Canadian Acoustics / Acoustique Canadienne 26(2) 11-19 (1998) **Research article / Article de recherche**

AIR CONDUCTED AND BODY CONDUCTED SOUND PRODUCED BY OWN VOICE

Mie Ø. Hansen* and Michael R. Stinson

Institute for Microstructural Sciences
National Research Council
Ottawa, Ontario K1A 0R6, Canada

ABSTRACT

When we speak, sound reaches our ears both through the air, from mouth to ear, and through our body, as vibrations. The ratio between the air borne and body conducted sound has been studied in a pilot experiment where the air borne sound was eliminated by isolating the ear with a large attenuation box. The ratio was found to lie between -15 dB to -7 dB, below 1 kHz, comparable with theoretical estimations. This work is part of a broader study of the occlusion effect and the results provide important input data for modelling the sound pressure change between an open and an occluded ear canal.

SOMMAIRE

Quand on parle, le son atteint nos oreilles à la fois à travers l'air, de notre bouche à l'oreille, et à travers notre corps comme vibrations. Le rapport entre le son se propageant dans l'air et le son se propageant dans le corps a été étudié dans un projet-pilote où le son se propageant par l'air fut éliminé en isolant l'oreille avec une grande boîte d'atténuation. Il fut découvert que ce rapport se trouve entre de -15 dB à -7 dB, au-dessous de 1 kHz, comparable aux évaluations théoriques. Ce travail fait partie d'une étude plus large sur l'effet d'occlusion et les résultats fournissent des données d'entrée importantes pour la modelisation du changement de la pression sonore entre un canal auditif ouvert et un qui est bloqué.

*On a work term from the Dept. of Acoustic Technology, Technical University of Denmark

1. INTRODUCTION

People using earplugs are aware that the sound of their own voice is unnaturally loud. In the following text, the term 'occlusion effect' (OE) refers to the difference in sound pressure level between closed and open ear canal while the person is speaking. A tight earmold can give an occlusion effect up to 30 dB. The occlusion effect has typically a maximum at 100-500 Hz and drops to 0 dB at 1-2 kHz, but can vary much between subjects.

The hearing organ detects both environmental sounds and the sounds we produce with our own body. The total sound pressure in the ear canal created by our own voice is thought to be mainly a sum of two components: An air conducted signal propagates from mouth through the air to the ear while a body conducted component propagates as vibrations from the larynx through the neck and skull to the soft cartilage of the ear canal, where it radiates into the canal. At low frequencies, the body conducted sound waves will radiate out of the unoccluded ear so that the open canal sound pressure will be nearly the same as the air borne sound. In the closed ear, however, the body conducted sound will be reflected by the earmold (or earplug) and an enhanced sound pressure from this component will arise in the closed ear canal.

Although the occlusion effect has been measured by various authors, additional knowledge of the air conducted and body conducted components is required to help understand and to model the occlusion effect. The availability of a model to describe the occlusion effect would be a very helpful tool for developing a method to eliminate the occlusion effect. Such a method would benefit hearing aid users, for example. An independent measure that is particularly pertinent is the relative magnitude of the body conducted sound to the air conducted component. This ratio is the object of interest in this paper.

One of the more comprehensive studies on bone conducted sound is that of Schroeter and Poesselt[1] (1986). They modelled the contributions from the middle ear and the ear canal walls to the total sound pressure in the occluded ear canal. However, they considered only the sound produced by a bone vibrator placed on the skull and not by one's own voice. The literature provides little real ear data on body conducted sound produced by one's own voice. The most relevant work is that of Békésy[2] in 1949. Békésy tried to eliminate the air conducted sound by the mean of large tubes placed on the ears. These tubes attenuated the air conducted sound by 30 dB between 150 Hz and 4 kHz, without causing an occlusion effect. The difference in loudness between open ears and covered ears of short words, using the subject's own vocalization, was measured by comparison of the words heard to the loudness of a 1 kHz tone produced by a bone-conductor on the forehead. Békésy found that the body conducted component was 0 - 10 dB lower than the air conducted component, depending on the sound produced.

This experiment needs some comment. First of all, the ratio was only measured indirectly, no real ear pressures being determined. Assuming that the loudness measurements are done using standard methods, tissue and bone act as a lowpass filter and the voice will then sound deeper when only the body conducted sound is heard. The difference in loudness between open and covered ear is therefore not necessarily a measure for the loudness at the same frequencies. Secondly, a reference microphone was not used to assure that the subjects spoke with the same level in the two situations. The present author[3] has performed an experiment showing that people automatically change their speech level when the ear is occluded. When the ear is covered and there is no occlusion effect, it must be anticipated that subjects still raise their own voice by a few dB. This phenomenon means that the ratio is smaller than that measured by Békésy.

It was therefore decided to perform an experiment in order to get an estimate of the ratio between air and body conducted voice sounds, making use of the probe tube technology that has developed since the time of Békésy.

2. SOUND CONDUCTION MECHANISMS

An overview of the mechanisms involved in the generation of the air and body conduction components will be presented here. The relevant anatomy is shown schematically in Fig. 1.

2.1. Speech Production

Speech contains voiced (vowels and voiced consonants) and unvoiced (consonants) sounds. Voiced sounds are primarily created by periodic oscillations of the vocal folds. Unvoiced sounds are aperiodic and generated by oral constriction. Oral constriction is also called occlusion in speech production terminology and should not be confused with the occlusion effect in the ear canal[4].

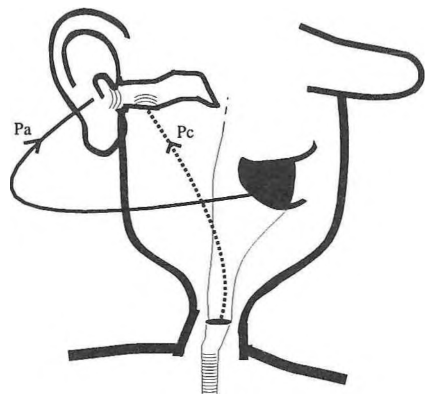


Figure 1. Sketch of a human head showing the air conduction (P_a) and body conduction (P_c) pathways.

The vocal folds (which form the glottis) are set into vibration by air flow from the lungs up through the vocal tract. The opening and closing of the vocal folds acts like a saw tooth generator. The time function of the velocity of volume passing through the glottis (the area between the vocal folds) is a train of approximately triangular pulses. Transformed into the frequency domain, these pulses give a line spectrum that decreases by approximately 12 dB per octave, as seen in Fig. 2. This spectrum is the glottal source spectrum. The fundamental frequency of a voice is determined by the strength and the length of the vocal folds. On average, the fundamental frequency is 125 Hz for men and 200 Hz for women.

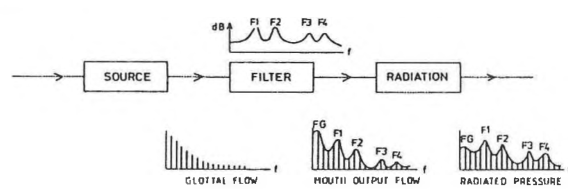


Figure 2. Vowel shaping.

The sound produced in the glottis is modified by the vocal tract and by the oral and nasal cavities and depends on the geometry of these structures. The shape of the oral cavity is controlled by the position of the tongue. Sound is radiated from the mouth, being controlled by lip rounding and opening of the mouth. There is also some radiation of sound from the nose. The air conduction pathway is completed by propagation of sound from mouth (and nose) around the head to the ear; this propagation is controlled by diffraction. A simple model of vowel shaping is illustrated in Fig. 2. The first line in the final spectrum represents the fundamental frequency, the next lines are the harmonics and the peaks of the envelope are the formants of the vowel. The formants are resonances in the acoustic system consisting of the vocal tract and the shaped oral cavities. Changing the shape of the oral cavity with the tongue, jaw and lip positioning leads to different vowel sounds. The three most distinct vowels with most extreme tongue positions are the cardinal vowels /i/, /a/ and /u/.

Continuous speech can be measured as a long-term average spectrum, over 60 seconds, for example. The radiated spectrum is language independent[5].

2.2. Ear Canal Anatomy and Acoustics

The average ear canal is about 23.5 mm long and 7 mm in diameter[6]. The outer part of the ear canal is soft tissue and the inner part is surrounded by the temporal bone. The soft part of the ear canal is surrounded by connecting tissue of the cartilage type which contains elastic fibers and makes it very flexible. The cartilage is covered by a 0.5-1.0 mm thick skin layer that holds hair cells. In the bony part of the ear canal the skin layer is thinner, approximately 0.2 mm and this skin is continuous with the skin covering the eardrum.

Seen from above, the ear canal is S-shaped. The ear canal terminates the eardrum, forming an angle of approximately 55° with the inferior wall. Oliveira's[7] scannings show that the first bend is located in the cartilaginous part while the second bend is surrounded partly by cartilage and partly by bone, implying that the soft tissue reaches 1/3-1/2 into the ear canal. A standard hearing aid earmold ends between the first and second bend. On average it reaches 1/3 into the ear canal. The ear canal shape and dimensions do differ significantly from person to person.

Despite these complexities, some simple acoustical models can be used. If the largest dimension of the ear canal is much less than a quarter of the wavelength, then the ear canal can be treated as a simple cavity. This simplification is valid for frequencies below about 1 kHz. At higher frequencies, up to 6 kHz or so, the ear canal can be approximated as a uniform cylindrical tube[8,9]; the influence of higher modes is negligible[10] at these frequencies. At even higher frequencies, the individual ear canal geometry has an effect. The surfaces of the ear canal can be treated as rigid regarding air borne sound transmission because the dilatation impedance in bone and cartilage is much larger than air[10]. The eardrum and middle ear impedance can be calculated using a lumped-element model such as that of Zwislocki[11].

2.3. Hypotheses of Body Conduction

The body conduction pathway extends between the larynx and the ear canal. The vocal folds are located within the larynx, in the front part of the neck, just behind the thyroid cartilage. Attached to this cartilage is the thyroid muscle, running up to the hyoid bone, and muscles from the hyoid bone are directly connected to the bone around the external ear, the temporal bone. It is anticipated that the sound in the vocal tract can couple mechanically to some or all of these structures.

Vibrations from the speech organ and the mouth cavity are transmitted, not only to the outer ear canal, but to the middle ear and probably to the inner ear as well[18]. All three pathways will contribute to the perception of the body conducted sound. In this paper, though, we are only concerned with the sound pressure level in the ear canal for which the dominant source is believed to be the vibrations of the ear canal walls. We neglect, therefore, the middle ear and inner ear pathways in the modelling.

The basic requirement for producing a sound in the ear canal is that a volume change takes place, for example a (tiny) deformation of the ear canal wall. The source of body conducted sound can be the vibrations of the vocal cords and the subsequent air pressure changes in the oral cavity, exciting the skull. Békésy[2] demonstrated that one's own voice sets the skull into vertical vibration. Another way to create body conducted sound is to clench the teeth and waves will be transmitted via the bone. The result is that the cartilagineous part of the ear canal vibrates whether it has been excited by skull vibrations or directly via the tissue from the vocal cords to the ear canal.

The cartilagineous wall in the canal not only vibrates during speech but deforms when the jaw moves. Lowering the jaw increases the cross-section of the soft ear canal[7], mostly because the ventral wall is moved outwards. Hence, jaw movements cause a volume change in the ear canal, but do not create excess sound in the speech frequencies because the jaw movements are slow: It takes about 0.4 seconds (2.5 Hz) to lower the jaw during speech[12].

However, the jaw does influence the vibration of the cartilage through its inertia. Franke et al.[13] (1952) measured the ear canal sound pressure when the lower jaw vibrates relative to the skull. They placed a bone conductor on top of the skull, producing a vertical oscillation of the skull. The ear canal was closed with a rubber ear plug and the sound pressure in the ear canal was measured with open and closed mouth positions. The sound pressure in the occluded ear canal increased by 8 dB at 200 Hz. When the ear was open the increase was only 2-2.5 dB at 200 Hz. Howell et al.[14] (1988) also measured the occluded ear canal pressure on 4 subjects with the jaw closed and open. They found that the phase did not change significantly. These measurements indicate that the sound pressure level created in the ear canal by a vibrating jaw is higher when the jaw is lowered. The upper part of the jaw (the condyle) is located closer to the ear canal when the mouth is closed than when it is open. This observation supports the hypothesis that the inertia of the jaw contributes to the damping of vibrations of the cartilage.

3. ESTIMATING P_C / P_A

During speech with the ear canal open, the sound pressure in the canal is a sum of an air borne component P_A and a component P_C passed through the body (see Fig. 1). Assuming that these components are uncorrelated, the total root-mean-square pressure is $(P_A^2 + P_C^2)^{1/2}$. In the following, we consider the effects of blocking the air borne path, both in producing the occlusion effect and as a means for establishing the relative magnitude of the two components.

3.1. The Occlusion Effect

The occlusion effect OE is calculated as the difference in sound pressure level between the closed ear and open ear conditions, when the canal is blocked with an earplug or hearing aid mold. With the ear closed, the air conducted component is reduced by the attenuation of the mold giving a sound pressure αP_A, where α is the earmold attenuation factor. On the other hand, closing the ear canal tends to increase the sound pressure due to the body conducted component, at least at lower frequencies. The acoustic load presented by the closed canal is higher than the load for the open canal and, with an effective volume velocity source, higher sound levels are anticipated. With the canal blocked, the sound pressure due to body conducted sound is βP_C, where β is an occlusion gain factor. The occlusion effect can then be calculated as

$$OE^2 = \frac{P_{\text{closed}}^2}{P_{\text{open}}^2} = \frac{\alpha^2 P_A^2 + \beta^2 P_C^2}{P_A^2 + P_C^2} \qquad (1)$$

The values of α and β depend on the depth of the earmold. Consequently, the OE also depends on the insertion depth, in agreement with clinical experiences. In the present study, only standard length earmolds were used, i.e., insertion depths about 1/3 of the total ear canal length.

Some measurements of the occlusion effect are shown in Fig. 3. For these, Wimmer[15] and Thorup[16] both inserted the probe tube through a hole in the earmold, avoiding leakage. May and Dillon[17] placed the probe tube between the ear canal wall and earmold.

The β factor has not been measured previously. As seen from Eq. (1), it is a quantity that can be derived once measurements of both OE and P_C / P_A are available. It is possible, though, to estimate β theoretically if assumptions are made about how the body conduction is taking place. A model, currently under development[18], to predict the occlusion gain factor will be used in this paper to provide rough estimates for β. In this model, an effective volume velocity source in the ear canal wall represents the radiation of the body conducted sound into the canal. The ear canal is terminated using a middle ear model (modified Zwislocki

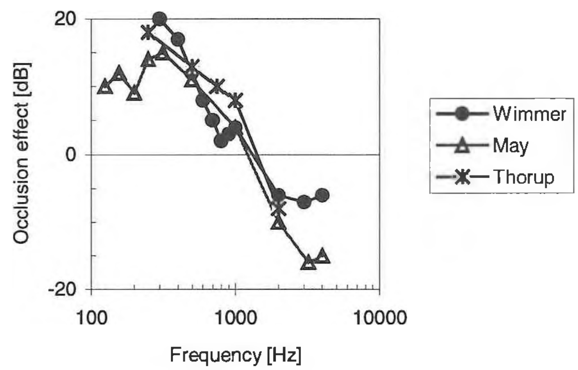


Figure 3. Occlusion effects, own speech (References 15, 16 and 17).

model). Initial calculations indicate that the value of β is much greater than unity. The sound pressure P_C is typically of the same order of magnitude[2] as P_A (though generally smaller) and the factor α is less than unity, so that $(\beta P_C)^2 \gg (\alpha P_A)^2$ should be a good approximation. With this assumption,

$$OE^2 \cong \frac{\beta^2 (P_C / P_A)^2}{1 + (P_C / P_A)^2} \Leftrightarrow \qquad (2)$$

$$(P_C / P_A)^2 \cong \frac{OE^2}{\beta^2 - OE^2}$$

This result allows us to obtain estimates of the ratio P_C / P_A from published values of OE.

The air conducted and body conducted sound have been assumed to be uncorrelated. However, P_C and P_A are likely highly correlated. It has not been possible, though, to find any reports on measurement of the phase created by vocalization, although some data on phase differences of signals traveling from one ear to the other and sound velocity in the skull has been reported. The phase difference could not be measured in this pilot study because the required two probe measurements were not done simultaneously.

3.2. Procedure to Measure P_C / P_A

Our method of determining the ratio P_C / P_A is similar to that used by Békésy[2] except for two main differences. First, the ear canal sound pressures are measured directly, using probe microphones. And second, an attenuation box, rather than large tubes, is used to reduce the air conduction component. With this technique, described later, the air borne component can be reduced by a substantial factor γ, without causing an occlusion effect. Then, the pressure with the attenuation box in place P_{boxed} is related to the pressure without the box P_{open} through

$$\frac{P_{\text{boxed}}^2}{P_{\text{open}}^2} = \frac{\gamma^2 P_A^2 + P_C^2}{P_A^2 + P_C^2} \qquad (3)$$

Now, if γ can be made sufficiently small that $(\gamma P_A)^2 \ll P_C^2$, then we have

$$(P_C / P_A)^2 \cong \frac{P_{\text{boxed}}^2}{P_{\text{open}}^2 - P_{\text{boxed}}^2} \qquad (4)$$

so a measurement of P_{boxed} and P_{open} yields the desired ratio P_C / P_A.

4. MEASUREMENT METHOD

This section describes the set up and equipment used in the pilot study. The three main issues are the design of the box, repeatability of the reference microphone and cushion pressure.

4.1. Set Up and Calibration

The test set up is sketched in Fig. 4. The subjects were seated in a chair in an anechoic chamber. The ear canal pressure in the test ear was measured with a probe microphone. At the same time, a reference signal was picked up near the opposite ear with a ½" Brüel & Kjaer microphone. The probe microphone signal was amplified with a Stanford Research System SR640 amplifier. The reference signal was amplified using a Brüel & Kjaer 2610 measuring amplifier. Both signals were lowpass-filtered with a Stanford Research System SR640 amplifier at 9 kHz in order to avoid aliasing. The signals were recorded onto DAT-tape with a sampling frequency of 44.1 kHz. In the data conversion procedure, the signals were down sampled to 22,050 Hz. The supervisor could check the signals via an external loudspeaker and a scope (not shown on the figure).

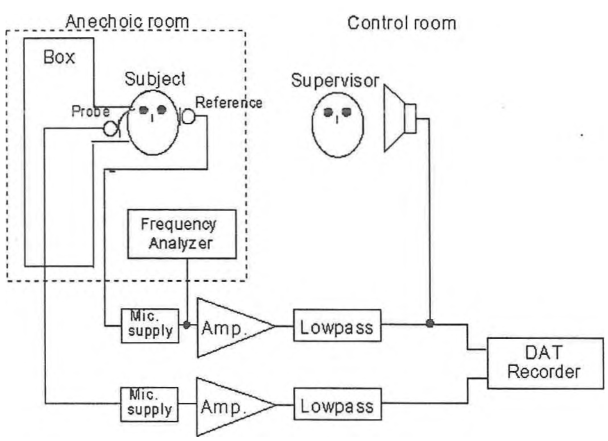


Figure 4. Test set up

Figure 4 shows the subject attached to the attenuation box. In the free field situation, the box was taken out of the anechoic room. As only one probe microphone was available it was necessary to have the subject repeat the same sound twice. Continuous speech (10 Harvard sentences) and the cardinal vowels /i/, /a/ and /u/ were used. The challenge here was to maintain identical conditions (formant frequencies and levels) for the two cases, with and without the attenuation box in place. The subject could look at the monitor of a simultaneous frequency analyzer, where the fundamental frequency of the first pronounced vowel was stored and the subject then had to hit the same fundamental frequency again. A successful reproduction required a few tries every time.

The system was calibrated using a Brüel & Kjaer calibrator Type 4231. The calibrator tone (94.0 dB ± 0.2 dB at 1 kHz) was fed into a 1/2" microphone, followed by a 70 Hz highpass filter and a 9.0 kHz lowpass filter and into measuring amplifier, with an amplification of 30 dB. Digital signals were then calibrated by calculating the rms-value of the calibration signal in the time domain and setting this value equal to 94.0 dB.

4.2. Attenuation Box Design

A critical element of the experimental apparatus was the attenuation box. This box provides sound isolation so that the air conduction path to the test ear is attenuated, leaving the body conducted signal as the dominant signal in the ear canal. The volume of the box is large (over 0.25 m^3) so that there will be no occlusion effect in the test ear.

The box was made of 2 cm thick plywood panels, with height of 104 cm, width of 54 cm and depth of 54 cm. The outside and inside of the box were covered with 10.2 cm (4") fiberglass in order to minimize reflections. A hole was cut in the front panel of the plywood box and a rectangular Plexiglas tube installed. This tube was 15.5 cm x 15.5 cm and extended out 10.5 cm from the box. A removable front plate, also made of Plexiglas, was made to fit to the end of the tube. A doughnut cushion was attached to this front plate. For an experiment, the cushion was placed over the subject's ear and the plate attached to the head with Velcro straps. It was then easy to place the probe microphone in the ear canal and buckle the plate to the box. It was assured that the plate was attached tightly to the box.

The cushion was fit around the pinna, so that the effect of the pinna on the radiated sound was essentially the same as for an uncovered ear. The effect of the box on the radiation from the ear canal is insignificant at frequencies below about 1.3 kHz.

Measurement of the attenuation provided by the box was not trivial and could not be made with a head-and-torso simulator because it could not be attached to the box. In order to get as close as possible to the experimental situation, where the subject would be vocalizing, a small loudspeaker was placed close to the mouth and chin of the subject. The speaker was placed either at a 0^o or at a 45^o angle to the subject's nose and the speaker center was raised to mouth height. A white noise test signal was used, giving 74 dB SPL at the position of the subject. The probe tube was inserted into the subject's ear canal to a depth corresponding to a standard earmold made especially for that subject (about 8 mm + 3 mm). The sound pressure was measured first without the attenuation box. This simulated free field situation was performed in the empty chamber, anechoic above 200 Hz. The box was then placed in the chamber and the test ear attached to the box. The box attenuation was calculated as the difference in sound pressure level between the free field condition and the boxed ear condition.

The box attenuation was about 5 dB different for the two speaker positions with the 45^o position providing the larger attenuation. Averaging the 45^o and 0^o results, we find that the box attenuates sound by **more than 25 dB** for frequencies between 250 Hz and 3 kHz.

4.3. Pressure from the Cushion

The subject had to lean against the cushion to avoid leakage between the head and the cushion, but there is a risk that the mechanical motion of the bones and tissue will be different than if the head was not attached to the cushion. The effect of the cushion pressure was measured with the subject sitting in the anechoic room. The front plate was detached from the box and strapped on the subject's head. The sound pressure in the ear canal was measured while the subject read aloud with the front plate either very loose, but still without leak, or very tight. The difference in sound pressure between loose and tight attachment is a measure of the effect of cushion pressure. If the ear canal is open, the difference is 2.5 dB or less below 2 kHz. This is comparable to the error of repeatability so cushion pressure effects may be ignored.

4.4. Reference Microphone

The reference microphone was positioned at the non-boxed ear at the top of the pinna and 1 cm out so it did not touch the skin. The effect of moving the reference microphone a few centimeters upwards, downwards and sideways was checked using a KEMAR mannikin in an empty anechoic room. A speaker was positioned 3 cm in front of the mouth of the mannikin with the centerline in mouth height. Displacement of the microphone 2.5 cm downwards and 1 cm backwards gave the largest measured difference of +1.25 dB at f < 2 kHz.

The disturbance of the box was also measured, using the same set up. The signals picked up by the reference microphone in the free field conditions were compared to those with the box in place. At frequencies below 1 kHz, the disturbance was of the order of measurement error.

5. RESULTS

5.1. Continuous Speech

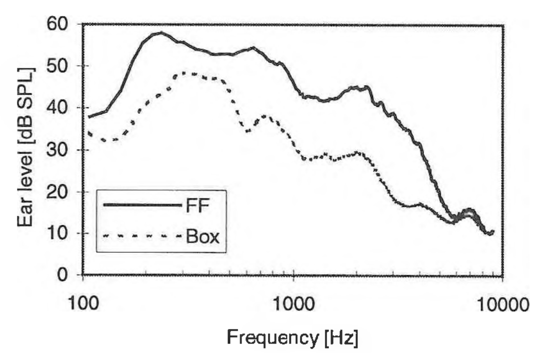
The long-term power spectrum of continuous speech was calculated with the Welch method:

$$P_{xx}(\omega) = \frac{1}{K}\sum_{i=1}^{K}\left|\sum_{n=0}^{M-1} x^i(n)w(n)e^{-j\omega n}\right|^2 ;i=1,2,\dots,K \quad (5)$$

where K is the number of windows, M is the window length (512 samples) and w is a Hanning window. An overlap of 50% was used. The optimal overlap is 75% but to save process time the overlap was reduced.

Spectra for the probe microphone recordings are shown in Fig. 5 for both left and right ears, comparing spectra with the attenuation box in place and with no box. Spectra are calibrated to correspond to dB SPL and have been corrected for the probe microphone frequency characteristics.

Right ear



Left ear

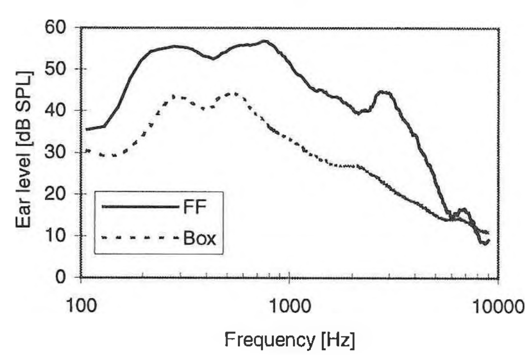


Figure 5. Real ear spectra using continuous speech. FF: free ear, Box: boxed ear. The spectra in each panel have not been normalised to account for differences in speech levels.

The measurements obtained for boxed ear and free field conditions have not been measured simultaneously so we must consider that the subject may have spoken with a different rms-level in the two situations. Therefore, it is necessary to normalize the boxed ear sound pressure level with respect to the free ear level, using the difference in sound power of the reference microphone signals. The normalization factor used is the power ratio,

$$\text{ratio} = \frac{1/2\pi \sum S_{\text{boxed}}(n)}{1/2\pi \sum S_{\text{freefield}}(n)} \tag{6}$$

where S_{boxed} is the spectrum measured by the reference microphone when one ear is boxed and the other open, $S_{\text{freefield}}$ is the reference spectrum when both ears are open (and the anechoic room empty, the simulated free field condition), and n is the frequency index.

With the above normalization, the ratio between the sound pressure in the open ear and in the boxed ear was calculated using Eq. (4). The ratio P_C/P_A is shown in Fig. 6 for the left and right ear of one subject.

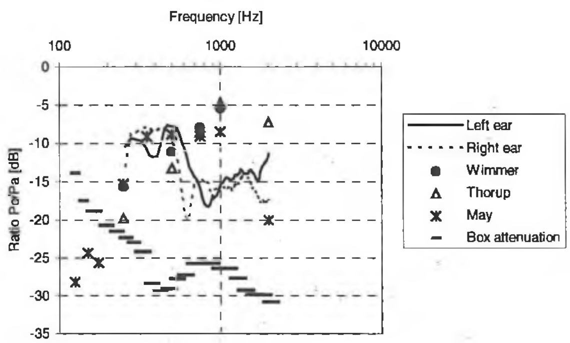


Figure 6. Ratio of body conducted and air conducted sound using continuous speech (1 subject). Measured ratios are compared with estimated values derived from literature data. Speech 1/3 octave smoothed.

Included on this figure is the measured box attenuation. Between 250 Hz and 2 kHz, the box attenuation is at least 6 dB below the pressure ratio. The error due to neglect of the box is then less than 1.3 dB over this frequency range. Below 200 Hz, the measured P_C/P_A is unreliable.

Also shown on Fig. 6 are pressure ratios estimated from occlusion effect data. The estimates make use of a model[18] for computing the function β and use Eq. (2) to transform the occlusion effect data of Fig. 3 to predicted ratios P_C/P_A. The Wimmer and Thorup data was measured using very tightly fitted earmolds, but the data from May and Dillon was measured with standard earmolds and effect of leakage[19] between the ear canal wall and the earmold had to be accounted for in the calculations. The estimates are consistent with our measurements. In the range 300-600 Hz, the measured data lie in the same range as Békésy's values[2] for loudness (-10 dB to 0 dB). Above 700 Hz the measured ratio is about 5 dB lower than measured by Békésy. In the Békésy experiments, short words were used and these sounds will have most of their energy at the fundamentals of the vowels, e.g. up to about 700 Hz.

The spectra for the right and left ears are similar but not exactly the same. This is due to the anatomical asymmetry that is particularly important for body conducted sound transmission.

5.2. Vowels

The ratio P_C/P_A was also measured for the three cardinal vowels /a/, /i/ and /u/. Each vowel was repeated three times. The subject was trained to reproduce the fundamental frequency within 30 Hz precision. The levels were also reproduced. It is known that the speech spectrum changes when a voice is softer or louder than normal speech level[20]. But in the pilot experiment a soft, normal and a loud level vowel did not give a consistent change in the ratio of boxed ear and free ear level. This indicates that the ratio of body to air conducted sound is level independent as long as the sounds are produced in the same way, i.e., whispering might not give the same result.

The differences between the boxed ear and the free field ear levels for the left ear are shown in Fig. 7; evaluations were performed using specific harmonics of the spectra. The differences correspond well to that for continuous speech. The higher harmonics of the /a/ were well reproduced and the ratio corresponds well with the ratio found for continuous speech.

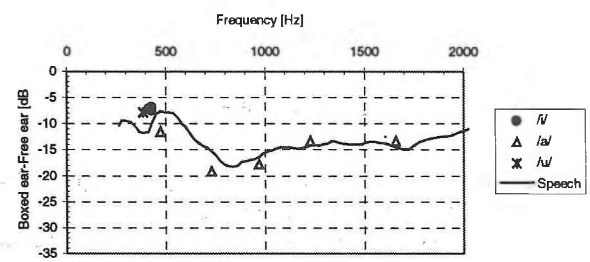


Figure 7. Ratio, for specific harmonics, of boxed ear to free ear levels produced by speech and vowels. Left ear. Note, the linear frequency axis. Speech 1/3 octave smoothed.

6. DISCUSSION

Several factors affect the accuracy of the speech spectra being evaluated. The reference signal was measured with an uncertainty of 2 dB. Considering all sources of uncertainty, the measured pressure ratio P_C/P_A is believed to be accurate to within 5 dB.

The measurements indicate that a significant fraction of the sound pressure in the ear canal is due to body conduction. To show that the sound pressures generated through body conduction are physically reasonable, the required motion of the canal walls can be estimated using the same ear canal model that was used to estimate β earlier. The body source is treated as a piston in the ear canal wall. The volume velocity entering the canal is $q = P_C / Z_C$, where P_C is the pressure generated in the canal and Z_C is the acoustic impedance presented by the canal. The corresponding piston displacement for angular frequency ω has an amplitude

$$x = q/(A\omega) \qquad (7)$$

where the effective piston area is A. The result of this calculation is shown in Fig. 8. The displacement at 300 Hz is approximately 0.05 μm. This is much less than the 20% change in diameter of the ear canal that was observed by Oliveira[7] during opening of the mouth. It must be concluded that the measured sound pressures correspond to cartilage displacements that are physically reasonable. The decrease of the displacement with frequency is consistent with the known glottal source spectrum and Eq. (7). It is known that the impedance of soft tissue increases with frequency[21] up to about 15 kHz where it becomes constant.

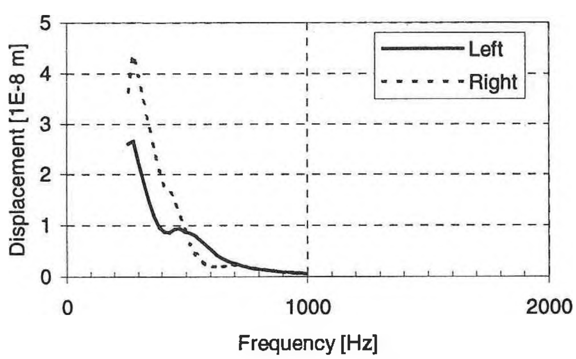


Figure 8. Displacement of the ear canal wall for body conducted sound due to own voice. The curves are valid above 250 Hz; at lower frequencies there was inadequate box attenuation. Note the linear frequency axis.

There has been considerable study of the transmission of air borne sound around the head, head-related transfer functions (HRTF). The body conduction path requires body-related transfer functions (BRTF). Sound transmission through muscles and soft tissue at audio frequencies has not been given as much attention. Some measurements of the point skull impedance have been made for bone-anchored hearing aids, e.g., Håkansson and Carlsson[22], but here the bone is directly stimulated and the results cannot be applied directly to speech production. Ishizaka *et al.*[23] measured the mechanical impedance on the cheek and neck in order to estimate the impedance of the vocal tract walls. Unfortunately they only measured up to 160 Hz. The impedance of the skin and soft tissue has been measured as a point impedance in most cases on the finger or the arm[21]. These data, although necessary and useful, are not sufficient in themselves to lead to the development of a model for body conduction of sound.

There are several body conduction routes that need to be addressed in developing a model. For example, consider Fig. 9 in which the open ear spectrum for a vowel sound is compared to the spectrum obtained with the ear closed with the attenuation box. The boxed ear spectrum shows the same resonant structure as the free ear spectrum and the formant (F1 and F2) are somewhat represented in the boxed ear signal too. The formants are an effect of the acoustic of the vocal tract length and the shape of the oral cavities. This comparison indicates that the body conducted sound is produced, not only by vibration of the vocal cords and attached structures, but through pressure changes in the vocal tract and oral cavity.

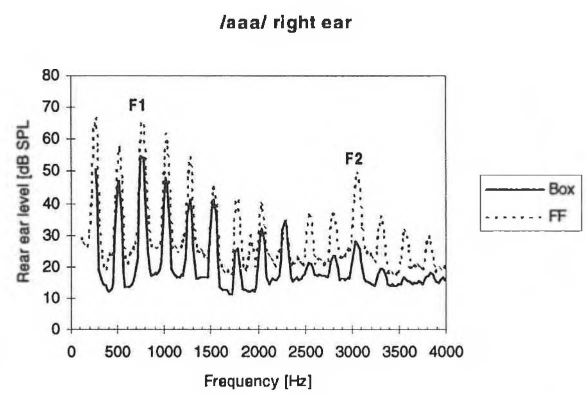


Figure 9. Spectrum of /a/ measured in the ear canal on the free ear and boxed ear. Note, the linear frequency axis.

7. CONCLUSION

The ear canal sound pressure produced by own voice has been investigated. A pilot study of the ratio between body conducted and air conducted sound was performed. The pilot study provides new information on sound produced by one's own voice. Previous studies have concentrated on bone conductor stimuli.

The measured spectra for vowel sounds indicate that body conducted sound originates from pressure changes in the oral cavity, vocal tract and vibrations of the vocal folds. The observed real ear levels due to vibrations of the soft tissue in the ear canal have been shown to be plausible, but the transmission mechanisms need further investigation.

The measurements here provide important input data for the modelling of the sound transmission in the ear canal and the occlusion effect. This modelling will be the main object of interest in the proceeding of the Ph.D. project[18]. The pilot experiment described in this paper has lead to further study by Carsten Bremmelgaard[24] in a master degree project. In this work, the ratio P_C / P_A is being measured in nearly the same way as in the pilot experiment, using 10 subjects (21 to 33 years).

REFERENCES

1. Schroeter J. and Poesselt C. (1986), "The use of acoustical test fixtures for the measurement of hearing protection attenuation. Part II: Modeling the external ear, simulating bone conduction, and comparing test fixture and real-ear data", J. Acoust. Soc. Am., Vol. 80, No. 2, p. 505-527.

2. Békésy G.V. (1949), " The structure of the middle ear and the hearing of one's own voice", J. Acoust. Soc.Am., Vol. 21, No.3, p. 217-232.

3. Hansen M.Ø. (1997), "Occlusion effects, Part I. Hearing aid users experience of the occlusion effect compared to the real ear sound level", Dept. of Acoustic Technology, Report No. 71, Technical Uni. of Denmark, ISSN 1397-0542

4. Borden G.J. and Harris K.S. (1980), " Speech Science Primer", Williams & Wilkins, ISBN 0-683-00941-9

5. Byrne D., Dillon H. and Tran K. (1994), "An international comparison of long-term average speech spectra", J. Acoust. Soc. Am., Vol. 96, No. 4 , p. 2108-2120.

6. Salvinelli F., Maurizi M., Calamita S., D'alatri L., Capelli A. and Carbone A. (1991), "The external ear and the tympanic membrane: A three-dimensional study", Scand. Audio., Vol. 20, p. 253-256.

7. Oliveira R.J. (1995), "The dynamic ear canal", in: Ballachanda (ed.) "The human ear canal", Singular Publishing Group, p. 84-111.

8. Stinson M.R. and Lawton B.W. (1989), "Specification of the geometry of the human ear canal for the prediction of sound-pressure level distribution", J. Acoust. Soc. Am., Vol. 85 No. 6, p. 2492-2503

9. Rabbitt R.D. and Holmes M.H. (1988), "Three-dimensional acoustic waves in the ear canal and their interaction with the tympanic membrane", J. Acoust. Soc. Am., Vol. 83, No. 3, p. 1064-1080.

10. Stinson M.R. (1985), "The spatial distribution of sound pressure within scaled replicas of the human ear canal", J. Acoust. Soc. Am. Vol.78, No.5, p.1596-1602.

11. Zwislocki J. (1962), "Analysis of the middle-ear function. Part I: Input impedance", J. Acoust. Soc. Am. Vol. 52, p.1514-1523.

12. Borghese N. A., Ferrigno G., Redolfi M. and Pedotti A. (1997), "Automatic intergrated analysis of jaw and lip movement in speech production", J. Acoust. Soc. Am. Vol. 101, No. 1, p. 482-487

13. Franke E.K, Gierke H.E., von Grossman F.M. and von Wittern W.W. (1952), "The jaw motions relative to the skull and their influence on hearing by bone conduction", J. Acoust. Soc. Am., Vol 24, no 2, p.142-146

14. Howell P., Williams M. and Dix H. (1988), "Assesment of sound in the ear canal caused by movement of the jaw relative to the skull", Scand. Audiol. Vol. 17, no. 2, p.93-98.

15. Wimmer V.H. (1986), "The occlusion effect from earmoulds", Hearing Instruments, Vol. 37, No.12, p.19+57.

16. Thorup A. (1996), "Okklusion", (in Danish), Master thesis at the Dept. of Acoustical Technology, Technical Univ. of Denmark.

17. May A. and Dillon H. (1992), "Comparison of physical measurements of the occlusion effect with subjective reports", Paper presented at the Audiological Soc. Conference, Barossa Valley National Acoustic Laboratory, Sydney.

18. Hansen M.Ø. (1997), "Occlusion effects. Part II", Ph.D. Thesis, Dept. of Acoustic Technology, Report No. 73, Technical Univ. of Denmark, ISSN 1397-0548.

19. Macrae J. and McAlister P. (1989), "A mathematical model of acoustic leakage through pathways past earmoulds", Australian Jour of Audiology, Vol 11, No. 2, p. 89-100.

20. Elberling C. and Nielsen C. (1993), "The dynamics of speech and the auditory dynamic range in sensorineural hearing impairment", In Beilin J. and Jensen G.R. (ed.): "Recent developments in hearing instrument technology", 15th Danavox Symposium, p.99-132.

21. Gierke H.E. and Brammer A.J. (1995), "Effects of shock and vibration on humans", In Harris C.M. (ed.): Shock and vibration handbook, 4.th. ed., McGraw-Hill, p.44.2-44.19

22. Håkansson B. and Carlsson P. (1986), "The mechanical point impedance of the human head, with and without skin penetration", J. Acoust. Soc. Am., Vol. 80, No. 4 , p. 1065-1075.

23. Ishizaka K., French J.C. and Flanagan J.L. (1975). "Direct determination of vocal tract wall impedance", IEEE Transactions on acoustics, speech and signal processing, Vol. ASSP-23, No.4, p.370-373.

24. Bremmelgaard C. (1997), "Overføringsfunktionen mellem lufledt of kropsledt lyd og okklusionseffect ved tale", (in Danish), Master Thesis, Dept. of Acoustic Technology, Technical Univ. of Denmark

Instructions for the Preparation of Abstracts

1) Duplicate copies of an abstract are required for each meeting paper; one copy should be an original. Send the copies to the Technical Program Chairperson, in time to be received by the deadline. Either English or French may be used. A cover letter is not necessary. 2) Limit the abstract to 250 words, including title and first author's name and address; names and addresses of coauthors are not counted. 3) Title of abstract and names and addresses of authors should be set apart from the abstract. Text of abstract should be one single, indented paragraph. The entire abstract should be typed double spaced on one side of 8 1/2 x 11 in. or A4 paper. 4) Be sure that the mailing address of the author to receive the acceptance notice is complete on the abstract, to insure timely deliveries. 5) Do not use footnotes. Use square brackets to cite references or acknowledgements. 6) At the bottom of an abstract give the following information: a) if the paper is part of a special session, indicate the session; b) name the area of acoustics most appropriate to the subject matter; c) e-mail address (if available), telephone and fax numbers, including area code, of the author to be contacted for information. Authors outside of Canada should include country; d) If more than one author, name the one to receive the acceptance notice; e) Overhead projectors and 35mm slide projectors will be available at all sessions. Describe on the abstract itself any special equipment needed.

Electronic submission of an abstract may be made via e-mail to cheesman@audio.hhcru.uwo.ca. Electronic submission is encouraged; abstracts will be posted on the CAA Web page (http://www.uwo.ca/hhcru/caa/awc98).

Instructions pour la Préparation des Résumés de Conférences

1) Deux copies du résumé sont requises pour chaque papier soumis; une des copies doit être un original. Envoyer les copies au Président du Comité technique, suffisamment à l'avance pour qu'elles soient reçues avant la date de tombée. L'anglais ou le français peut être utilisé. Une lettre de présentation n'est pas requise. 2) Limiter le résumé à 250 mots, incluant le titre, le nom et l'adresse du premier auteur; les noms et les adresses des co-auteurs ne sont pas comptabilisés. 3) Le titre du résumé, les noms et les adresses des auteurs doivent être séparés du texte. Le texte du résumé doit être présenté en un seul paragraphe. Le résumé entier doit être dactylographié à double interlignes sur une face d'une page 8 1/2 x 11 pouce ou du papier A4. 4) S'assurer que l'adresse postale complète de l'auteur qui doit recevoir l'avis d'acceptation est inscrite sur le résumé afin d'assurer une livraison rapide. 5) Ne pas utiliser les notes de bas de page. Utiliser les crochets pour les références et les remerciements. 6) A la fin du résumé, fournir les informations suivantes: a) si la communication fait partie d'une session spéciale, indiquer laquelle; b) identifier le domaine de l'acoustique le plus appropié à votre sujet; c) l'adresse d'e-mail (si disponible), les numéros de téléphone et de télécopieur, incluant le code régional, de l'auteur avec qui l'on doit communiquer pour information. Les auteurs étrangers doivent indiquer leur pays; d) S'il y a plus d'un auteur, mentionner le nom de celui qui doit recevoir l'avis d'acceptation; e) Des projecteurs à acétates et à diapositives seront disponibles dans chaque session. Indiquer les besoins spéciaux, si nécessaire.

Le résumé peut être parvenu via courrier électronique à cheesman@audio.hhcru.uwo.ca. Les soumis électroniques sont encouragés; les résumés seront publiés aussi au 'World Wide Web' (à l'adresse http://www.uwo.ca/hhcru/caa/awc98).

Instructions pour la Préparation des Articles à être Publiés dans le Cahier des Actes du Congrès

Général - Soumettre un article prêt-à-copier d'un maximum de deux pages présenté en deux colonnes. Ne pas inclure de sommaire. Tout le texte en caractères Times-Roman. Disposer les figures dans le haut ou le bas des pages si possible. Lister les références dans un format logique à la fin du texte. Envoyer l'article au président du Programme Technique avant la date de tombée . Le format optimal peut être obtenu de deux façons:

Méthode directe - Imprimer directement sur une feuille 8.5" x 11" en respectant des marge de 3/4" dans le haut et sur les côtés et un minimum de 1" dans le bas. Titre en 12pt, caractères gras, en simple interligne (12pt), centrés sur la page. Le reste du texte en 9pt en 0.75 (9pt) interligne, dans un format en deux colones, avec une largeur de colonnes de 3.4" et une séparation de 1/4". Noms des auteurs et adresses centrés sur la page avec les noms en caractères gras. Les titres de sections en caractères gras.

Méthode indirecte - Dactylographier ou imprimer comme suit, réduire au trois-quart (s.v.p., s'assurer de bonnes photocopies) et assembler l'article sur un maximum d'une page 8.5" x 11" avec les côtés et un minimum de 1" dans le bas. Titre en 16pt avec 1.33 (16pt) interligne, centré sur la page. Le reste du texte en 12pt avec simple (12pt) interligne. Noms et adresses des auteurs centrés sur la page avec les noms en caractères gras. Titres des sections en caractères gras. Imprimer les colonnes de texte sur quatre feuilles 8.5" x 14" avec une largeur de colonnes de 4.5", une longueur maximum de 12.25", en laissant de la place pour le titre, les noms et les adresses sur la première page.

Instructions for Preparation of Articles to be Published in the Conference Proceedings Issue

General - Submit the camera-ready article on a maximum of two pages in two-column format. Do not include an abstract. All text in Times-Roman font. Place figures at the top and/or bottom of the pages, if possible. List references in any consistent format at the end. Send to the Chairperson of the Technical Programme by the deadline. The optimum format can be obtained in two ways:

Direct method - Print directly on one sheet of 8.5" x 11" paper with margins of 3/4" top and sides, and 1" minimum at the bottom. Title in 12pt bold with single (12pt) spacing, centred on the page. All other text in 9pt with 0.75 (9pt) line spacing, in two-column format, with column width of 3.4" and separation of 1/4". Authors' names and addresses centred on the page with the names in bold type. Section headings in bold type.

Indirect method - Type or print as follows, reduce to three-quarters size (please ensure good copies) and assemble article on a maximum of one 8.5" x 11" page with margins of 3/4" top and sides, and 1" minimum at the bottom. Title in 16pt bold type with 1.33 (16pt) line spacing, centred on the page. All other text in 12pt with single (12pt) line spacing. Authors' names and addresses centred on the page with the names in bold type. Section headings in bold type. Print individual text columns on four sheets of 8.5" x 14" paper with a column width of 4.5", a maximum length of 12.25", and leaving room for the title and names and addresses on the first page.