Audio Engineering Society

Convention Paper 6492

Presented at the 118th Convention
2005 May 28–31 Barcelona, Spain

Balloons of Directivity of Real and Artificial Mouth Used in Determining Speech Transmission Index

Fabio Bozzoli[1], Michel Viktorovitch[2] and Angelo Farina[3]

[1] Industrial Engineering Dept., University of Parma, Parma, Italy
bozzoli@ied.unipr.it

[2] Rieter Automotive Management AG, Winterthur, Switzerland
Michel.Viktorovitch@rieterauto.com

[3] Industrial Engineering Dept., University of Parma, Parma, Italy
farina@unipr.it

ABSTRACT

In room acoustics, one of the most used parameters for evaluating the speech intelligibility is the Speech Transmission Index (STI). The experimental evaluation of this STI generally employ an artificial speaker (binaural head) and listener (artificial mouth). In many cases (i.e. big rooms or system of telecommunications) the precision of directivity of artificial mouth doesn't influence too much the result; on the contrary inside cars, but also in other cases, the shape of the whole balloon of directivity is important for determining correct and comparable values and different mouths give really different results in the same situation. For these reasons we have measured, in an anechoic room, the 3D directivities of a statistical population of speakers. The post-processing of the results enabled us to determine the average and the standard deviation of human speech directivity. These results constitute a valuable source of information for assessing the compliances of artificial mouth to reality.

1. INTRODUCTION

The optimal listening conditions inside a car compartment are of paramount importance for car makers, as this is one of the most relevant point in assessing the "comfort" of the vehicle. One of the intelligibility's parameters that has been reveled to be more sensitive inside cars is the Speech Transmission Index. The STI, as exposed in IEC standard n.60268-16, [1], is based on the reduction of the modulation index of a test signal simulating the speech characteristic of a real talker, due to the environment. The test signal is emitted by a sound source situated at the talker's position and is recorded by a microphone at any listener's position. This signal consists of a noise carrier with a speech-spread frequency spectrum and a sinusoidal intensity modulation at frequency F.

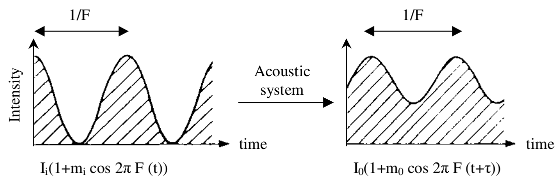

Figure 1: Modulated signal emitted by the artificial mouth (left) and received at the listener position (right), showing a smaller modulation at the receiver

This kind of measurement, in many uses, is not strongly connected to the directivity of the artificial mouth that is employed. For room acoustics applications, because of the substantial distance between speaker and receiver and the numerous reflections, the directivity of the artificial mouth doesn't influence too much the result; equally, in telecommunication acoustics, the receiver microphone is so close to the mouth that only the frontal near field affects the global intelligibility.

On the contrary, inside cars, the order of magnitude of the distance between the speaker and the listener is round one meter. Thus, only a small number of image sources are considerable and some cases (for example when the mouth is in the driver position and the receiver is in a rear seat) the signal emitted by the back of the speaker is as important as the signal stemming from the front; for all these reasons the whole balloon of directivity of the artificial mouth is important for a reliable assessment of the STI in the automotive field (Figure 2).

In other respects, there aren't many recent publications about the true human directivity and there are no norms that define the whole balloon of directivity. The only really good paper that we have been able to find is an internal report of National Research Council Canada, [2].

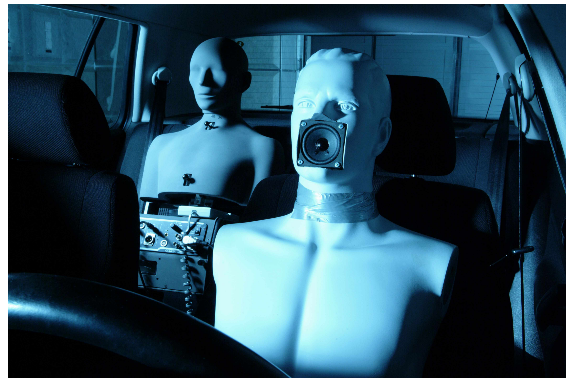

Figure 2 : Set-up for STI measurement. In foreground a Rieter artificial mouth prototype (the talker), in background, the artificial mouth (the receiver).

Hence, the principal aim of this work is to obtain the complete balloons of directivity of a statistical population of speakers, at all frequencies, on a sphere of radius 1 meter centered at the mouth of the speaker.

2. MEASUREMENTS OF HUMAN SPEECH

The measurements were carried out in an anechoic room and it was asked of 10 male speakers to repeat the same sentence, specially chosen for its wide frequency content. The directivity was investigated on a sphere of radius one meter and centered on the mouth of the speaker.

The measurements are carried out using five microphones (Figure 3).

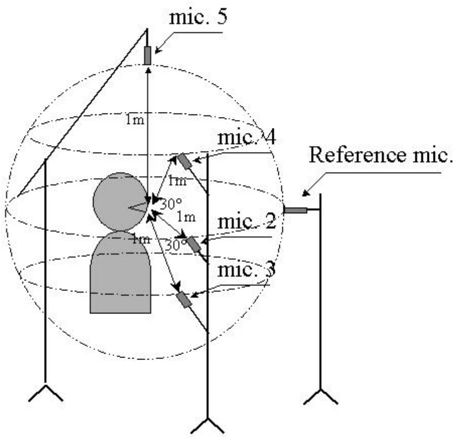

Figure 3 : The speaker repeats the sentence for each new horizontal position of the microphones (every 15 degrees).

Microphone 1 remains in the horizontal plane of the mouth in front of the speaker during all the measurements. This is the reference microphone.

Three other microphones are positioned on a mobile stand. They are located in a vertical plane and they always remain one meter far from the mouth. Microphone 2 is in the horizontal plane passing by the mouth, microphone 3 is located 30 degrees below, and microphone 4 is positioned 30 degrees above.

A measurement is also performed with a microphone located vertically above the mouth and one meter away from the mouth.

Horizontally, the measurements are carried out every 15 degrees (by shifting the microphone stand).

For each horizontal position, the recordings for microphones 1, 2, 3 and 4 are carried out in one take with a multi-channel recorder. Each recording is then post-treated in octave frequency bands. Seven octave bands are considered: 125 Hz, 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz and 8000 Hz. Finally, for each octave band, the outputs from microphones 2, 3 and 4 are subtracted to the output of the reference microphone This procedure enables to get rid of the unavoidable differences occurring in the spectra of each repetition of the speaker.

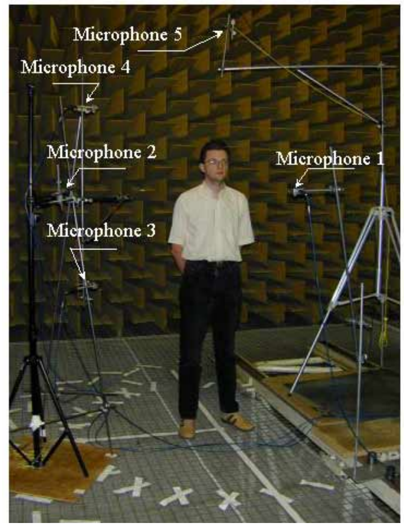

Figure 4 : Illustration of the set-up for the directivity measurement.

3. MEASUREMENTS OF ARTIFICIAL MOUTH

The first step of this work has bean measuring the directivity of a commercial artificial mouth (Brüel & Kjær mouth simulator type 4230). This mouth is built according with ITU recommendations, [3]. Inside an anechoic room we have employed MLS-based impulse response, which was repeated hundredths of times, rotating the transducer by means of an automated turning table. It has been exposed in [4].

Figure 5 : Brüel & Kjær mouth simulator type 4230 on rotating table.

The results can be easily transformed in frequency response spectra by means of an FFT: consequently, it is possible to get the frequency response for any angle of emission or listening, and to plot polar pictures of the response in a given frequency band.

4. RESULTS

We have processed directivities of ten real speaker, calculating average and standard deviation.

The results obtained for the different subjects are homogeneous, even though the deviation increases with the frequency. Maximum standard deviation is 3 dB, at 4000 Hz.

The results obtained in the framework of this study are comparable to those obtained by Chu, [1].

We have noticed that, even if B&K fits ITU recommendations, the differences with the real speaker in some directions exceed 5 dB and they are present mainly in the back of the head probably because artificial head are made of materials with a density definitely lower than real body's one.

Here we report average directivities of real speaker and directivities of B&K mouth at 1000 Hz in horizontal plane , 30 degrees below and 30 degrees below.

In figure 6-7-8 we have used dashed lines for representing the average real directivity plus standard deviation and the average real directivity minus standard deviation.

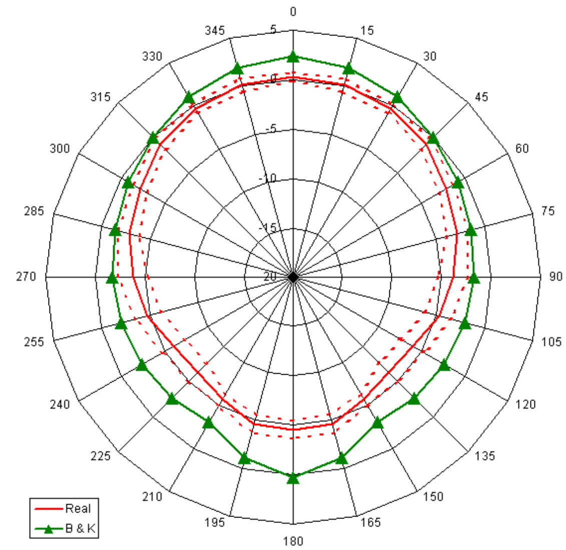

Figure 6 : directivities of real speaker and B&K mouth at 1000 Hz, 30 degrees above horizontal plane.

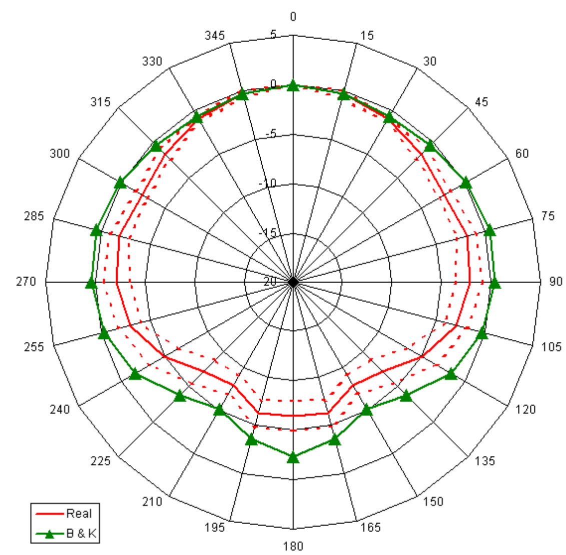

Figure 7 : directivities of real speaker and B&K mouth at 1000 Hz, horizontal plane.

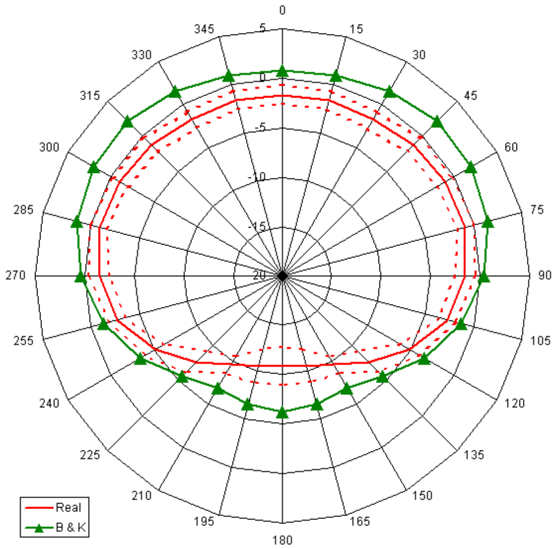

Figure 8 : directivities of real speaker and B&K mouth at 1000 Hz, 30 degrees below horizontal plane.

5. CONCLUSIONS AND FUTURE WORK

With this study we have underlined that the lack of a norm about artificial mouth's balloon of directivity generate sources with different behavior: in some cases this produce different values of Speech Transmission Index especially when mouth's lateral and posterior emission are more important than frontal one for transferring signal from speaker to listener. We have analysed ten real speaker and obtained the directivity of speech emission of an average human talker.

We have compared it with the directivity of the Bruel and Kjaer mouth simulator and we have found a good agreement.

Afterwards we'll study how much the differences between artificial mouths influence the STI determination using different artificial mouths (all according to ITU recommendations) in some critical cases for example inside car with speaker in the driving position and listener in the back seat.

6. ACKNOWLEDGEMENTS

This work was supported by Rieter Automotive (Winterthur, Switzerland).

7. REFERENCES

[1] CEI-IEC standard n. 60268-16 – "Sound System Equipments – Objective rating of speech intelligibility by speech transmission index" – Second Edition, March 1998

[2] A W.T. Chu, A.C.C. Warnock, 'Detailed Directivity of Sound Fields Around Human Talkers', Report B3144.6 prepared for Public Works and Government Services Canada, September 2001.

[3] ITU-T Recommendation P51, "Telephone Transmission Quality – Artificial Mouth" – August 1996

[4] A. Farina, F. Bozzoli, 'Measurement of the speech intelligibility inside cars', Pre-prints of the 113th AES Convention, Los Angeles, 5-8 October 2002.