Audio Engineering Society
Convention Paper
Presented at the 113th Convention
2002 October 5–8 Los Angeles, California, USA

Measurement of the speech intelligibility inside cars

Angelo Farina[1], Fabio Bozzoli[2]

[1] *Industrial Eng. Dept., Università di Parma, Via delle Scienze - 43100 Parma – Italy – E-mail: farina@unipr.it*

[2] *Industrial Eng. Dept., Università di Parma, Via delle Scienze - 43100 Parma – Italy – E-mail: bozzoli @ied.unipr.it*

ABSTRACT

The paper describes a measurement system developed for assessing the speech intelligibility inside car compartments. This relates directly to the understandability of the voices being reproduced through the radio receiving system (news, traffic info, etc.), but in the future it will be used also for assessing the direct speech communications between the passengers, and the performance of hands-free communication devices.

The system is based on the use of two head and torso simulators, one equipped with an artificial mouth, the second equipped with binaural microphones. Only the second is used when the sound is being reproduced through the car's sound system.

The MLS-based version of the STI method is used for performing the measurements, taking into account the effect of the background noise and the electro-acoustic propagation inside the compartment.

GOALS AND LIMITS

This paper reports on just the very first part of a research program, which was recently started at the University of Parma. The goal of this paper is simply to describe the hardware and software system which has been set up and validated: this system allows for the measurement of the Speech Intelligibility Index STI [1] (and of other related indexes, such as SII [2] and SIL [3]) inside car compartments, modeling properly both the head and torso of the talker and the head and torso of the listener. Depending on the various cases, only one or both torso simulators are needed.

The paper describes the transducers, which have been qualified through careful measurement in anechoic room.

Then the paper describes the software tools employed for automating the measurement.

No experimental results coming from actual measurement taken on cars are presented here: these will be the topic of subsequent papers.

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 car.

Typically "sound quality" methods were used in recent years for assessing the perceived noisiness and harshness of the background noise, but till now the effect of the noise on the speech intelligibility was obtained simply as a post-processing of the noise measurement results, based on the old A.I. (articulation index) method [5], without taking into account the effect of internal reflections, echoes and resonances inside the cavity.

Furthermore, recently highly absorbing treatment systems appeared on some cars: these systems revealed to be very effective in reducing the background noise with a limited weight of material installed inside the car, but this approach changes substantially the acoustical environment. Often problems of intercommunication between front and rear passengers have been reported.

Finally, on most cars nowadays very sophisticated sound system are installed, devoted to three main functions: radio receiver reproduction, hands-free communication over cellular phones, and infotainment systems (GPS navigation for the driver, DVD reproduction for the passengers on the rear seats). Proper listening conditions must be ensured, and the simple evaluation of the frequency response curve of the sound system revealed to be of minor significance in terms of understandability of the message, localizability of warning signals and disturbance between front and rear passengers.

It can be predicted that not necessarily the cars equipped with powerful, "high fidelity" sound systems will get the better values of STI, because the speech intelligibility is not depending too much on the "musical" performance of the system, but instead are also dependent on position and directivity of the loudspeakers.

These facts indicate that a more advanced assessment technique is required for quantifying the effective speech intelligibility inside cars. What is required is a realistic emission of sound (with directivity and waveform curvature comparable with an human talker), and a realistic pickup of it through a proper binaural probe.

This can be done, in principle, using a true human talker, and a true "listener" wearing portable binaural microphones fitted at the ear channel entrance. Although this can be practical in some situations, in general this approach is not suitable for reproducible tests. Very often, an objective measurement with high reproducibility is preferred, because only this makes it possible to assess different configurations of the car fittings, which have to be compared not simultaneously (because a single prototype of the car is available, and each system must be mounted, measured and dismounted).

For the above reasons, the measurement system developed is based on a pair of HATS (head and torso simulators). One is employed as an artificial mouth, the second as a binaural microphone. The measured quantity is the STI (Speech Transmission Index, by Houdgast and Steeneken). Two different signal processing methods were developed for doing the measurement. The first one is based on the use of an MLS (Maximum Lenght Sequence) signal, which makes it possible to measure the binaural impulse response, and to derive the MTF (Modulation Transfer Function) matrix following the approach initially suggested by M. Schroeder and refined by D. Rife [4]. This method is fast and reliable, although it tends to underestimate the effect of the background noise.

The second method (not fully implemented yet) is based on the generation of amplitude modulated signals, as it was initially suggested by Houtgast and Steeneken. At that time, this approach was very slow and difficult, but with today's computing capability it was easy to create a software capable of generating the test signal and computing the MTF matrix, and consequently the STI, from the recorded response of the cavity. This process is slower, but by definition it is the "true" STI, and constitutes a reference against which other faster measurement methods have to be assessed.

ARTIFICIAL “TALKER” AND “LISTENER”

Two independent transducers were developed for the two ends of the transmission chain: an artificial mouth simulator, mounted on a simplified head and torso model, and a standard binaural mannequin providing “microphonic ears”.

Both mannequins were heavily tested inside an anechoic chamber, and the respect of the most relevant electroacoustic standards was verified.

The tests were conducted thanks to the facilities of ASK Automotive Industries, Reggio Emilia, Italy and Rieter Automotive, Winterthur, Switzerland. They provided the anechoic chambers and most of the measuring equipment.

Fig. 1 – the complete mouth simulator mannequin

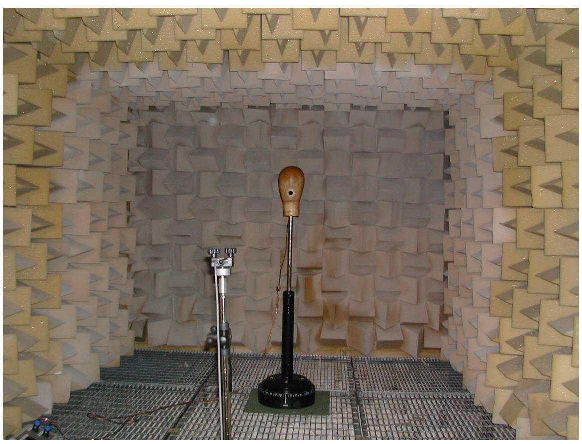


Fig. 2 – anechoic measurement of the directivity of the mouth simulator

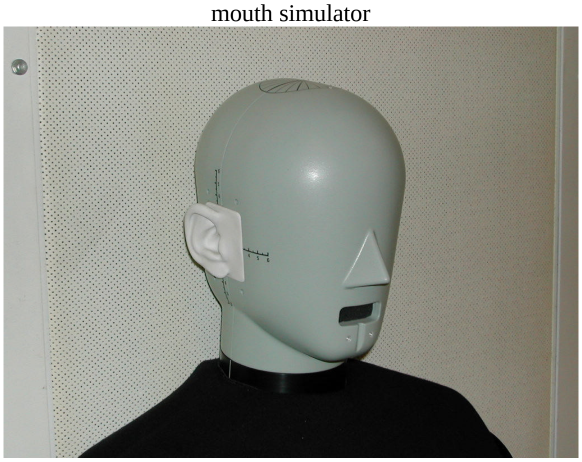


Fig. 3 – the binaural microphone set (B & K type 4100)

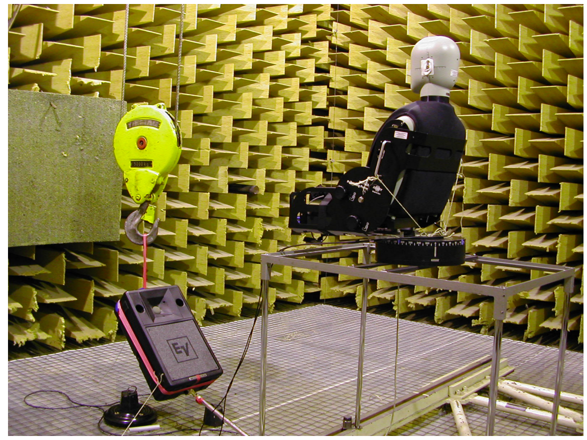


Fig. 4 – anechoic measurement of the directivity of the binaural microphone

Basically, the measure was an MLS-based impulse response, which was repeated hundredths of times, rotating the transducer by means of an automatized turning 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 bands.

Some of these results can be seen in the following page.

The directivity of the artificial mouth resulted to be within the limits of the current standard [6], and the directivity of the binaural microphone was also compliant with the specifications for dummy heads [7].

Consequently, both transducers were considered perfectly suitable for the goals of this research work.

SOFTWARE TOOLS

The measurements are performed by means of the Aurora software package [8,9]. This software allows for the generation of the MLS (maximum length sequence) signal, which is inherently white. Thanks to the filtering tools already incorporated in the host program (CoolEditPro), it is easy to apply frequency-domain equalization so that the spectrum generated in front of the artificial mouth, at 1m distance, complies with the specification of the IEC 60268-16 code. This was verified, in anechoic conditions, placing a B&K2260 real-time analyzer in the reference position.

The impulse response measurement needs to e performed on just a single repetition of the MLS period, for avoiding that the averaging process artificially increases the S/N ratio. Indeed, the measurement is performed playing continuously the

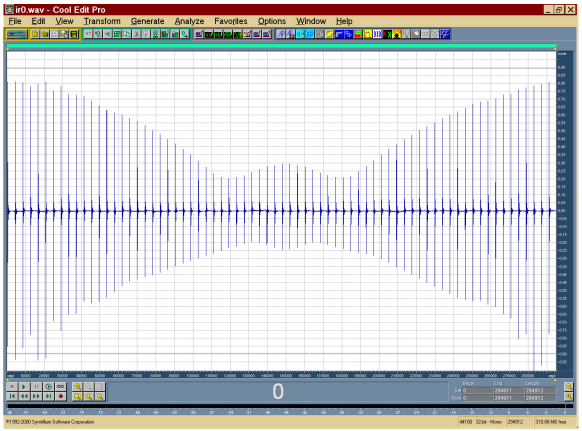


Fig. 5 – results (impulse responses) of the directivity of the artificial mouth – 0° elevation

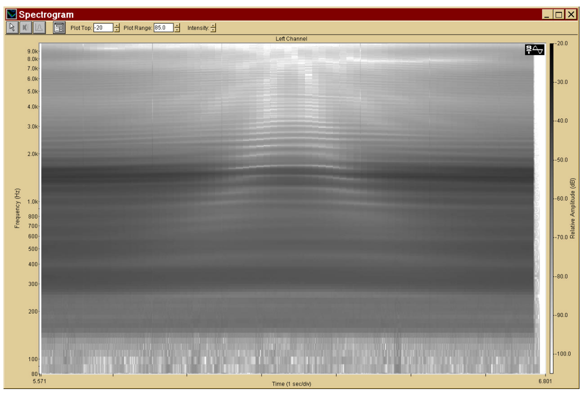


Fig. 6 – results (sonograph) of the directivity of the artificial mouth – 0° elevation

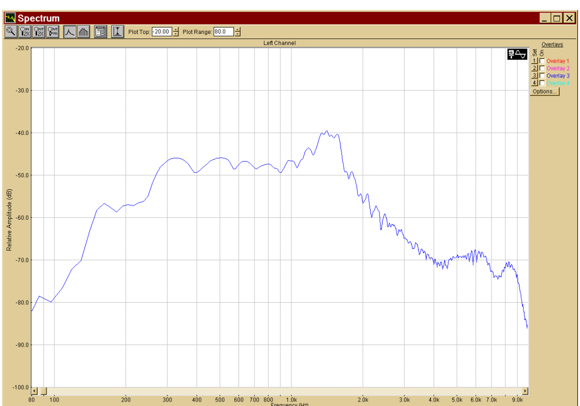


Fig. 7 – on-axis frequency response of the artificial mouth – 0° elevation

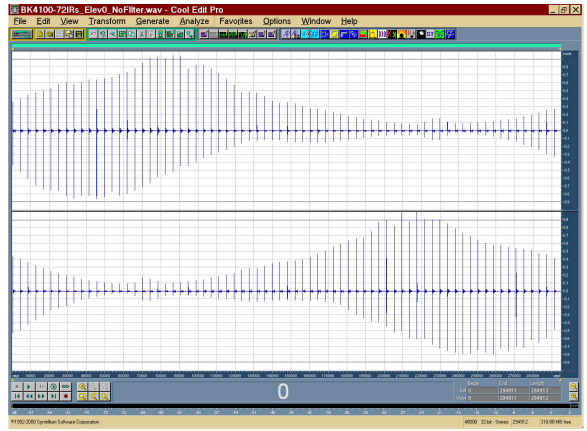


Fig. 8 – results (impulse responses) of the directivity of the binaural microphone – 0° elevation

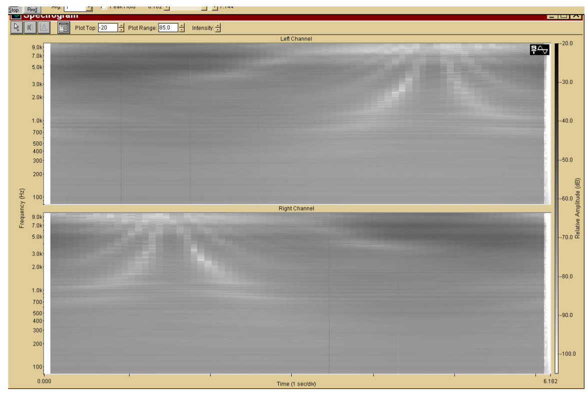


Fig. 9 – results (sonograph) of the directivity of the binaural microphone – 0° elevation

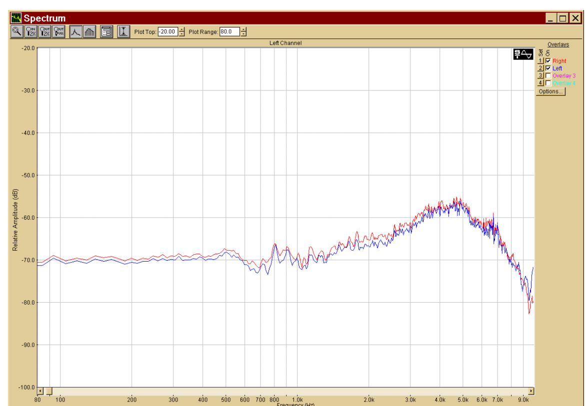


Fig. 10 – on-axis frequency response of the binaural microphone – 0° elevation

MLS signal, so that the system is perfectly in steady-state during the measurement.
After the impulse response has been recovered, it is saved in the MLSSA .TIM file format.
This allows for the subsequent post-processing with the good, old MLSSA software, which already includes the subroutine for computing the STI value.
In the future it is planned to develop a new Aurora plugin, replicating this function directly inside the CoolEdit host program.The following pictures show an impulse response measured inside a car with Aurora, and the results of the STI analysis of the same IR performed with the MLSSA software.

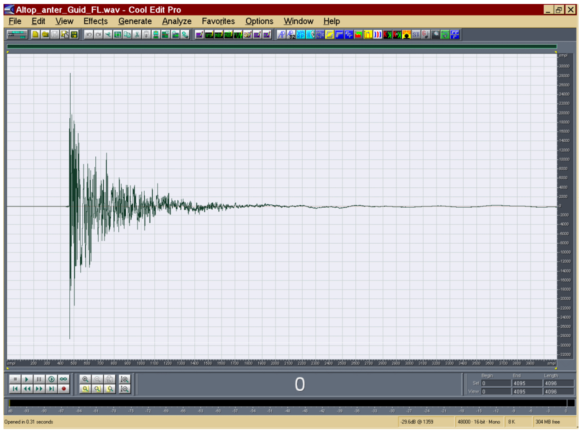
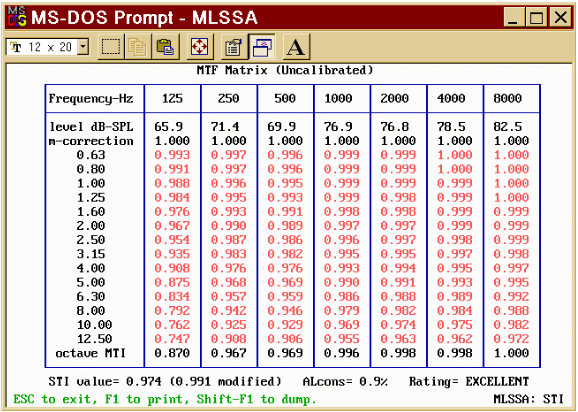


Fig. 11 – impulse response measured inside a car with the original sound system (left ear)



MS-DOS Prompt - MLSSA

MTF Matrix (Uncalibrated)

Frequency-Hz	125	250	500	1000	2000	4000	8000
level dB-SPL	65.9	71.4	69.9	76.9	76.8	78.5	82.5
m-correction	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.63	0.993	0.997	0.996	0.999	0.999	1.000	1.000
0.80	0.991	0.997	0.996	0.999	0.999	1.000	1.000
1.00	0.988	0.996	0.995	0.999	0.999	0.999	1.000
1.25	0.984	0.995	0.993	0.999	0.998	0.999	1.000
1.60	0.976	0.993	0.991	0.998	0.998	0.999	0.999
2.00	0.967	0.990	0.989	0.997	0.997	0.999	0.999
2.50	0.954	0.987	0.986	0.996	0.997	0.998	0.999
3.15	0.935	0.983	0.982	0.995	0.995	0.997	0.998
4.00	0.908	0.976	0.976	0.993	0.994	0.995	0.997
5.00	0.875	0.968	0.969	0.990	0.991	0.993	0.995
6.30	0.834	0.957	0.959	0.986	0.988	0.989	0.992
8.00	0.792	0.942	0.946	0.979	0.982	0.984	0.988
10.00	0.762	0.925	0.929	0.969	0.974	0.975	0.982
12.50	0.747	0.908	0.906	0.955	0.963	0.962	0.972
octave MTI	0.870	0.967	0.969	0.996	0.998	0.998	1.000

STI value= 0.974 (0.991 modified) ALcons= 0.9% Rating= EXCELLENT

ESC to exit, F1 to print, Shift-F1 to dump. MLSSA: STI

Fig. 12 – STI evaluation performed by means of the MLSSA software

FUTURE WORK AND CONCLUSION

The measuring system described in this paper will subsequently be employed for performing a comparative study between some cars, for evaluating the sensitivity of the system to the different variables: background noise (which of course depends on the speed and RPMs of the engine), absorption inside the cockpit, number of passengers inside the car.

Other measurements will be conducted for assessing the intelligibility of electronics communication inside the car: understandability of the news transmitted by the radio, two-way intelligibility over an hands-free telephone system, effectivity of warning messages from the navigation system.

In conclusion it is now clear that looking seriously at the speech intelligibility in cars will bring to the design of sound insulation, sound absorption and electronic sound reproduction systems which are significantly different from those optimized looking at other criteria, such as the reduction of the background noise and the performance for music reproduction. The tools developed for this purpose and presented here will be the practical means to check that the design's goals have been attained.

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] ANSI standard S3.5-1997 – "Methods for calculation of the speech intelligibility index" – (1997).

[3] ISO standard 9921-1 – "Speech interference level and communication distances for persons with normal hearing capacity in direct communication (SIL method)" – 1993.

[4] D. Rife, "Modulation Transfer Function Measurement with Maximum-Length Sequences", (JAES), vol. 40, No. 10, pp 779-790, 1992 October.

[5] ANSI standard S3.5-1969 – "Methods for calculation of the Articulation Index" – (1969).

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

[7] ITU-T Recommendation P58, "Telephone Transmission Quality – Head and torso Simulator for Telephonometry" – August 1996

[8] A. Farina, F. Righini, 'Software implementation of an MLS analyzer, with tools for convolution, auralization and inverse filtering', Pre-prints of the 103rd AES Convention, New York, 26-29 September 1997.

[9] A. Farina, E. Ugolotti, "Automatic Measurement System For Car Audio Application", Pre-prints of the 104th AES Convention, Amsterdam, 15 - 20 May, 1998.