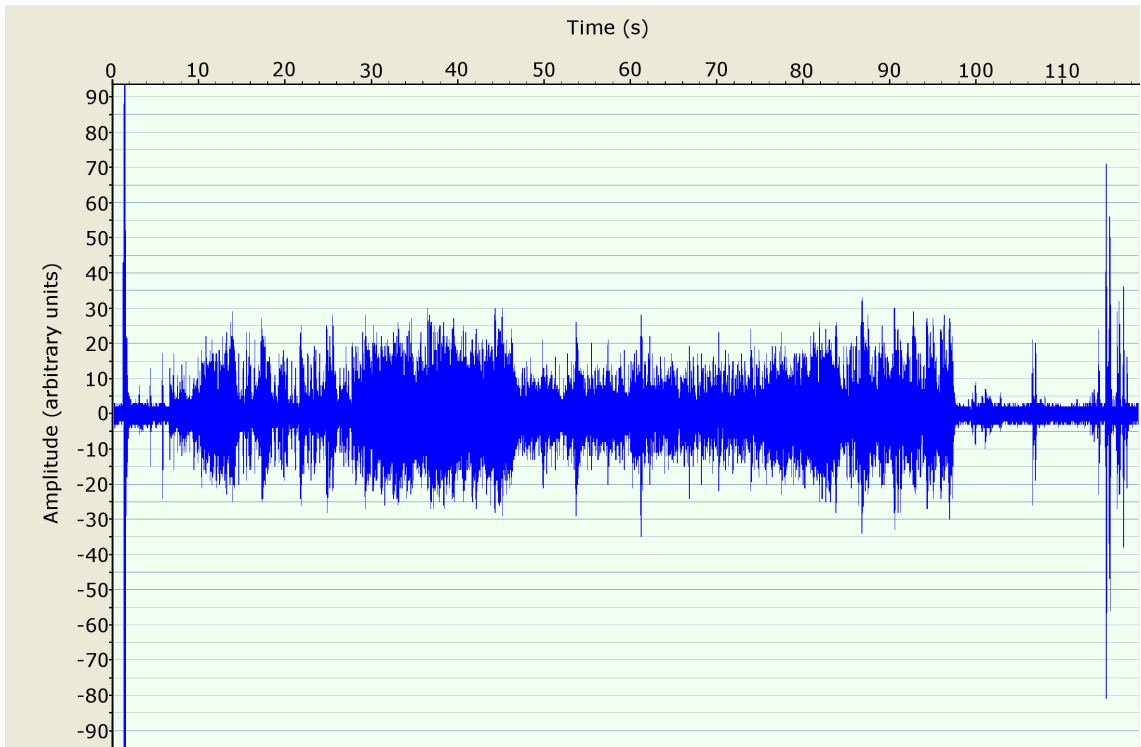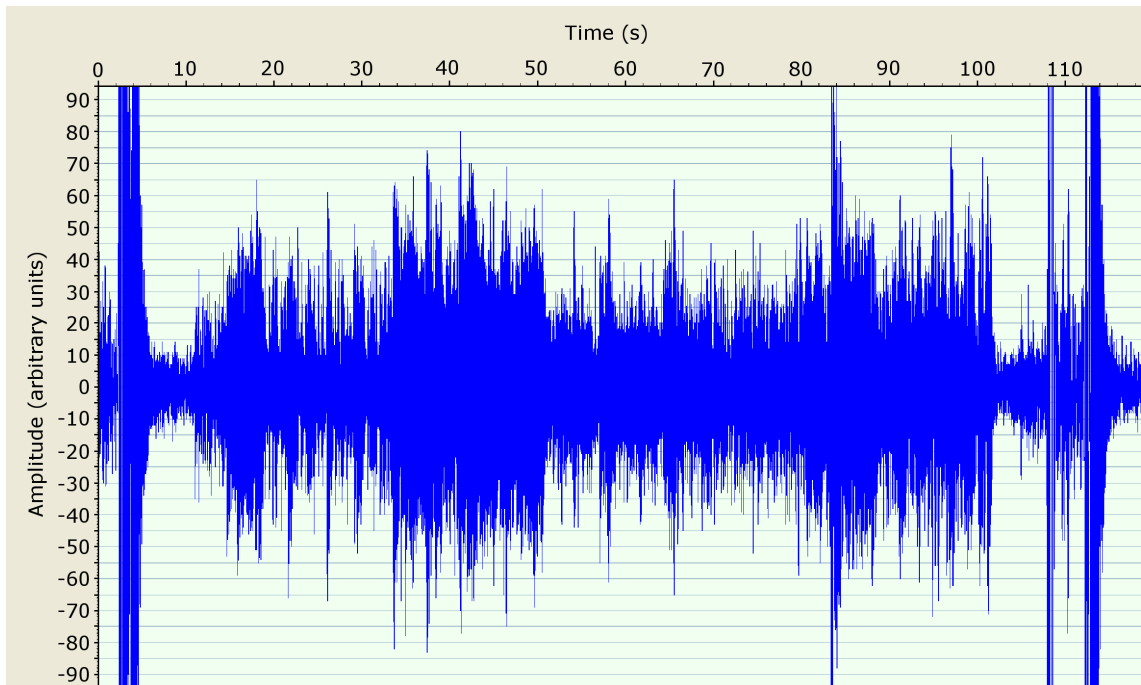
Figure 5.2: microphone 1

Figure 5.3: microphone 2

may have resulted on the way from the computer to the line out or from the line in to the computer. The FR, the Noise level, the Dynamic Range (DR), the Total Harmonic Distortion plus Noise (THD+N), Intermodulation Distortion (IMD), IMD with swept tones and stereo cross-talk (CR) were measured from both sound cards.

The full-scale **FS** input voltage is the input voltage level that causes the A/D converter output to be just equal to full scale. This means that no clipping occurs at positive or negative peaks. This level is defined as a level where total harmonic distortion plus noise (THD+N) is -40dB. In practise, when the amplitude of the 1kHz signal is increased, FS is achieved when the amplitude of 2kHz or another harmonic signal amplitude is found to be -40dB referenced to the 1kHz signal. [83] This happens much before visually observing the clipping of the sine-wave in the time domain as is explained later. The 0dB reference level of the TUT sound card measurements is the FS level. Another reference level was used in the PSHP measurements.

FR represents the effects of the signal frequency to the measured amplitude of the signal. In an ideal case the device measures the constant amplitude of the sine-wave constantly, independent of the frequency of the signal. In such a case a straight line is seen at 0dB all over the frequency domain. If the line deviates from 0dB those frequencies are suppressed or enhanced, depending on whether they are below or above the 0dB levels respectively. In practise, a constant amplitude sine-wave is applied to the input of the measuring instrument. The frequency is swept over the frequencies under interest and each amplitude of a certain frequency is measured. The 0dB reference level is the amplitude of the 1kHz signal.[83, 84]

Dynamic Range (**DR**) gives a ratio of the FS signal level to the RMS noise floor in the presence of a signal. It is like SNR in the presence of the signal, but the term SNR should not be used because the exact definition of SNR is different. The DR is measured while a 1kHz -60dB FS test signal is present. The DR measurement is similar to the THD+N measurement but a smaller amplitude minimizes any large signal non-linearities. [83, 84]

THD+N is the ratio of the amplitude of the harmonics to the 1kHz -3dB FS test signal. The +N means that the noise level which is present in the output signal is also included in the amplitude of the harmonics. If the system has a low distortion, the THD+N and DR measurements give the same kinds of results. If significant distortion occurs in the system, the THD+N value is much worse than the DR value. It is important to notice that converters may produce also other frequencies than harmonics of the test signal and these may be seen in the THD+M measurement.[83, 84]

Intermodulation distortion (**IMD**) is the same kind of phenomenon as THD. In THD the harmonics of the test signal are observed. These harmonic frequencies are

multiples of the fundamental frequency f, i.e the harmonics are 2f, 3f, etc. If there are two test signals f_1 and f_2, in addition to harmonics, second order intermodulation frequencies $f_1 + f_2$ and $f_1 - f_2$ may exist. There are also third order intermodulation frequencies $2f_1 + f_2$, $2f_1 - f_2$, $f_1 + 2f_2$ and $f_1 - 2f_2$. In addition, the intermodulation frequency components have their own harmonic frequencies which makes the phenomenon more complicated. The IMD is measured with two different methods; by two discrete frequencies and by sweeping the tones. The discrete frequencies were 60Hz and 7kHz. The values of IMD for the swept tones were determined at 5kHz, 10kHz and 15kHz. [83, 84, 85]

Crossover (**CR**) is an information leakage from one channel to another, for example from the left stereo channel to the right or vice versa. The -20dB FS test signal is applied to one channel. The other channels are terminated to the signal ground via a 50Ω resistor. The signal in the undriven channel is analyzed in order to test the presence and amplitude of the test signal. The amplitude found in the undriven channel is expressed as a dB ratio to test the signal amplitude. It is recommended that the test signal frequency be altered with steps less than an octave.[83, 84]

The results of the Sound Blaster Audigy 2NX sound card are shown in figures 5.19 to 5.10. The results of the Sound Blaster X-Fi model are presented in figures 5.20 to 5.17. The X-Fi model was measured because the model Audigy 2NX is no longer available for purchase and it was planned that the X-Fi model would replace the Audigy 2NX model in the future. Furthermore, we were interested to see what kind of differences occur between different models from the same manufacturer.

The results are not discussed here in detail, instead the overall performance is evaluated. When the overall summary and the figures are studied it can be said that the minimum undistorted level for the Audigy 2NX is approximately -80dB referenced to the test signal amplitude. This value is determined from the THD+N measurement. The crosstalk level is higher but it does not affect the tracheal sound measurements since the HeLSA uses a mono channel. However, if the measurements are done simultaneously with both HeLSA-microphones for comparison purposes, the crosstalk effect must be noticed. When THD+N and DR measurements were compared it was clearly seen that more harmonics are seen in the THD+N measurements, indicating that at least a moderate distortion exists. Moreover, the frequencies which are not multiples of the test signal and are produced by the converters can be seen in THD as was mentioned by Harris [83].

The X-Fi card has a similar distortion when THD and DR measurements are compared, even though the distortion level is lower than with the Audigy 2NX. When the summaries of the sound card are compared, it can be seen that the Audigy 2NX and X-Fi have a quite similar overall performance while differences do occur between different fields. FR of the Audigy 2NX is slightly better, while noise levels

are excellent in both cards. Similarly the DR is very good for both card models. The worst results for both cards was the THD+N measurement, which was rated by the RMAA software as average. IMD+N was rated very good for both Audigy 2NX and X-Fi. Stereo crosstalk was very good for Audigy 2NX and excellent for X-Fi. The general performance was rated as very good for both models.
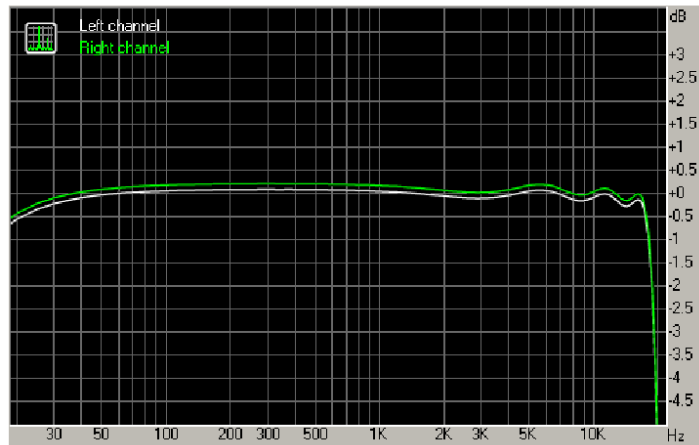
Frequency response

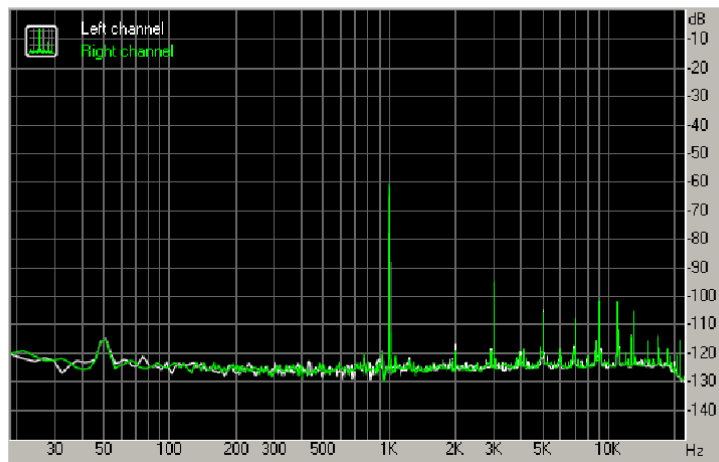Figure 5.4: Frequency response of the Audigy 2NX

Dynamic range

Figure 5.5: Dynamic range of Audigy 2NX

THD + Noise (at -3 dB FS)
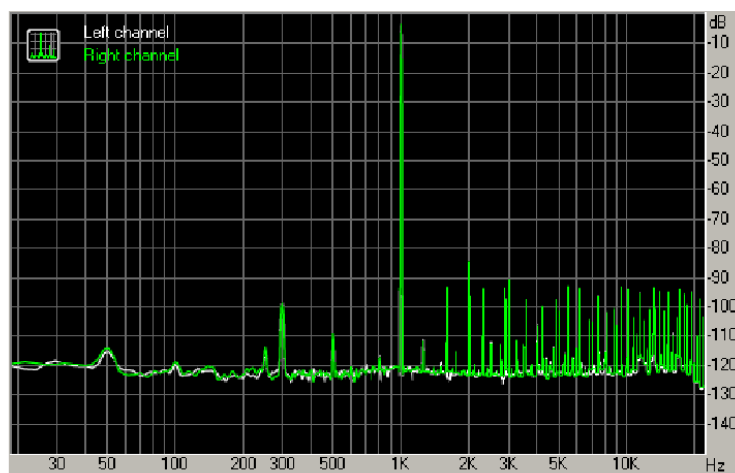
Figure 5.6: Total harmonic distortion and noise of the Audigy 2NX
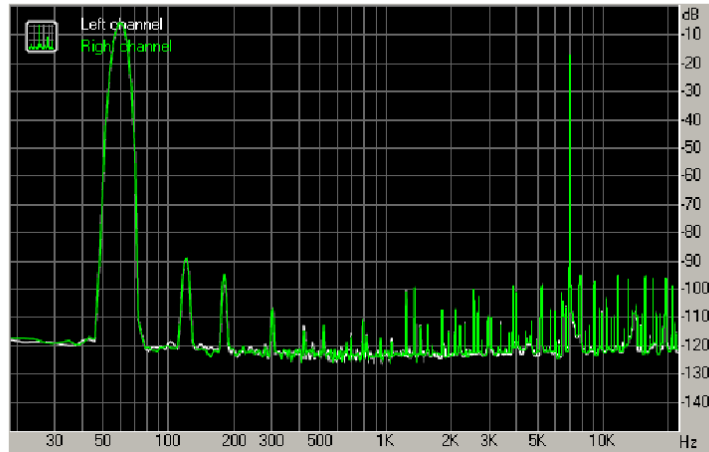
Figure 5.7: Intermodulation distortion of the Audigy 2NX
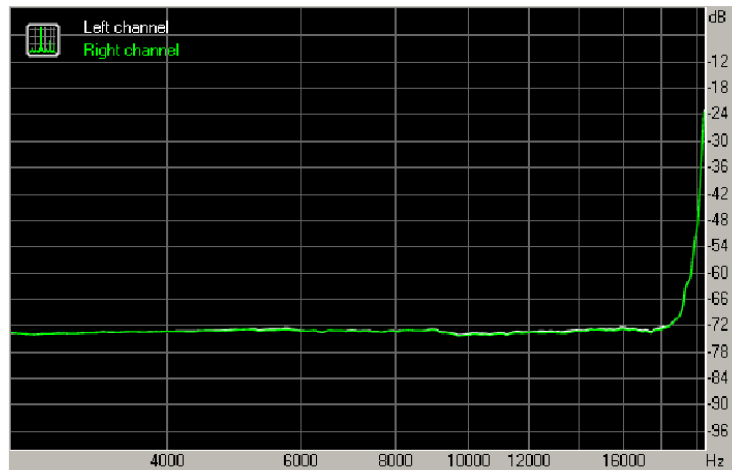

Figure 5.8: Intermodulation distortion with swept tones of the Audigy 2NX

Figure 5.9: Stereo crosstalk of the Audigy 2NX
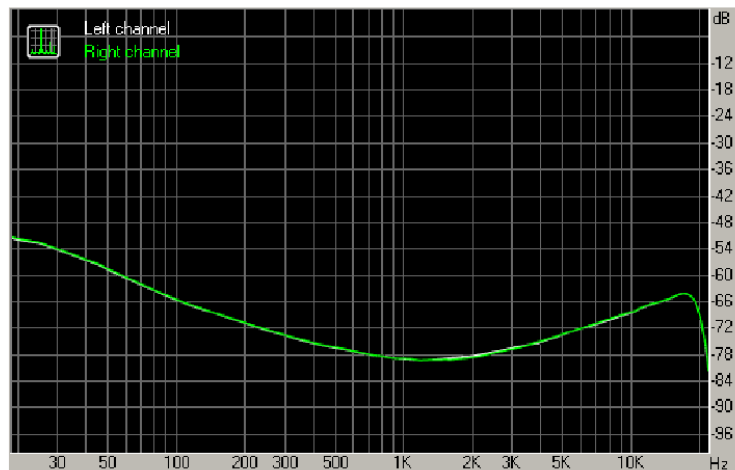
RightMark Audio Analyzer test report

Testing device	USB Sound Blaster Audigy 2 NX
Sampling mode	16-bit, 44 kHz
Interface	MME
Testing chain	External loopback (line-out - line-in)
RMAA Version	6.2.2

20 Hz - 20 kHz filter	ON
Normalize amplitude	ON
Level change	-0.9 dB / -0.8 dB
Mono mode	OFF
Calibration singal, Hz	1000
Polarity	correct/correct

Summary

Frequency response (from 40 Hz to 15 kHz), dB	+0.09, -0.28	Very good
Noise level, dB (A)	-99.8	Excellent
Dynamic range, dB (A)	90.5	Very good
THD, %	0.012	Good
THD + Noise, dB (A)	-74.8	Average
IMD + Noise, %	0.017	Very good
Stereo crosstalk, dB	-78.9	Very good
IMD at 10 kHz, %	0.022	Good
General performance		**Very good**

Figure 5.10: The summary of the measurements with the model Audigy 2NX
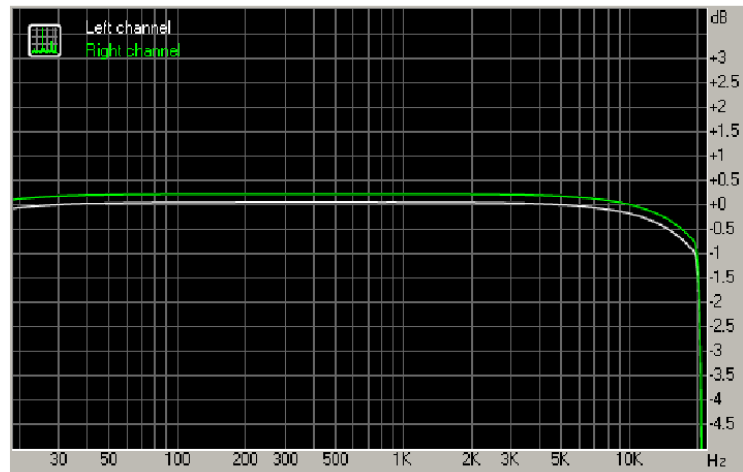
Frequency response

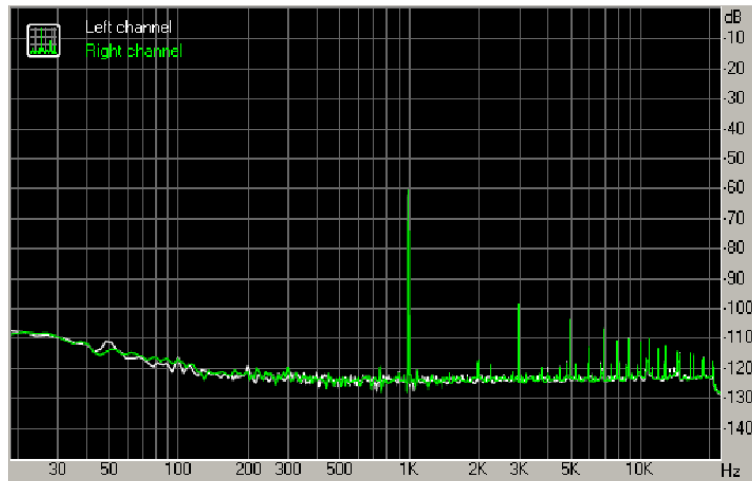Figure 5.11: The frequency response of the X-Fi

Dynamic range

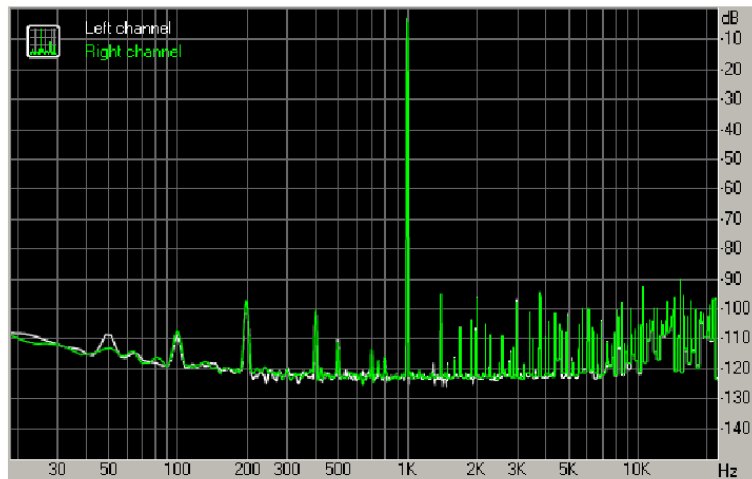Figure 5.12: The dynamic range of X-Fi

THD + Noise (at -3 dB FS)
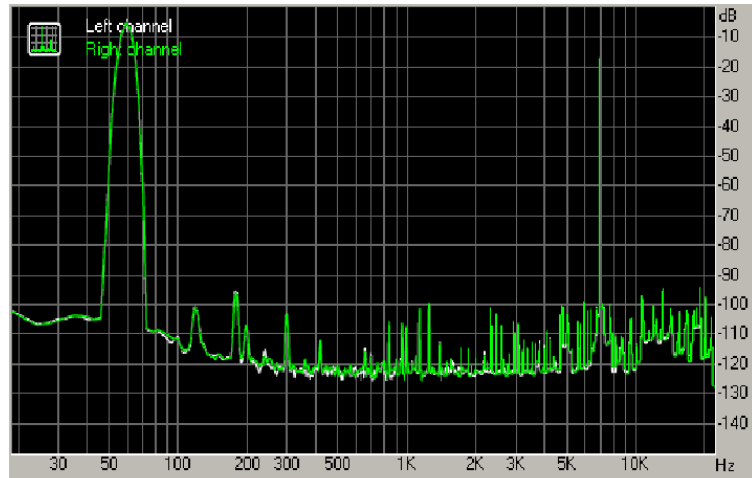

Figure 5.13: The total harmonic distortion and noise of the X-Fi
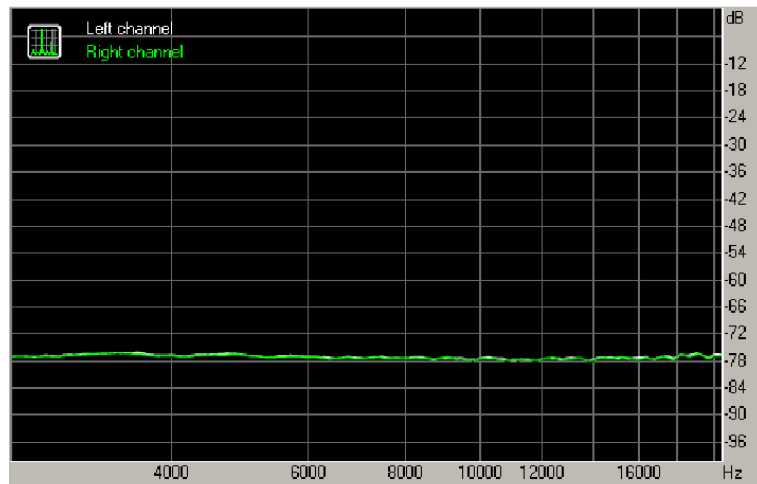
Figure 5.14: Intermodulation distortion of the X-Fi

Figure 5.15: Intermodulation distortion with swept tones of the X-Fi
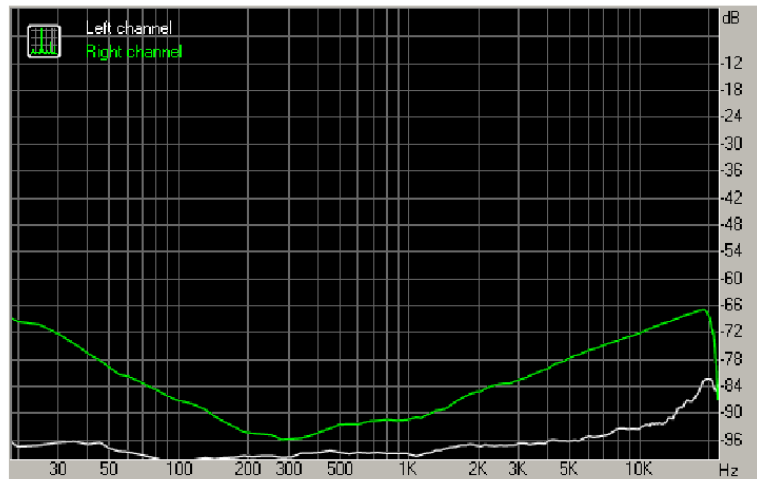

Figure 5.16: Stereo crosstalk of the X-Fi

RightMark Audio Analyzer test report

Testing device	SB X-Fi Surround 5.1
Sampling mode	16-bit, 44 kHz
Interface	MME
Testing chain	External loopback (line-out - line-in)
RMAA Version	6.2.2

20 Hz - 20 kHz filter	ON
Normalize amplitude	ON
Level change	-5.5 dB / -5.3 dB
Mono mode	OFF
Calibration singal, Hz	1000
Polarity	inverted/inverted

Summary

Frequency response (from 40 Hz to 15 kHz), dB	+0.05, -0.51	Good
Noise level, dB (A)	-97.1	Excellent
Dynamic range, dB (A)	90.6	Very good
THD, %	0.0074	Very good
THD + Noise, dB (A)	-78.8	Average
IMD + Noise, %	0.015	Very good
Stereo crosstalk, dB	-95.3	Excellent
IMD at 10 kHz, %	0.014	Very good
General performance		**Very good**

Figure 5.17: The summary of the measurements made by X-Fi

Measurements at PSHP The FR of the sound cards was measured at the Sleep Laboratory of the PSHP. In the measurement setup the signal source was the Thandar TG503 pulse and function generator (Thurlby Thandar Instrument Ltd, Huntington, Cambridgeshire, United Kingdom). The sine-wave voltage signal was transmitted to an oscilloscope and to a sound card. The oscilloscope model Gould 4072 (Gould Electronics, Ohio, USA) was used for monitoring the signal going into the sound card. The open source software Audacity®version 1.2.6. was used for 24bit sound signal recordings.

Discrete frequencies from 10Hz to 5.5kHz of sine-wave signals were used. In the range of 10Hz-100Hz the frequency step was 10Hz, in the range of 100Hz-5kHz step it was 100Hz and above 5kHz a step of 10Hz was used again. According to the specification of reference [83] the amplitude should be set to -20dB from the full scale (FS) of the line-level. The FS should have been determined as a level where the harmonic distortion reaches the level of -40dB. At the PSHP the FS was determined from the sound signal in the time domain when the clipping was observed in the screen of the oscilloscope. This was a rough error and the full scale was determined to be 6V peak to peak when the correct value determined by the engineers at TUT was half of this. However, this erroneous FS level was used and therefore the amplitude used was 0.6V. It was explained in the reference [83] that increasing the amplitude from a -20dB level may produce distortion components that affect the FR measurements. However, this same amplitude was used in all of the measurements done at the PSHP and therefore, the PSHP measurements can be compared with each other as long as it is remembered that there might be errors in the absolute values.

The amplitude of the signals were calculated as an average of the values within 1% of the signal maximum. This method was used because it was observed that as the signal frequencies approached one half of the sampling frequency, the error of the amplitude measurement grew. This amplitude reduction occured because some maximums of the signal were missed when the sampling moment and time of the signal maximum were not the same. This resulted in the dynamical distortion of the signal. During the distortion it was confirmed with a monitoring oscilloscope that the measured signal was sine shaped and no other distortions were observed. An example of this amplitude cut-off is presented in figure 5.18. Without a 1% selection rule this would have reduced the calculated amplitude.

The results of the measurements are presented in frequency-decibel plots, see figures 5.19 and 5.20. The reference level, corresponding to 0dB, is the amplitude at a 1kHz frequency. It can be seen that all the FRs of the sound cards are flat over the used frequency space except the Audigy 2NX with a 11kHz sampling frequency. It is very unlikely that the sound cards would use different hardware with different
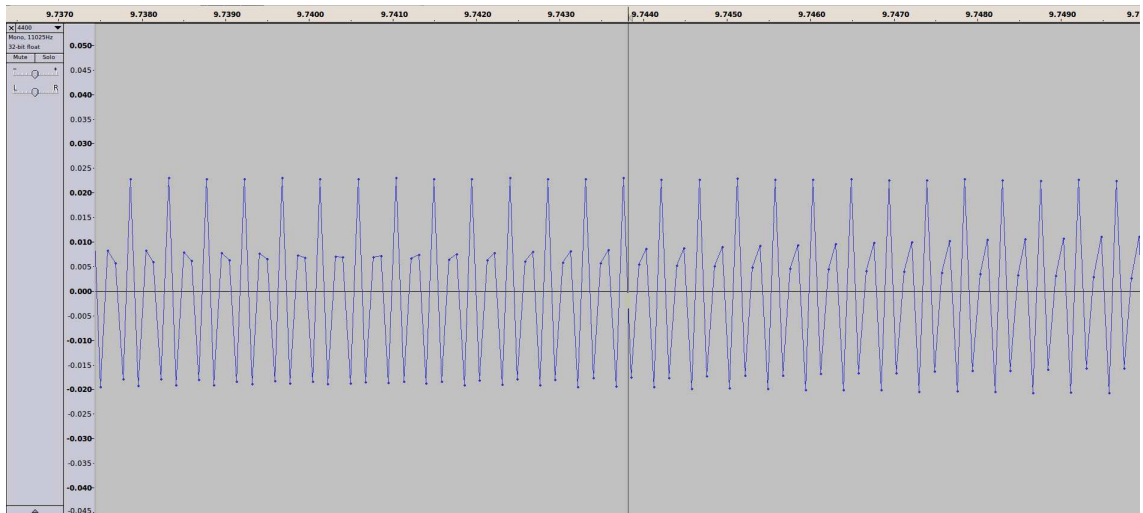
Figure 5.18: Signal distortion when the signal frequency approaches one half of the sampling frequency

sampling frequencies. Therefore, it was concluded that because the FR with the 44kHz sampling rate was flat, the reason for the distortion with the 11kHz rate had to be software based.
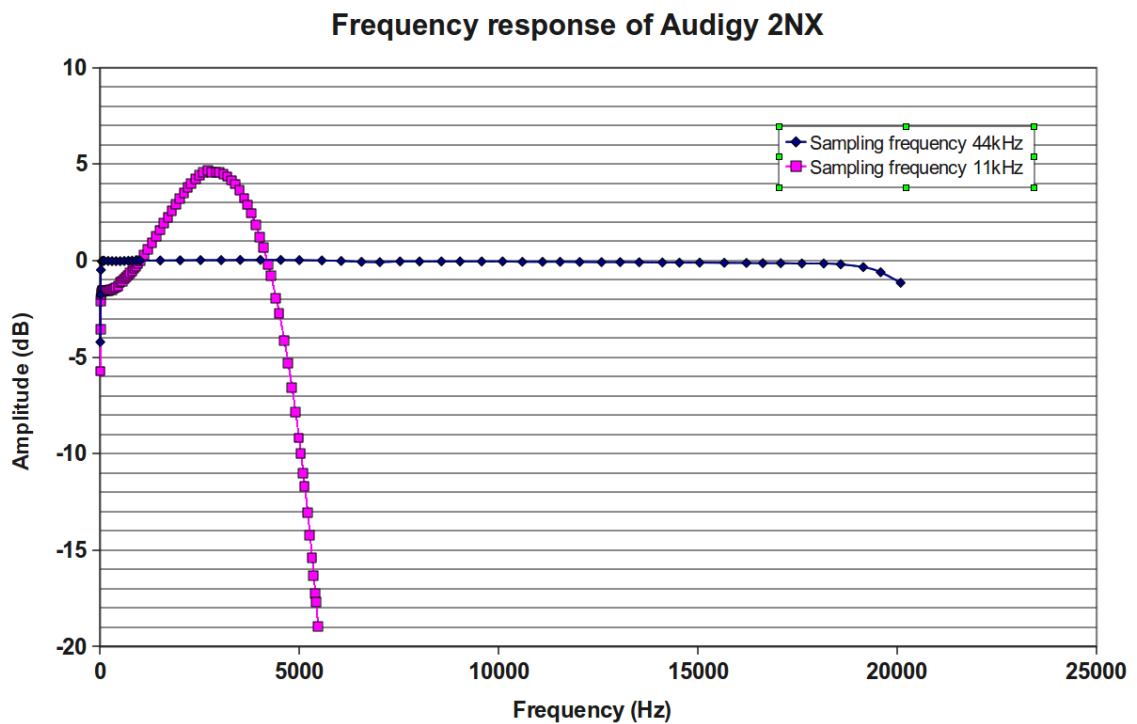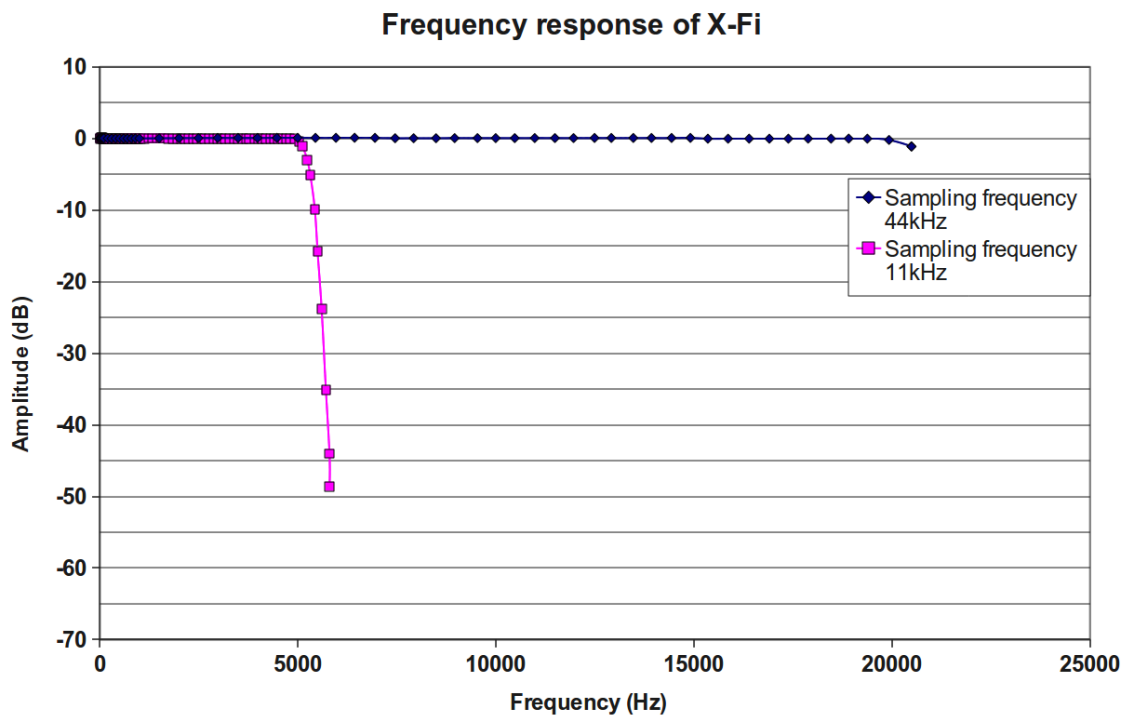

Figure 5.19: Frequency response of the Audigy 2NX

Figure 5.20: Frequency response of the X-Fi

5.0.2 Free field measurements of the microphone

The FR and the THD+N were measured for the HeLSA-microphone in a sound proof chamber at TUT. The measurement software was APWIN made by Audio Precision. The measurement instrument was Audio Precicion, system two, dual domain 2322. The precision microphone 1/2" Brüel & Kjaer model 4189 was used as a reference microphone for the measurements. The FR of the reference microphone is not totally flat either, but this variation is smaller than the precision of the measuring instruments. The microphone under test was placed onto a stable holder and the distance to the loudspeaker was constantly kept at 2m. This same setup was used for all free field measurements of the microphones.

In figure 5.21 the FR of the HeLSA microphone no 1 is presented. The measurements were corrected in a way that the 0dB level is the FR of the reference microphone. This means that any observed deviations from 0dB represent the difference of the FR between the B&K reference microphone and the HeLSA microphone. It can be seen that the response increases in the range of 2kHz to 7kHz. An experienced engineer at TUT reported that this is a normal observation with many different microphone models. However, the positive deviation in the range of 7-10kHz is not natural and it is caused by the external plastic coupler which is used to attach the microphone to the skin. The effects of the plastic coupler were
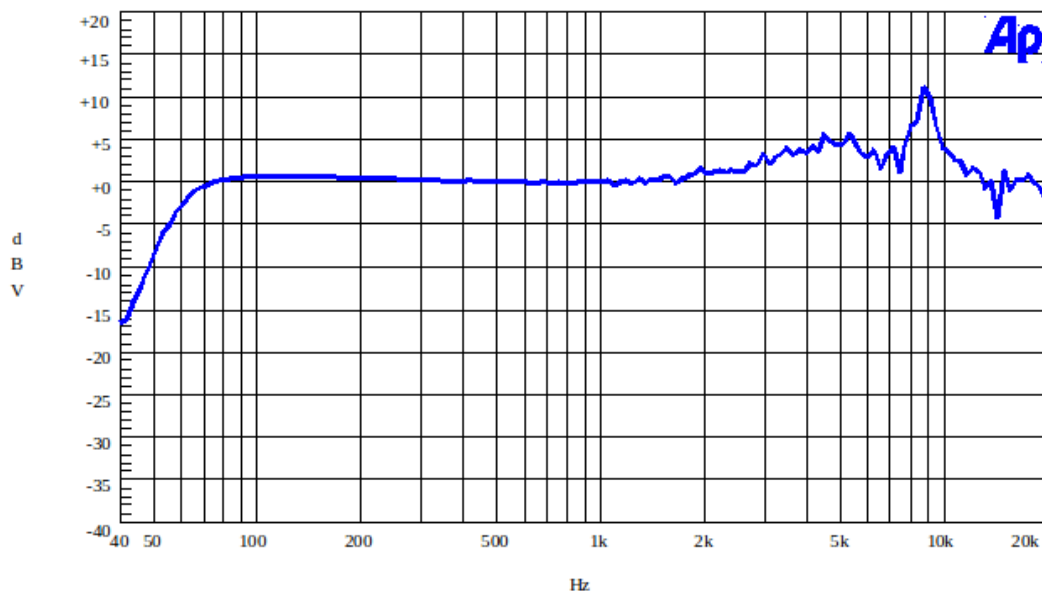

Figure 5.21: Frequency response of the HeLSA microphone

demonstrated in a way that a similar plastic collar as used in the coupler of the HeLSA was attached to the B&K reference microphone. The results of this FR measurement with a modified B&K microphone is presented in figure 5.22.
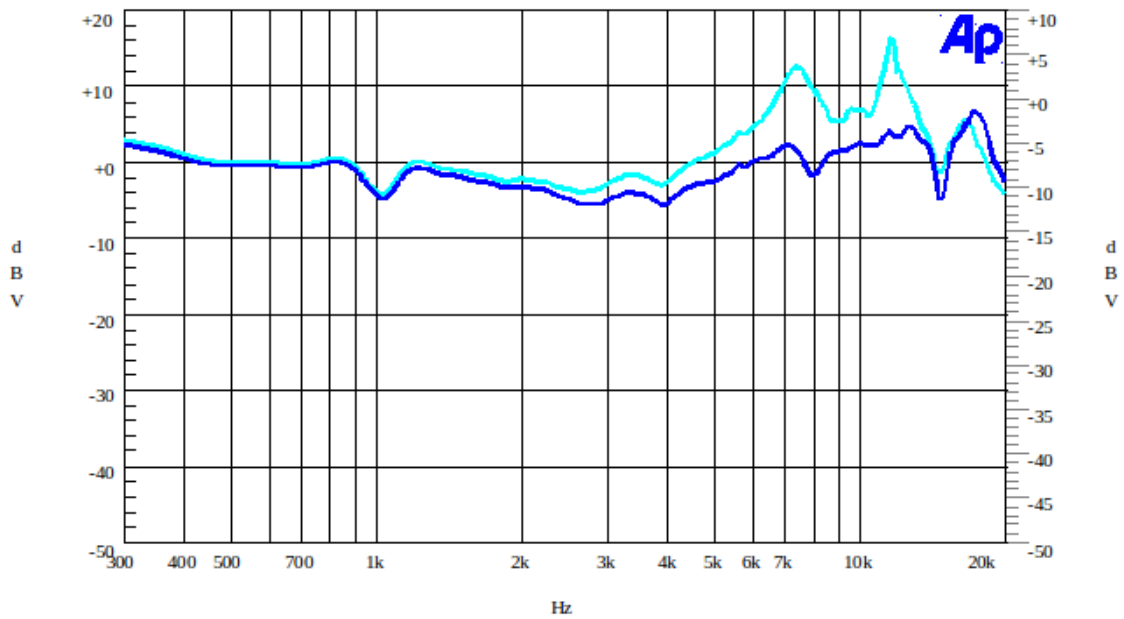
Figure 5.22: Frequency response of the HeLSA microphone

The upper light blue curve is the FR with the collar and the lower dark blue is the FR without the collar. The measured FR is not linearized and therefore the unlinear effects of the loudspeaker are included in the response. There are different y-scales for both measurements. The scale for the light blue curve is at the left hand side and the scale for the dark blue curve is at the right hand side. In figure 5.22 the two spectral peaks are seen in the light blue curve at the frequencies of 7kHz and above 10kHz. When compared to the FR of the HeLSA-microphone the spectral peak is located approximately at 9kHz whereas in the B&K microphone with the collar, the peak is shifted above 10kHz. Even if the 7kHz peak is not observed with the FR of the HeLSA it can be said that the shape of the peak and the frequency ranges are similar to the modified B&K microphone. The explanation for this effect is that the plastic collar acts as an aperture for the sound waves arriving to the microphone and this effect has a dependency to the frequency of the sound waves. It seems that the dimensions of the collar allows certain wavelengths to affect the microphone more easily than the others.

The THD was measured for both HeLSA microphones. The sound pressure level (SPL) used in the measurements was 1Pa. The loudspeaker was driven with a 1kHz sine-wave signal and a spectrum analyzer of the Audio Precicion measuring instrument was used. In figures 5.23 and 5.24 the results of the HeLSA-microphones no. 1 and no. 2 are presented, respectively.

The light blue curve is the curve of the HeLSA-microphone and the dark blue is the curve of the B&K reference microphone. Interestingly, it was observed that both
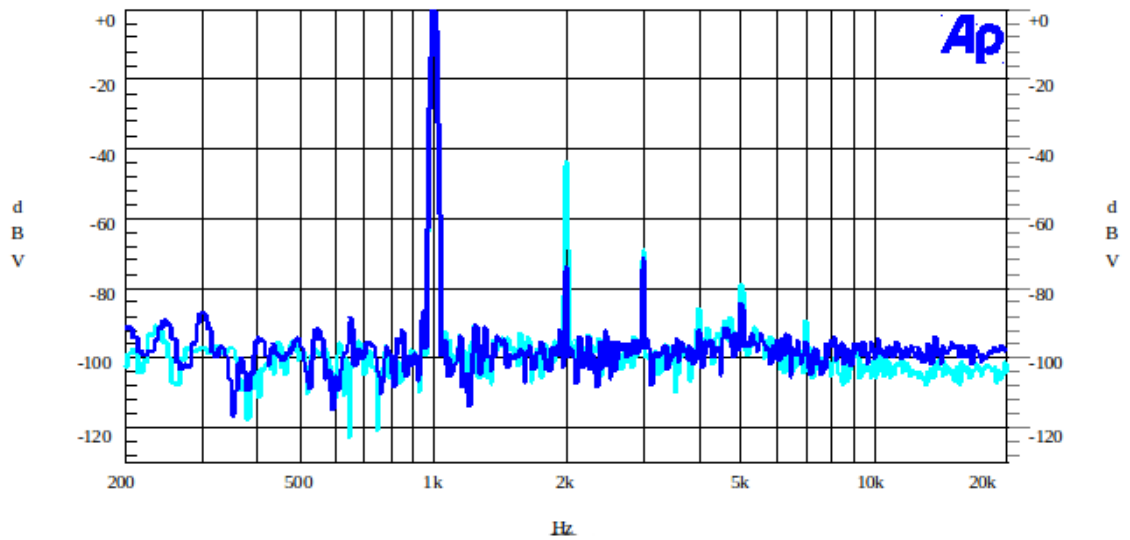
Figure 5.23: Total harmonic distortion of the HeLSA microphone no. 1
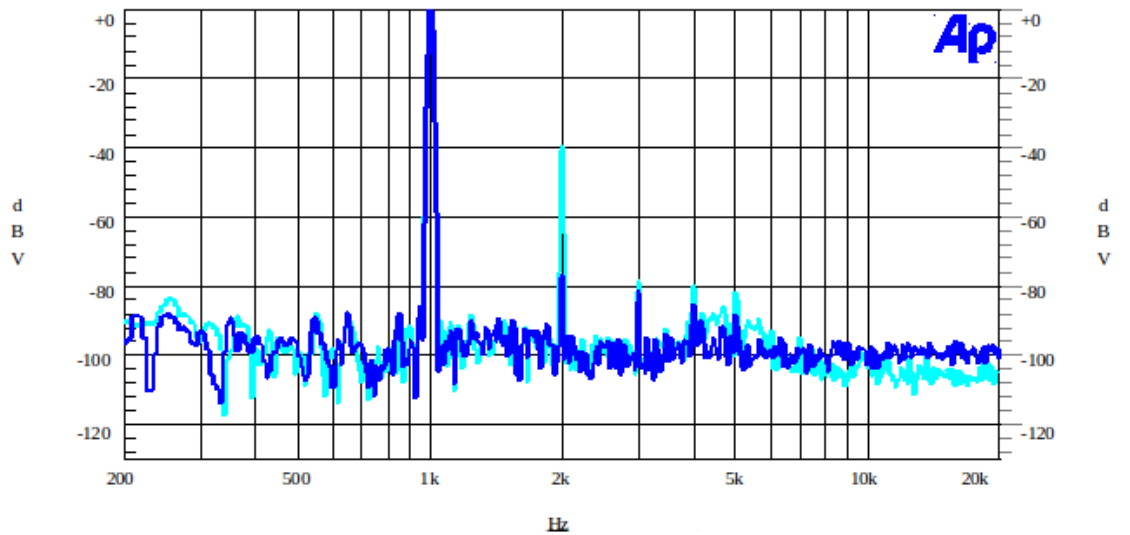

Figure 5.24: Total harmonic distortion of the HeLSA microphone no. 2

HeLSA-microphones produced a strong second harmonic at the frequency of 2kHz. The reason for this is unknown. It might be the microphone amplifier that is used in the HeLSA. However, clear harmonic behaviour could not be found in the measured tracheal sounds. If the second harmonic is produced by the amplifier it should be seen with both free field and a contact measurements. When, for example figure 5.29 is observed, it can be seen that the amplitude of the tracheal sound decreases in the range of 1kHz to 2kHz, while the amplitude increases in the range of 500Hz to 1kHz. If the sound frequencies in the range of 500Hz to 1000Hz produced second harmonics in the amplifier we should see an increase in the amplitude also in the range of 1kHz to 2kHz. Instead, a decrease in the amplitude is observed, and therefore it seems that the amplifier is not causing the distortion. This is supported also by the fact that the distortion increased with the increasing sound pressure level. This means that if the distortion had been caused by the amplifier, in addition to an absolute increase of the amplitude in the range of 1kHz to 2kHz, there should have also been a relative amplitude increase. This would happen because the higher amplitude of the 'true' sound at 1kHz would produce a relatively higher distortion than the lower amplitude at 500Hz.

In the free field measurements the harmonic distortion was observed the following way. When the sound pressure level caused a 170mV RMS voltage to the amplifier output the distortion was 1%. When the SPL level was increased to 388mV RMS, the distortion was 2.5% and for the SPL level equal to 1V RMS, the distortion was 6.8%. The 3rd and 5th harmonics are produced by the loudspeaker. However, the second harmonic is quite high. For reference it can be mentioned that when a sine-wave is distorted to the square wave the second harmonic is 30%, but in that case other harmonic frequencies are observed as well. No significant difference in the THD was observed between HeLSA-microphone no. 1 and no. 2.

5.0.3 Measurements with an audiometer

The Medelec ST 10 Sensor audiometer (Surrey, United Kingdom) was used in the measurements. The audiometer is used to determine the hearing level (HL) of human beings. The system contains a main unit and headphones. The reference level of 0dB can be presented as a curve which represents the ideal frequency response of the ear. The device produces sinusoidal sounds with discrete frequencies which can be heard from the headphones. If a person's hearing is reduced it can be expressed as HL +30dB at 4000Hz. This means that sound must be 30dB louder than in the ideal case. This same procedure was used for the HeLSA-microphone. Two constant power levels were used, the 65dB HL and 80dB HL. Microphone no. 1 and no. 2 were both connected to the sound card at the same time. This was done by using an adapter cable which connected microphone 1 to the left stereochannel and microphone 2 to the right channel. However, simultaneous measurements with both microphones were not done. An advantage was that it was possible to prepare the measurement setup at once and the measured microphone was chosen from the main unit of the audiometer and from the registering computer. The electret microphone of the HeLSA is permanently mounted into a plastic coupler. During the measurements the plastic coupler was mounted to the earmuffs of the headphones. The plastic coupler fit the loudspeaker well, opening in the center of the earmuffs by chance. By gently stretching the muff material it was possible to mount the coupler to the headphones. As a result, there was an air proof cavity between the diaphragm of the loudspeaker and the microphone. Discrete frequencies up to 4000Hz were measured. Higher frequencies were not measured because the tracheal sounds observed are below 4000Hz. During the measurements the 44kHz sampling rate was used with the Audigy 2NX sound card. As before, the Audacity-software was used on a laptop PC for the sound recordings. The Audacity records the data as a percentual value from the FS in a way that the values were in a range of -0.5 to 0.5. The data was exported from the Audacity in a wav format. The wav-files were read by Matlab (version 7.0.4) and the absolute values of the signal were calculated. After that, the average of the signal was calculated from the measurement points in the range of 200000-700000. The range was selected in this way in order to prevent the distortion caused by the background noise at the beginning and at the end of the recordings. The background noise existed, because the signal was switched on and off while the recording was on.

The measurements were done at three separate occasions. At the first time both 65dB and 80dB HL levels were used. Discrete frequencies of 125Hz, 250Hz, 500Hz, 750Hz, 1000Hz, 1500Hz, 2000Hz, 2500Hz, 3000Hz and 4000Hz were measured. During the measurements it was observed that the two microphones had different fre-

quency response characteristics. The same kinds of results were achieved for the HL with 85dB. The shapes of the curves remained the same and the absolute amplitude increased, which was predictable. Because there were no large differences between the microphones in the freefield measurements, it was concluded that the reason for this may have been the bad positioning of the microphones (like tilting) when they were mounted onto the headphones. The measurements that followed were done immediately after the first ones, but before them the position of the microphones was fixed. The third set of measurements were done two days after the second set. The purpose was to make a quick check that the results of the second set of measurements could be reproduced at least once. In the third set the frequencies of 750Hz, 1500Hz and 2500Hz were not measured at all. The results of the second and third sets of measurements are showed in figures 5.25 and 5.26. The averages of both microphones with frequencies common to measurements 2 and measurements 3 were calculated from the results. The solid line represents the fitted curve of the average measurements.

Figure 5.25: The average of the second and third sets of measurements while the SPL is 65dB over hearing level. The continuous line is fitted to the average values.

As can be seen from the measurements, the shapes of the curves were almost identical to the levels of 65dB and 80dB. The absolute level increases when sound pressure is increased, as can be seen by comparing figures 5.25 and 5.26.
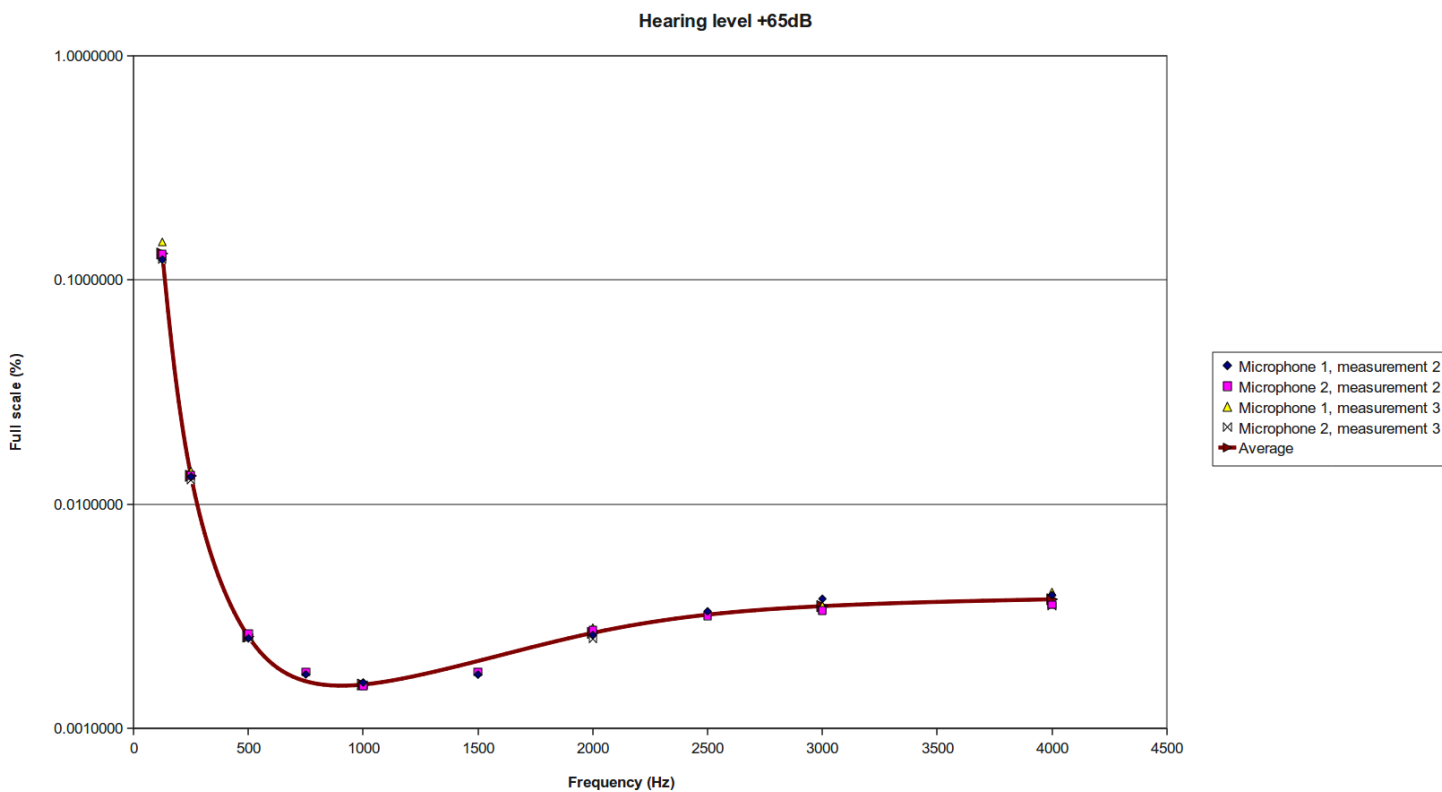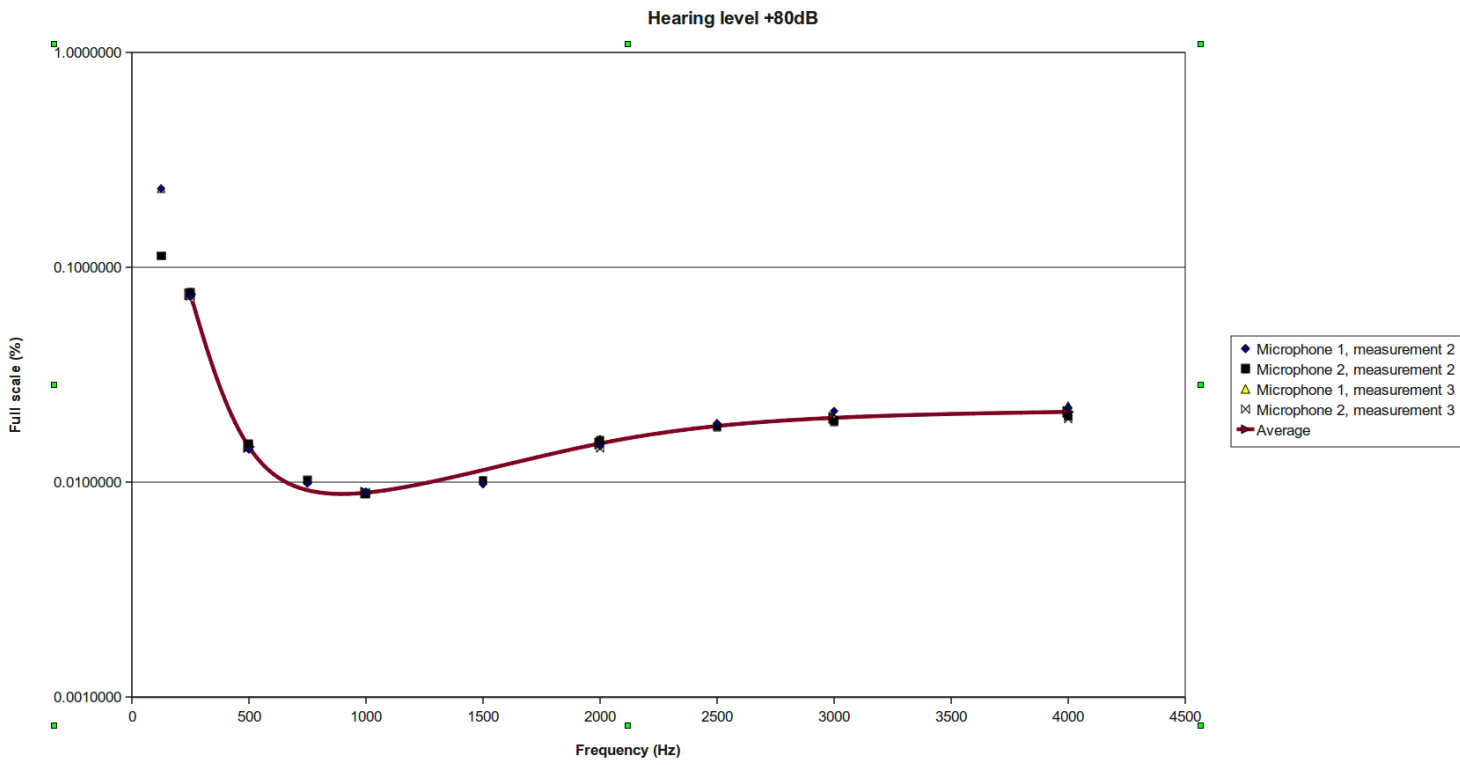
Figure 5.26: The average of the second and third sets of measurements while the SPL is 80dB over hearing level. The continuous line is fitted to the average values.

5.1   Gain and amplitude resolution of the HeLSA

At the Sleep Laboratory of the PSHP, clinical tracheal sounds are recorded with a polysomnography software (Somnologica). Figures 5.2 and 5.3) are screen captures from the Somnologica window. The signal amplitude is misleadingly named kV, because the amplitude calibration is not performed in the software and at the same time the signal type is set to be voltage. The correct unit of the amplitude axis would be kbit. The software is a 16bit measuring system and therefore the maximum bit value is $2^{16} = 65536$. When this value is divided into a positive and a negative side on the amplitude scale, the maximum value is 32768 and the minimum value is -32767. The maximum voltage signal of the line level of the Audigy 2NX is 3V from peak to peak. This means that the bit resolution of the software is $3.0V/65536 = 457.8\mu V/$bit. The correct voltage value for the y-scale in the micro volts in the Somnologica signal is achieved by multiplying the bit value by 457.8 micro volts. The bit resolution of the sound card is higher, because 24bit analog to digital (A/D) converters are used. However, the software reduces it to 16bits. The output of the amplifier is 40mV/Pa and if the FS of the sound card is reached it is equivalent to the sound pressure level of $\frac{1.5V}{40mV/Pa} = 37.5 Pa$. When converted to decibels it is 125.5 dB (SPL), which is the maximum sound pressure level the HeLSA system can record without clipping.

When the microphones were sent to the manufacturer for repairing, the headroom issue was also studied. If the gain of the microphone amplifier is too great the signal will clip. Clipping means that the voltage range of the sound card line in is exceeded. In the case of the Audigy 2NX this limit is 3V AC from peak to peak. Normally the amplifier gain is fixed to 40mV/Pa. For test purposes, the manufacturer added a switch which enabled the gain to be adjusted to 100mV/Pa for microphone no. 1 and to 200mV/Pa for microphone no. 2. Figure 5.27 shows the difference in the gain values of 40mV/Pa and 200mV/Pa. It can be seen that with the value of 200mV/Pa the signal clipped during the artificial snoring of the author of this thesis. The level of 100mV/Pa was occasionally too high. Therefore, it was estimated that the optimum gain would have been 70mV/Pa. The gain was not changed to this value eventually. The reason for this was that the inprovement of the bit resolution of the amplitude axis would have been quite small. As a result, only a small portion of the "bit depth" of the sound card is used during normal breathing. The rest of the bits are reserved to louder sounds such as snoring. The recommendations for the headroom were presented in references [86] and [87].
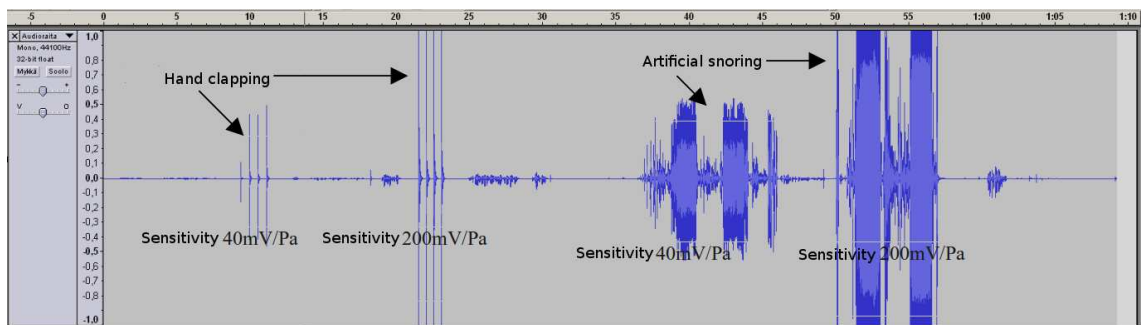
Figure 5.27: Different sensitivities 40mV/Pa and 200mV/Pa.

5.2   Tracheal breath sound (TBS) measurements

Three different measurement setups were done, in which the breathing pathway was altered. Oral, nasal and oronasal breathing were studied separately. The study subject (i.e. the author) was awake in the supine position in all measurements and the HeLSA microphone was fixed to the sternal notch with adhesive tape. The breathing was spontaneous free breathing without targeted flow-rates. The recording room was quiet, but not sound proof. At first the TBS were recorded while mouth breathing was applied. The nostrils were sealed with the tape to prevent leaking through the nose. In the next phase the breathing was done through the nose and the mouth was sealed with the tape. Finally, oronasal (mixed) breathing was recorded. The sound during breath hold was recorded in all the breathing setups. The sounds were recorded with the Audigy 2NX sound card. Both the 11kHz and 44kHz sampling rates were used. The Audacity software was used with the laptop in order to capture and save the sound on the hard disk. For analysis the wav-files were imported to the Matlab v.7.0.4 software. Data from the expirations were selected manually with the help of a graphical presentation of the breath sound signal. The Pwelch-function was applied to the expiration data segments. The Pwelch-function was adjusted to calculate the 2048 point FFT from every expiratory data segment. The Hanning window with a 50% overlap was used. As a result, the power spectrum of each expiration was obtained. Similarly the spectrum of breath holds was calculated. Finally, the average of the expiration spectra was calculated individually for oral, nasal and oronasal breathing.

The power spectra of the expirations are presented in figures 5.28 and 5.29 and 5.30.

In the spectrum plot the amplitude was presented as a logarithmic scale whereas the frequency scale was linear (semilogarithmic spectrum). It was observed that the spectrum was the same kind as was presented by [2]. There was an attenuating curve above 600Hz. Some shallow peaks were observed. This has been explained to be caused by the resonances of the respiratory tract[32]. Figure 5.28 presents the power spectrum of nasal expiratory sounds with different sound amplitudes. The measurements were made with the Audigy 2NX sound card and a 11kHz sampling frequency was used. Around 3.7kHz there exist two spectral peaks which were identified to be artefacts. Those peaks were related to the 11kHz sampling frequency. In figure 5.29 a 44kHz sampling frequency was used and peaks were not detected. Similarly, when the X-Fi card was used with a 11kHz sampling frequency, those artefact peaks were not detected in figure 5.30. This supports the previously noticed distortions in FR measurements of the Audigy sound card when the 11kHz sampling frequency was used.
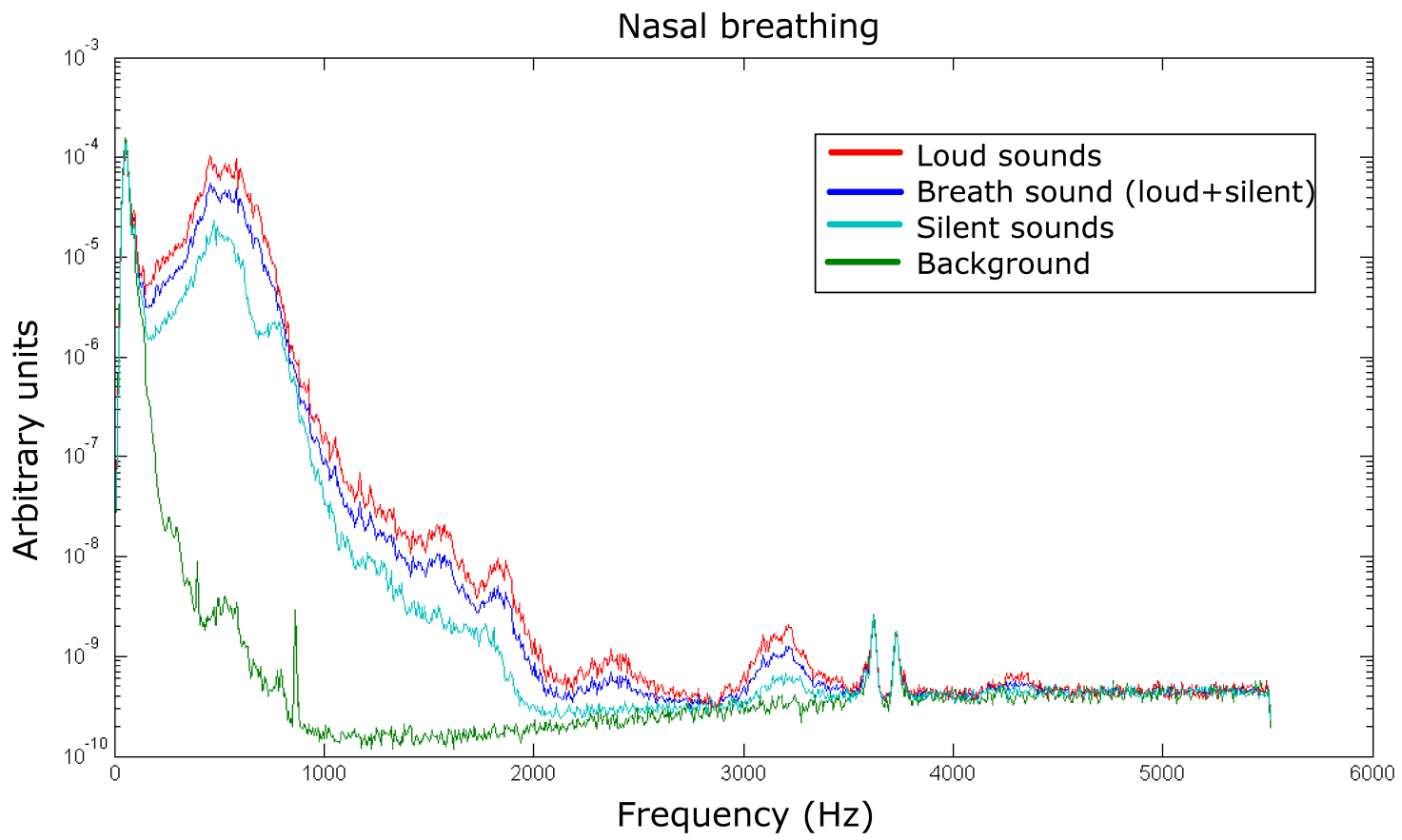
Figure 5.28: Power spectrum of nasal breathing. Difference between loud and silent breathing.
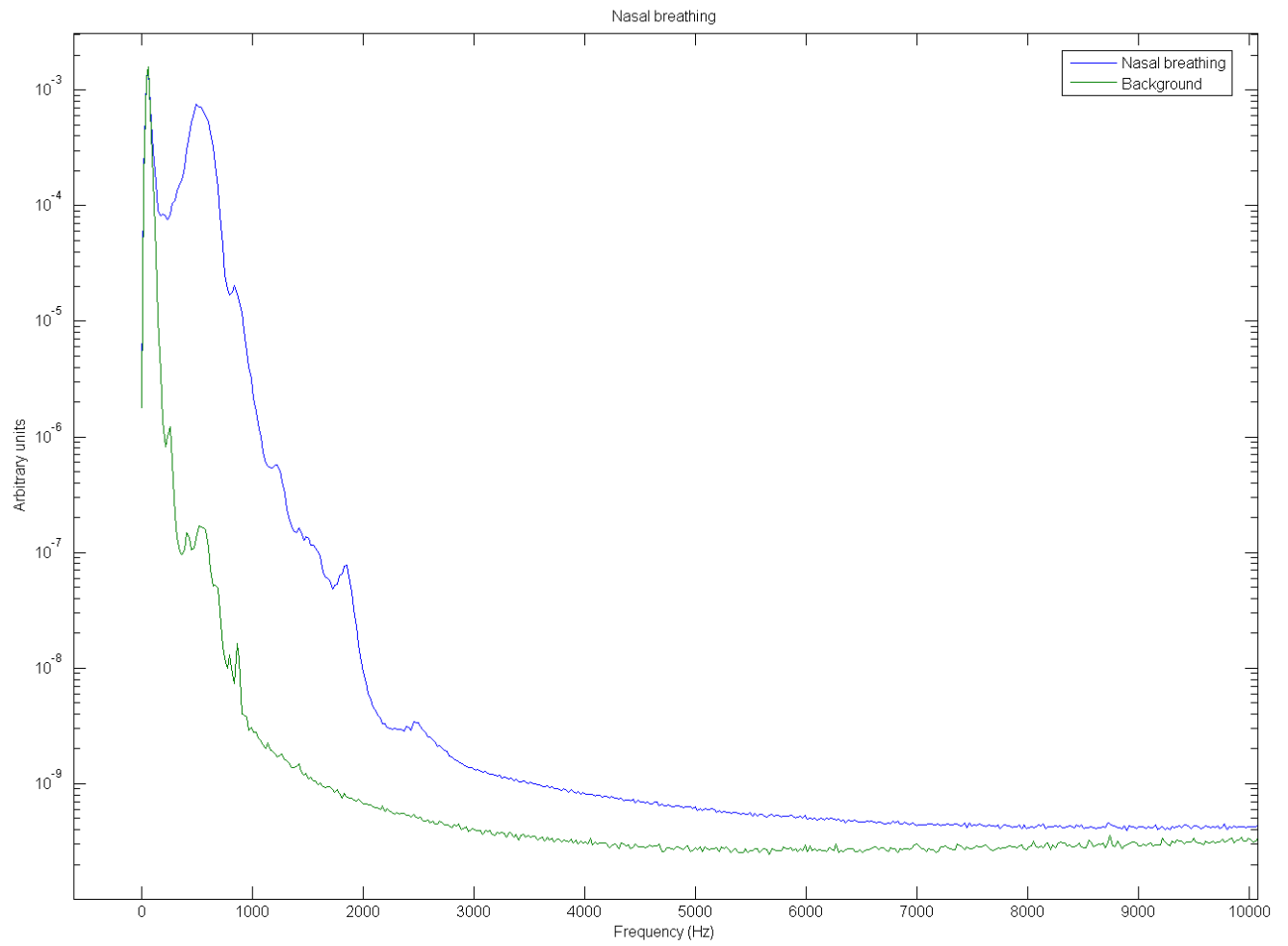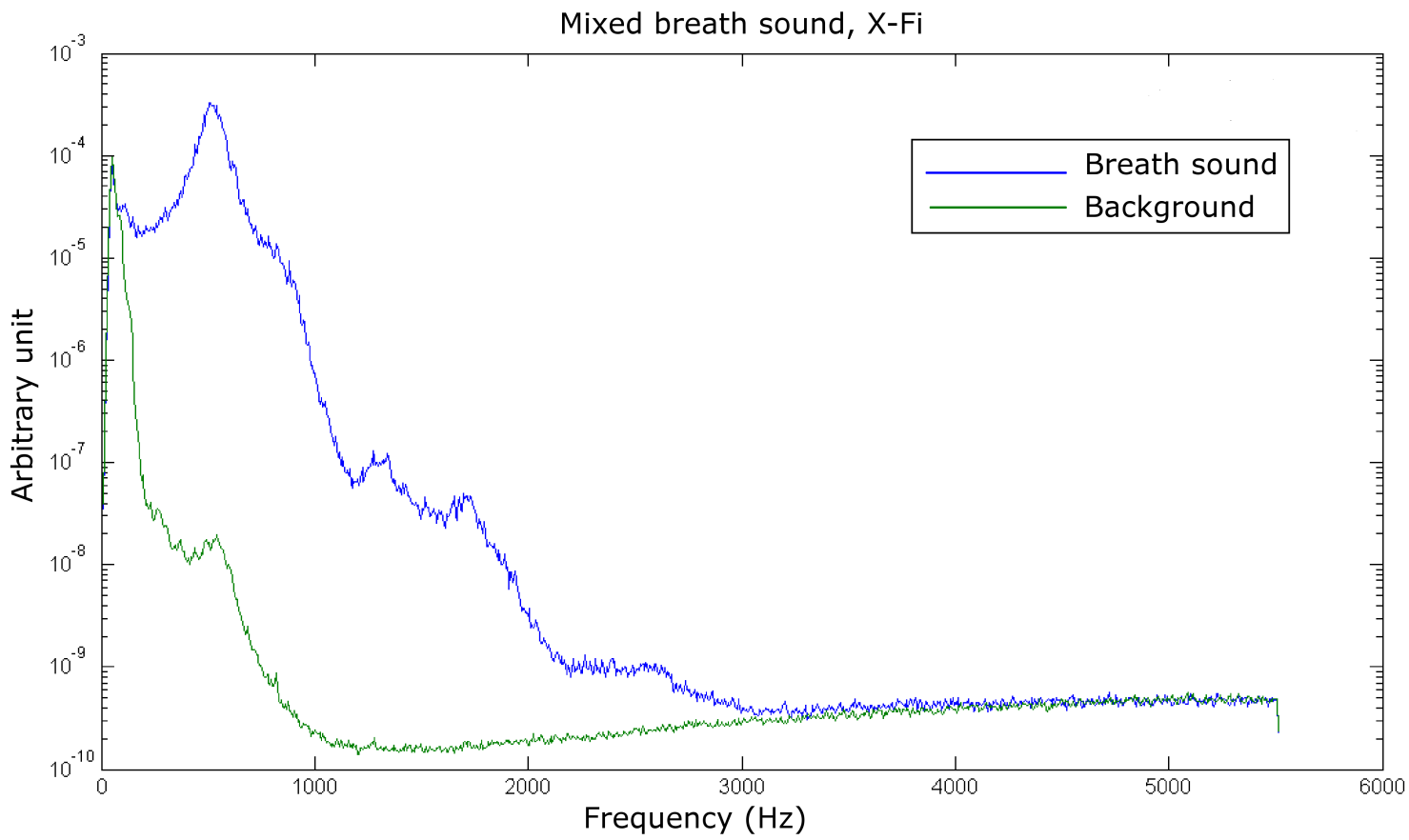
Figure 5.29: Nasal breathing. Sampling rate 44kHz.

Figure 5.30: Oronasal breathing. X-Fi sound card and 11kHz sampling rate used in measurement. Artefact spikes could not be found.

It was observed that when the amplitude of the sound increased, the overall amplitude of the power spectrum increases as well. The intersection point of the background (breath hold) and the sound spectrum were increased at the same time, which means that the highest observed frequency increased when the sound amplitude increased. This is in agreement with the observations of [34].

The base line of the spectra of the breath sounds and backgrounds was tilted. The trend of the tilt seemed to increase with the frequency. The reason for the baseline tilt is not understood. Tilt occured independently of the sampling frequency and the sound card model.

No clear differences between oral, nasal and oronasal(mixed) breathing could be detected from the spectrum. However, too few and too different an amount of expirations were included in the breathsound spectrum. Therefore, detailed physiological conclusions should not be made based on these measurements. In addition, different amounts of data was used for the background and breath sound spectrum calculations. This increases the uncertainty of the curves even more.
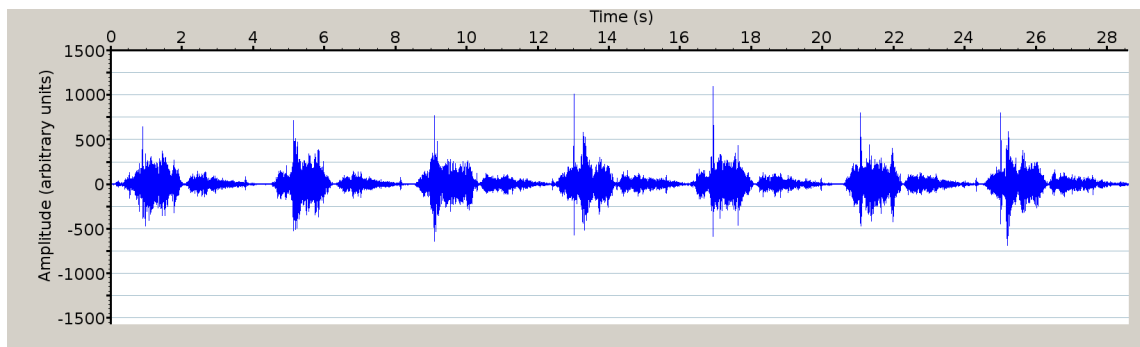
Figure 5.31: Period when inspirations were louder than expirations.
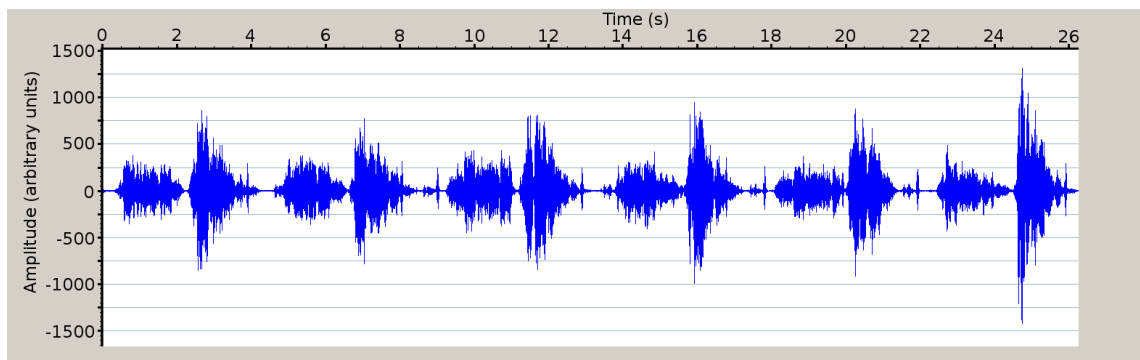
Figure 5.32: Period when expirations were louder than inspirations.

Summary The results of the quality measurements were presented. Some distortions occured in the THD measurements of the microphone-amplifier combination. In addition to these free field measurements, it was discovered that the Audigy 2NX has distortion in the FR when a 11kHz sampling frequency was used.

A calibration measurement setup for the HeLSA microphone was developed. The method is based on audiometry measurements. The sensitivity of the measurement setup for a microphone-headphone attachment was observed. These calibration measurements are not well suited for a general contact sensor comparison. Instead it can be used for a quality check of the HeLSA microphone.

Expiratory breath sounds of the author were recorded with the HeLSA. The power spectrum obtained had similar characteristics to the power spectra presented in the review chapters of this thesis.

6. DISCUSSION

6.1 Tracheal sound sensors

The acoustic impedance of soft tissue has been a topic of research since the 1960's when the method of phonocardiography was developed. Several studies were made, and soon it was realized that when the acoustic impedance is measured, the measuring sensor affects the results when in contact with the tissue.[9, 10] Researchers have discovered that for optimal performance the acoustic impedance of the sensor and the tissue must be matched.[9, 10] This impedance matching proved to be important also with tracheal breath sound sensors. It was presented in the previous literary reviews that the design of the air cavity of the air coupled microphone has significant effects on the sensor performance.[3, 17, 24] Especially the depth of the air cavity affects the measured tracheal sound spectrum. It was also pointed out that venting the cavity with a needle usually reduces the performance of the sensor especially with high frequencies. The measurements showed that the cavity diameter had only a small effect on the sensor performance. However, the mathematical model done by the same authors clearly showed that the surface area and the circumference of the cavity have an effect on the acoustic impedance. The different results of the cavity diameter between the measurements and the mathematical model was explained with the fact that in the measurements the cavity volume was changing at the same time that the diameter was altered. Overall, it can be said that the bad design of the cavity damped the high frequencies in the sound spectrum. This same effect has been observed with phonocardiograph microphones in the 1960's.[9, 10]

There was less variation in the performance of different piezoelectric accelerometers, but clearly some differences were observed. This implies that the design of the accelerometer is important and the best performing accelerometers were specially designed to be used as breath sound sensors (PPG accelerometers).[27] A common issue for both accelerometers and air-coupled microphones is the mass loading effect. The greater the sensor mass, the more it affects the local vibrations of the skin and thereby changes the observed signal. It was also shown that external noise may corrupt the recorded signal.[27] Therefore, a silent recording environment should be used. Air-coupled microphones were proven to be more sensitive to external noise than accelerometers. However, the effects of the external noise in the air-coupled microphones can be reduced with specially designed sound protectors.[27]

It seems that in general, a properly designed accelerometer is better for recording tracheal sounds than an air coupled electret microphone. The air coupled electret microphone is cheap and if the cavity is properly designed the acoustic impedance matching between the chest wall and the microphone diaphragm can be quite good. For impedance matching, the vibrations of the skin have to be transferred to the oscillations of the gas inside the coupler. This transformation cannot be 100% efficient and therefore, a part of the acoustic energy is lost. In contrast, accelerometers can measure the tissue vibrations as acceleration without energy conversions. When the sound spectrum is analyzed it can be seen that the advantages of the accelerometer become more apparent when the sound frequency increases.[22] This difference might be important because some studies imply that high frequencies of tracheal sound contain important information even if their power level is low.[29] In fact, the accelerometer model used for respiratory sound measurements is designed with a frequency response up to 6kHz[88]. Ultrasonic gel has not been applied between the skin and the accelerometers since the 1960's even if it was pointed out that this increased the amplitude of the phonocardiograms.[6] The effects of ultrasonic gel would be interesting to see when applied to tracheal sound accelerometers.

In sleep research, contact microphones have some advantages over external microphones. The source of the sound stays quite fixed when compared to external microphones. External microphones in free field have high sensitivity, but when the patient moves and changes his position during sleep, the position of the sound source and the acoustic field detected by the external microphone change. In addition, some tissue conductive sounds cannot be detected with external microphones. On the other hand, the sensitivity and the available frequency range of the external microphones are much better than with contact microphones.

In this research, the effects of absorption and scattering were not studied. This is mainly because they are phenomena which cannot be affected by the design of transducers and because a large intersubject variability exists. The impedance of the tissue is more important from the point of view of a transducer, because the impedance of the transducer can be coupled with the impedance of the tissue. Sound absorption and scattering are presented by Kaniusas[7] and Morse[5].

6.2   A review of respiratory sounds

A literature review showed that tracheal breath sounds were found to be characterized by an attenuating slope in frequencies above 500Hz.[45] The y-scale of the power spectrum plots has been linear or logarithmic. It seems that the choice is done depending on whether the authors are interested in tracheal sound frequencies above 1000Hz or whether they are not. Aeroacoustic sound production due to a turbulence flow in a tube was reviewed in detail. As a result, the shape of the breath

sound spectrum can be explained with the characters of single vortices and their convection through the constrictions in a respiratory tract.[37] Shallow resonance peaks caused by the geometry of the airways were found on the attenuation slope. It is important to note that these peaks are not affected by tissue properties, but instead the frequencies are determined by characteristics of oscillating air within an open-closed ended tube.[32] Based on these assumptions the resonance frequencies of the respiratory tract were predicted theoretically with good accuracy.[34] It was also shown that the frequency locations of these peaks were shifted to higher frequencies during Heliox(80-20%) breathing and the amplitude of the sound decreased. The flow had no effect on spectral peak locations.[47] This reinforces the idea of a turbulent origin of breath sounds. The amplitude of the sound was shown to be related to flow rate. However, the size of the glottal aperture seems to be changing when low and high flow rates are compared during voluntary breathing. Moreover, it was shown that there exists a subject related threshold level in the flow rate below which no sound is observed.[34] This is also in agreement with sound generation by turbulent flow.[37]

The highest observed tracheal sound frequency increased hand in hand with sound amplitude. It has been widely presented that flow could be calculated from tracheal breath sounds after calibration measurements with a pneumotachography. Some of the presented results were really promising, but more work is definitely needed.[60] From the point of view of sleep polygraphy, flow determination is important especially if the flow curve and the difference between oral and nasal breathing could be detected. An uncalibrated flow signal is used presently and the shape of the flow is more important than the actual spirometrical values. It is commonly accepted that the thermistor which is presently used for oral flow detection is not accurate and easy to use. Noninvasive flow detection based on tracheal sounds would be a sophisticated method in sleep studies. If flow from nasal cannula is measured at the same time as tracheal sounds, oral flow might be possible to determine from tracheal sounds.

Adventitious tracheal sounds are widely researched in sleep sciences, but in most cases the studies are focused on snoring sound analysis. Another alternative recording site might be the nasal and oral cavities, but more studying is definitely needed. When studying the aeroacoustic theory, it shows that if any obstruction is created in the upper airways change in the tracheal breath sounds should be produced. The theory predicts that when the speed of flow increases in the obstruction the amplitude and the frequency of the sound increases as well. [37] Some studies show that the power of higher frequencies increased in relation to lower frequencies when obstructions were created in the upper airways. However, a study of Tenhunen [29] showed that the amplitude and the highest observed frequency increased, but the

relationship of the high and low frequency components remained the same (i.e. there was no change in the shape of the spectrum). However, long epochs were used without discrimination between the inspirations, the expirations and the epochs without sound production and this might have affected the results. The aeroacoustic theory assumes that the walls of the tube are rigid. This is not the case with the real respiratory tract in which some vibrations, such as flutter and snoring, might exist.[65] It might be possible that these vibrations mask the change in the aeroacoustically produced sound or perhaps even prevent the creation of these high frequencies.Therefore more studying of tracheal sounds with high fidelity instruments is needed for clarification.

6.3   Measurements with HeLSA system

Quality measurements of the HeLSA-system were made. It was discovered that the measured tracheal sounds were distorted by harmonics generated by a microphone-amplifier combination and that the frequency response of the sound card was unflat when a 11kHz sampling frequency was used. The calibration method for a microphone 'quick check' was studied. The method can be used in situations such as suspecting that a microphone is not measuring correctly. Distortion artefacts were also found in the expiratory sounds when a 11kHz sampling frequency was used. Two spectral peaks at around 3.7kHz were observed. With a 44kHz sampling frequency artefacts were not observed. It was noticed that the highest observed frequency increased when the sampling frequency was increased to 44kHz.

Similar effects in tracheal breath sounds were observed as have been previously presented in literature.[34] An analysis of the inspirations of the author of this thesis is not presented in this thesis. Expirations were chosen, because it was observed that expirations were louder than inspirations. Therefore, it was assumed that they reveal more when studying the highest observed frequencies. Inspirations are as important as expirations, because usually breathing disorders are related to inspirations. Interestingly, it was noticed that in the polysomnography of the author of this thesis, the tracheal sounds were louder during inspirations than during expirations. The reason is not understood. Maybe the effects of glottal size alterations could explain this. It was also noticed that there exist epochs when the expiration sounds became louder than the inspirations during the author's sleep (see. 5.31 and 5.32). The measured expiration sounds of the author show that similar characteristics of the tracheal sound spectrum can be obtained with the HeLSA as the other research groups have obtained with targeted breaths. These kinds of results have not been observed previously with the HeLSA-microphone at the PSHP Sleep Laboratory. In previous studies the linear amplitude scale has been used in spectrum plots. Therefore, it has been thought that spectral components above 1kHz are meaning-

less because their power level is very small. Moreover, the background level has not been determined from the breath hold events or from the late expirations. Therefore there has been no chance of accurately evaluating the range of frequencies where the breath sound signal is above the background noise level. Moreover, the breath sound spectra have been calculated from long epochs of breathing sounds. These epochs have contained inspirations, expirations and silent periods (background) and they have all been mixed together when the spectra have been calculated. According to the articles presented in this thesis the inspiration and expiration sounds have different characteristics and therefore they should be treated separately. Instead of long epochs the statistics should be improved by making an average of a larger amount of individual inspiration or expiration spectra. However, care should be taken when presenting spectra in a logarithmic scale: the calculations should be made in a linear scale. Otherwise misconceptions may occur.[89]

6.4   Propositions for future studies

A new approach to tracheal sound analysis could be done at the Sleep Laboratory of the PSHP. Instead of using long epochs of tracheal sounds, the sound signal could be chopped into segments. Inspirations and expirations should be treated separately and respective power spectra should be calculated. The expiration spectra should be averaged as well as the inspiration spectra. The background signal should be collected from late expirations and similar averaging should be used. First this analysis should be done for a healthy person. Epochs where inspirations are louder than expirations and vice versa should be analyzed separately. The effects of nasal and oral breathing as well as the body position could be included in the analysis. The second case would be the subject having flow limitation periods in sleep. As previously, the inspirations and expirations should be treated separately and the background should be calculated. In addition to the cases of normal breathing, breaths before, during, and after the flow limitation periods should be analyzed separately. The signal could be improved by replacing the HeLSA with a PPG-accelerometer especially if it seems that high frequencies contain important information. The spectra of the breathing sound should be compared with the background spectrum and the intersection of sound and the background spectrum should be determined.

6.5   Future issues of tracheal sound research

Further sensor development is needed to be sure that all the sounds that are above the background threshold level can be observed. The power of these frequency components may be small, but it does not mean that they could not carry some information. If a general upper frequency limit for tracheal sounds is determined

by a scientific community, it should be done based on analytical measurements with devices which have so high a quality, that the sensor design does not limit the sound observed. In addition to high quality standards, tracheal and breath sounds should be detected with several sensor models based on different technologies. If scientific reliability is sought, the measurements should be done by independent research groups in order to define the high frequency limit of tracheal sounds. The state of tracheal sound research is not of such good quality that these kinds of decisions could be made at the moment. Therefore, more studying is needed and more information should be shared among the researchers on how to design high quality breath sound sensors and how to estimate the quality of the sensors with a quality measurement. There is urgent need for quality measurement setups for both types of contact sensors, the accelerometers and the air coupled microphones. Perhaps some kind of water filled phantom with a semirigid surface could be designed. Water has acoustic characteristics quite similar to soft tissue. A system of this kind was presented by Pasterkamp et al. [2] However, water is not used in this acoustic testing system to mimic tissue properties. Instead, special materials that may not be generally available, have been used.

On the other hand, if the quality standards for tracheal sound microphones are developed, there is no need to determine the artificial limit of the upper frequency of tracheal sounds. The measured spectrum can be reflected to the quality measurements which are done for that specific measurement instrument. This is the principle used in the field of basic research and it could be adopted also to these applied measurements. One consequence would be that instead of determining the upper limit of breath sound frequencies, the scientific society should standardize how the upper limit of a sound spectrum is determined. If the signal during breath hold is defined to be the background, what should be the intensity of the sound signal in the amplitude or power spectrum so that the components from breathing sounds can be assumed to exist? In analytical chemistry the IUPAC has defined the detection limit of the measured signal as $x_L = \bar{x}_{bg} + k s_{bg}$, where $\bar{x}_{bg}$ is the average of the background measurements, s_{bg} is the standard deviation of the background and k is a numerical factor chosen according to the confidence level. [90] An exact expression for the detection limit can be achieved when the confidence level is fixed to a certain value. This means that calculated datapoints of the spectrum should be analyzed statistically when the highest observed sound frequency is determined from the spectrum. The SNR is closely related to the detection level[91]. The role of the detection level can be demonstrated with a practical example. Two people are talking to each other near a street full of traffic noise. When the other person speaks with a soft voice, the words are not heard at all and the traffic noise dominates. When the speaker increases the level of his voice to a moderate level the

listener can hear the sound of the other one speaking above the traffic noise, but he cannot resolve the spoken words. Finally, when the speaker applies a loud voice the listener hears and understands the individual words. These examples present cases in which the confidence level of the signal varied from 0 to 1. Other signal processing parameters, like the quality of the spectrum estimate, should be evaluated as well. What could be done immediately in future breath sound work is presenting the background level during breath holds or during late expirations.

Why are these high precision measurements needed? When new methods are developed for detecting respiratory tract dysfunctions, researchers should have measuring instruments that do not distort the measured signal uncontrollably and which have sensitivity that is high enough. The research of normal tracheal breath sounds during sleep are also very important. Without such measurements it is very difficult to determine which sound is normal and which is related to a pathological status. It is obvious that if the measuring instrument is of low quality the separation of normal and adventitious sounds is difficult.

One of the goals of tracheal sound research in the field of sleep research would be the development of a portable instrument that could be given for a patient to be used at home, so that the data recorded over night could be analyzed later in the laboratory. For such a device, perhaps lower quality sound recording instruments could be used. However, without high quality measurements and without a proper understanding of the mechanisms producing tracheal sounds, a reasonable design of a lower quality instrument is difficult. This is because one of the key questions is how much the overall quality and spectral bandwidth of the recording system can be reduced while still finding the signal which indicates the pathological state. Unfortunately, it seems that the evolution of tracheal sound research has proceeded in the reversed order. First, something abnormal has been observed and studied from tracheal sounds. Later normal tracheal sounds were studied by a few research groups, but the amount of published results has been small. Since the 1990's, some studies concerning quality and sensitivity issues have been made and the standardization of methods has been proposed by several authors. This means that there is a large field of tracheal sound studies where the detailed description of the instruments used is missing and the authors have used instruments which have not been properly designed for tracheal sound recordings.

Despite the lack of quality issues and standardization, the tracheal area as a source of breathing sounds seems to be interesting even nowadays. The Masimo company has released a new FDA approved respiration rate (RR) sensor. This new respiration rate sensor is a contact sensor which is attached to the side of the neck and the respiration phase is detected from the vibrations of the skin.[92] It can be used for monitoring the RR's of post-operative patients or patients under opioid

medication. There is still a long way from RR's to the spectral analysis of sound, but perhaps one day this will also be possible when high fidelity tracheal sound measurements are standardized enough.
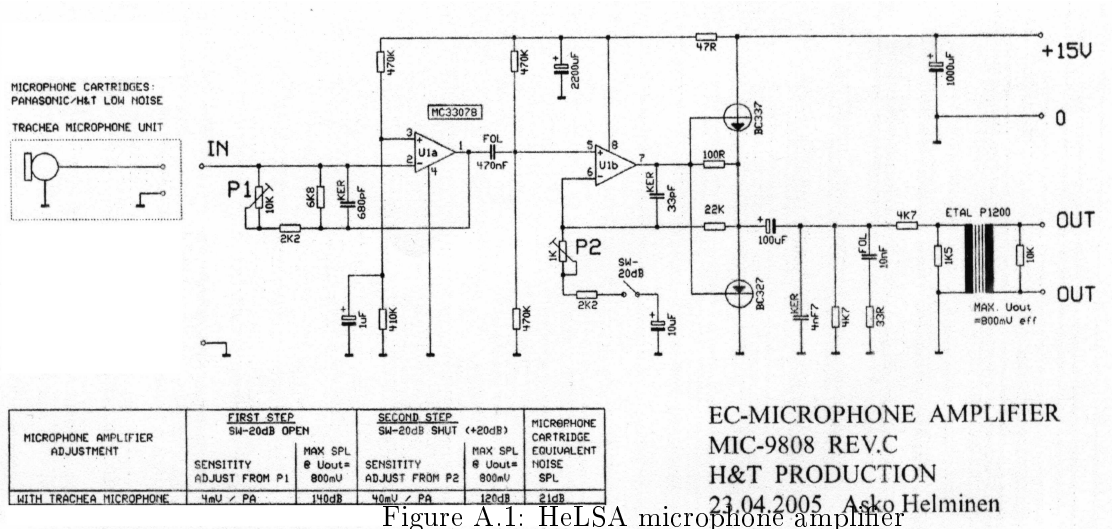
A.   APPENDIX: HELSA MICROPHONE AMPLIFIER

Figure A.1: HeLSA microphone amplifier

BIBLIOGRAPHY

[1] F. Dalmay, M.T. Antonini, P. Marquet, and R. Menier. Acoustic properties of the normal chest review. *European Respiratory Journal*, 8:1761–1769, 1995.

[2] S.S. Kraman, G.A. Pressler, H. Pasterkamp, and G.R. Wodicka. Design, construction and evaluation of a bioacoustic transducer testing '(batt)' system for respiratory sounds. *IEEE Transactions on Biomedical Engineering*, 53:1711–1715, 2006.

[3] S.S. Kraman, G.R. Wodicka, G.A. Pressler, and H. Pasterkamp. Comparison of lung sound transducers using a bioacoustic transducer testing system. *Journal of Biomechanical Engineering*, 101:469–476, 2006.

[4] H. Benson. *University Physics*, pages 325–364. John Wiley & Sons, Inc., 1995.

[5] P.M. Morse and U.K. Ingard. *Theoretical acoustics*, chapter 9,11. McGraw-Hill Book Company, 1968.

[6] R.S. Schwartz, J.T Reeves, I.E. Sodal, and F.S. Barnes. Improved phonocardiogram system based on acoustic impedance matching. *Am J Physiol Heart Circ Physiol*, 238:H604–H609, 1980.

[7] C.T. Leondes, editor. *Biomechanical Systems Technology*, chapter 1. Acoustical signals of biomechanical systems. World Scientific Publishing Co, 2007. Author E. Kaniusas.

[8] N. Suzumura and K. Ikegaya. Characteristics of the air cavities of phonocardiographic microphones and the effects of vibration and room noise. *Acoust. Sci. & Tech.*, 22:5–11, 2001.

[9] E. Van Vollenhoven. Calibration of contact microphones applied to the human chest wall. *Med. & biol. Engng.*, 9:365–373, 1971.

[10] E. Van Vollenhoven and J. Wallenburg. Calibration of air microphones for phonocardiography. *Med. & biol. Engng.*, 8:309–313, 1970.

[11] J. Hasegawa and K. Kobayashi. Blood flow noise transducer for detecting intracranial vascular deformations. *Acoust.Sci.&Tech.*, 22:5–11, 2001.

[12] Astra-Med,Inc. Grass product group. *Live demonstration: Acoustic Myography (AMG)*, 2004.

[13] K. Akataki M. Watakabe, K. Mita and Y. Itoh. Mechanical behaviour of condenser microphone in mechanomyography. *Med.Biol.Eng.Comput.*, 39:195–201, 2001.

[14] P. Heracleous, Y. Nakajima, H. Saruwatari, and K. Shikano. Audible (normal) speech and inaudible murmur recognition using nam microphone. In *Proceedings of XII European Signal Processing Conference (EUSIPCO 2004)*, volume 1, pages 329–332, 2004.

[15] P. Heracleous, Y. Nakajima, H. Saruwatari, and K. Shikano. A tissue-conductive acoustic sensor applied in speech recognition for privacy. In *ACM International Conference Proceedings Series, Proceedings of the 2005 joint conference on Smart objects and ambient intelligence:innovative context-aware services: usages and technologies*, volume 121, pages 93–97, 2005. Joint sOc-EUSAI conference.

[16] P. Heracleous, T. Kaino, H. Saruwatari, and K. Shikano. Unvoiced speech recognition using tissue-conductive acoustic sensor. *EURASIP Journal on Advances in Signal Processing*, pages 1–11, 2007.

[17] G.R. Wodicka, S.S. Kraman, G.M. Zenk, and H. Pasterkamp. Measurement of respiratory acoustic signals. effect of microphone air cavity depth. *Chest*, 106:1140–1144, 1994.

[18] T. Hirahara, S. Shimizu, and M. Otani. Acoustic characteristics of non-audible murmur. In *Japan-China Joint Conference on Acoustics, JCA 2007*, 2007.

[19] E.K. Franke. Mechanical impedance of the surface of the human body. *J Appl Physiol*, 3:582–590, 1951.

[20] S. Takagi. Trial and success in the technical realization of requirements for phonocardiographic microphones. *Med. Electron. Biol. Engng.*, 2:123–134, 1964.

[21] B. F. G. Katz. Acoustical absorption measurement of human hair and skin within the audible frequency range. *J. Acoust. Soc. Am.*, 108:2238–2242, 2000.

[22] H. Pasterkamp, S.S. Kraman, P.D. DeFrain, and G.R. Wodicka. Measurement of respiratory acoustic signals. comparison of sensors. *Chest*, 104:1518–1525, 1993.

[23] C.K. Druzgalski, R.L. Donnerberg, and R.M. Cambell. Techniques of recording respiratory sounds. *Journal of clinical engineering*, 5:321–333, 1980.

[24] S.S. Kraman, G.R. Wodicka, Y. Oh, and H. Pasterkamp. Measurement of respiratory acoustic signals. effect of microphone air cavity width, shape and venting. *Chest*, 108:1004–1008, 1995.

[25] M. Munakata, H. Ukita, I. Doi, and et al. Spectral and waveform characteristics of fine and coarse crackles. *Thorax*, 46:651–657, 1991.

[26] H. Pasterkamp, G.R. Wodicka, and S.S. Kraman. Effect of ambient respiratory noise on the measurement of lung sounds. *Medical and Biological Engineering and Computing*, 37:461–465, 1999.

[27] M. Zanartu, J.C. Ho, S.S. Kraman, H. Pasterkamp, J.E. Huber, and G.R. Wodicka. Air-borne and tissue-borne sensitivities of bioacoustic sensors used on the skin surface. *IEEE Transactions on Biomedical Engineering*, 56:443–451, 2009.

[28] M. J. Mussell, Y.Nakazono, and Y. Miyamoto. Effect of air flow and flow transducer on tracheal breath sounds. *Med. & Biol. Eng. & Comput.*, 28:550–554, 1990.

[29] M. Tenhunen, E. Rauhala, E. Huupponen, A. Saastamoinen, A. Kulkas, and S-L. Himanen. High frequency components of tracheal sound are emphasized during prolonged flow limitation. *Physiol. Meas.*, 30:467–478, 2009.

[30] R. Beck, M. Odeh, A. Oliven, and N. Gavriely. The acoustic properties of snores. *European Respiratory Journal*, 8:2120–2128, 1995.

[31] H. Pasterkamp, J. Schäfer, and G.R. Wodicka. Posture-dependent change of tracheal sounds at standardized flows in patients with obstructive sleep apnea. *Chest*, 110:1493–1498, 1996.

[32] V.P. Harper, S.S. Kraman, H. Pasterkamp, and G.R. Wodicka. An acoustic model of respiratory tract. *IEEE Transactions on Biomedical Engineering*, 48:543–550, 2001.

[33] M. Yonemaru, K. Kikuchi, M. Mori, A. Kawai, T. Abe, T. Kawashiro, T. Ishihara, and T. Yokoyama. Detection of tracheal stenosis by frequency analysis of tracheal sounds. *J. Appl. Physiol.*, 75:605–612, 1993.

[34] V.P. Harper, H. Pasterkamp, H. Kiyokawa, and G.R. Wodicka. Modelling and measurements of flow effects on tracheal sounds. *IEEE Transactions on Biomedical Engineering*, 50:1–10, 2003.

[35] I. Sanchez and H. Pasterkamp. Tracheal sound spectra depend on body height. *The American review of respiratory disease*, 148:1083–1087, 1993.

[36] R. Beck, G. Rosenhouse, M. Mahagnah, R.M. Chow, D.W. Cugell, and N. Gavriely. Measurements and theory of normal tracheal breath sounds. *Annals of Biomedical Engineering*, 33:1344–1351, 2005.

[37] M.H. Krane. Aeroacoustic production of low-frequency unvoiced speech sounds. *Journal of the Acoustical Society of America*, 118:410–427, 2005.

[38] J.L. Flanagan and L. Cherry. Excitation of vocal-tract synthesizers. *The journal of the Acoustical Society of America*, 45:764–769, 1969.

[39] M. Yonemaru, K. Kikuchi, M. Mori, A. Kawai, T. Abe, T. Kawashiro, T. Ishihara, and T. Yokoyama. Detection of tracheal stenosis by frequency analysis of tracheal sounds. *J. Appl. Physiol.*, 80:5–13, 1996.

[40] A. Yadollahi and Z.M.K Moussavi. Acoustical respiratory flow. *IEEE in Medicine and Biology Magazine*, 26:56–61, 2007.

[41] H. Pasterkamp, S.S. Kraman, and G.R. Wodicka. Respiratory sounds. advances beyond the stethoscope. *American Journal of Respiratory and Critical Care Medicine*, 156:974–987, 1995.

[42] K. Ishizaka, M. Matsudaira, and T. Kaneko. Input acoustic-impedance measurement of the subglottal system. *The journal of the Acoustical Society of America*, 60:190–197, 1976.

[43] H. Pasterkamp, S. Patel, and G.R. Wodicka. Asymmetry of respiratory sounds and thoracic transmission. *Medical and Biological Engineering and Computing*, 35:103–106, 1997.

[44] T. Brancatisano, P.W. Collet, and L.A. Engel. Respiratory movements of the vocal cords. *Journal of Applied Physiology*, 54:1269–1276, 1983.

[45] S.S. Kraman, H. Pasterkamp, M. Kompis, M. Takase, and G.R. Wodicka. Effects of breathing pathways on tracheal sound spectral features. *Respiration Physiology*, 111:295–300, 1998.

[46] J.L. Flanagan. *Speech Analysis Synthesis and Perception*, chapter 2-3. Springer-Verlag, 1972.

[47] H. Pasterkamp and I. Sanchez. Effect of gas density on respiratory sounds. *Am J Respir Crit Care Med*, 153:1087–1092, 1996.

[48] E. Seren. Effect of nasal valve area on inspirator nasal sound spectra. *Otolaryngology-Head and Neck Surgery*, 134:506–509, 2006.

[49] E. Seren. Frequency spectra of normal expiratory nasal sounds. *American Journal of Rhinology*, 19:257–261, 2005.

[50] R. Tahamiller, D. Edizer, and S. Canakcioglu. Nasal expiratory sound analysis in healthy people. *Otolaryngology-Head and Neck Surgery*, 134:605–608, 2006.

[51] R. Tahamiller, D. Edizer, S. Canakcioglu, M. Guvench, E. Inci, and A. Dirican. Nasal sound analysis: A new method for evaluating nasal obstruction in allergic rhinitis. *Laryngoscope*, 116:2050–2054, 2006.

[52] R. Tahamiller, D. Edizer, S. Canakcioglu, and A. Dirican. Odiosoft-rhino versus rhinomanometry in healthy subjects. *Acta Oto-Laryngologica*, 128:181–185, 2008.

[53] J.A. Fiz, J. Gnitecki, S.S. Kraman, G.R. Wodicka, and H. Pasterkamp. Effect of body position on lung sounds in healthy young men. *Chest*, 133:729–736, 2008.

[54] H. Pasterkamp, J. Schafer, and G.R. Wodicka. Posture-dependent change of tracheal sounds at standardized flows in patients with obstructive sleep apnea. *Chest*, 110:1493–1498, 1996.

[55] C-L. Que, C. Kolmaga, L-G. Durand, S. Kelly, and P. Macklem. Phonospirometry for noninvasive measurement of ventilation: methodology and preliminary results. *J. Appl. Physiol.*, 93:1515–1526, 2002.

[56] M. Golabbakhsh and Moussavi Z. Relationship between airflow and frequency-based features of tracheal respiratory sound. In *IEEE CCECE*, pages 751–754, 2004.

[57] Z.K. Moussavi, M.T. Leonpando, H. Pasterkamp, and G.Rempel. Computerised acoustical respiratory phase detection without airflow measurement. *Med.Biol.Eng.Comput.*, 38:198–203, 2000.

[58] G. Soufflet, G. Charbonneau, M. Polit, P. Attal, A. Denjean, P. Escourrou, and C. Gaultier. Interaction between tracheal sound and flow rate: A comparison of some different flow evaluations from lung sounds. *IEEE Trans.Biomed. Eng.*, 37:384–391, 1990.

[59] A. Yadollahi and Z. Moussavi. Comparison of flow-sound relationship for different features of tracheal sound. In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 805–808. IEEE, 2008.

[60] A. Yadollahi and Z.K. Moussavi. A robust method for estimating respiratory flow using tracheal sounds entropy. *IEEE Trans. Biomed. Engn.*, 54:662–668, 2006.

[61] Y.L. Yap and Z.K. Moussavi. Acoustic airflow estimation from tracheal sound power. In *IEEE CCECE*, 2002.

[62] A. Gold and A. Schwartz. The pharyngeal critical pressure. the whys and hows of using nasal continous positive airway pressure diagnostically. *Chest*, 110:1077–1088, 1996.

[63] J. P. Kirkness, V. Krishnan, S. P. Patil, and H. Schneider. *Sleep Apnea. Upper Airway Obstruction in Snoring and Upper Airway Resistance Syndrome*, volume 35 of *Prog Respir Res.*, pages 79–89. Karger, 1995.

[64] A.R. Schwartz, P.L. Smith, R.A Wise, A.R. Gold, and S. Permutt. Induction of upper airway occlusion in sleeping individuals with subatmoshperic nasal pressure. *J Appl Physiol*, 64:535–542, 1988.

[65] N. Gavriely and O. Jensen. Theory and measurements of snores. *Journal of Applied Physiology*, 74:2828–2837, 1993.

[66] N. Gavriely, Y. Palti, G. Alroy, and J.B. Grotberg. Measurement and theory of wheezing breath sounds. *Journal of Applied Physiology*, 57:481–492, 1984.

[67] N. Gavriely, T.R. Shee, D.W. Cugell, and J.B. Grotberg. Flutter in flow limited collapsible tubes: a mechanism for generation of wheezes. *Journal of Applied Physiology*, 66:2251–2261, 1989.

[68] F. Dalmasso and R. Prota. Snoring:analysis, measurement, clinical implications and applications. *European Respiratory Journal*, 9:146–159, 1996.

[69] C.M. Rembold and P.M. Suratt. An upper airway resonator model of high-frequency inspiratory sounds in children with sleep-disordered breathing. *J Appl Physiol*, 98:1855–1861, 2005.

[70] P. Foragacs. Editorial:breath sounds. *Thorax*, 33:681–683, 1978.

[71] J.A. Fiz, J. Gnitecki, S.S. Kraman, G.R. Wodicka, and H. Pasterkamp. Effect of body position on lung sounds in healthy young men. *Sleep Med. Rev.*, 14:131–144, 2010.

[72] G. Liistro, D.C. Stanescu, C. Veriter, D.O. Rodenstein, and G. Aubert-Tulkens. Pattern of snoring in obstructive sleep apnea patients and in heavy snorers. *Sleep*, 14:517–525, 1991.

[73] S. Agrawal adn P. Stone, K. McGuinness, and J. Morris. Sound frequency analysis and the site of snoring in natural and induced sleep. *Clin. Otolaryngol.*, 27:162–166, 2002.

[74] H. Nakano, T. Ikeda, M.Hayashi, E. Ohshima, and A. Onizuka. Effects of body position in apneic and nonapneic snorers. *Sleep*, 2:169–172, 2003.

[75] A. Yadollahi and Z. Moussavi. Formant analysis of breath and snore sounds. In *Proceedings of the 31th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2563–2566, 2009.

[76] A.K. Ng, T.S. Koh, E. Baey, T.H. Lee, U.R. Abeyratne, and K. Puvanendran. Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea? *Sleep Med.*, 9:894–898, 2008.

[77] S.J. Quinn, L. Huang, P.D.M. Ellis, and J.E. Ffowcs Williams. The differentiation of snoring mechanisms using sound analysis. *Clin. Otolaryngol.*, 21:119–123, 1996.

[78] M. Herzog, E. Schieb, T. Bremert, B. Herzog, W. Hosemann, H. Kaftan, and T. K'uhnel. Frequency analysis of snoring sounds during simulated and nocturnal snoring. *Eur Arch. Otorhinolaryngol*, 265:1553–1562, 2008.

[79] C.M. Rembold and P.M. Suratt. Children with obstructive sleep-disordered breathing generate high-frequency inspiratory sounds during sleep. *Sleep*, 27:1154–1161, 2004.

[80] C.M. Rembold and P.M. Suratt. An upper airway resonator model of high-frequency inspiratory sounds in children with sleep-disordered breathing. *J Appl Physiol*, 98:1855–1861, 2005.

[81] A. Yadollahi and Z. Moussavi. Acoustic obstructive sleep apnea detection. In *Proceedings of the 31th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 7110–7113, 2009.

[82] L.P. Malmberg, L. Pesu, and A.R. Sovijärvi. Significant differences in flow standardised breath sound spectra in patients with chronic obstructive pulmonary disease, stable asthma and healthy lungs. *Thorax*, 50:1285–1291, 1995.

[83] S. Harris and C. Sanchez. *Personal Computer Audio Quality Measurements.* Cirrus Logic, Crystal Audio Division, version 1.0 edition, September 1999. MEAS100.

[84] A.J. Aude. *Audio Quality Measurement Primer.* Intersil, October 1998. Application note AN9789.

[85] Anritsu. *Intermodulation Distortion (IMD) Measurements Using the 37300 Series Vector Network Analyzer*. Application note.

[86] European Broadcasting Union (EBU). *EBU Technical Recommendation R68-2000 Alignment level in digital audio production equipment and in digital audio recorders*.

[87] S. Klar and G. Spikofski. *On levelling and loudness problems at television and radio broadcast studios*. Audio Engineering Society (AES), May 2002. Conventio Paper 5538.

[88] Z. Wang adn Y. Xu. Desing and optimization of an ultra-sensitive piezoresistive accelerometer for continuous respiratory sound monitoring. *Sensor Letters*, 5:450–458, 2007.

[89] A. Yadollahi and Z. Moussavi. On arithmetic misconceptions of spectral analysis of biological signals, in particular respiratory sounds. In *Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 388–391. IEEE, 2009.

[90] M. Nic, J. Jirat, and B. Kosata. *Compendium of Chemical Terminology, 2nd ed. (the "Gold Book")*, chapter Orange Book, p. 5. Blackwell Scientific Publications, 1997. XML on-line corrected version: http://goldbook.iupac.org (2006-).

[91] J. Ripp. Analytical detection limit guidance & laboratory guide for determining method detection limits. Technical report, Wisconsin Department of Natural Resources, 1996. Laboratory Certification Program.

[92] Masimo product overview (8.4.2010) http://www.masimo.com/rra/.