euronoise
Naples 2003

paper ID: 266/p.1

Comparative Study of Speech Intelligibility Inside Cars

A.Farina[a], F.Bozzoli[b] and P.Strasser[c]

[a]*Industrial Eng. Dept., Università di Parma, Via delle Scienze, 43100 Parma, Italy, E-mail: farina@unipr.it*
[b]*Industrial Eng. Dept., Università di Parma, Via delle Scienze, 43100 Parma, Italy,*
E-mail: bozzoli@ied.eng.unipr.it
[c] *Rieter Automotive Systems, Schlosstalst, 43 CH 8406 Winterthur, E-mail: patricia.strasser@rieterauto.com*

The paper describes the results of a comparative study between two different groups of techniques for computing Speech Transmission Index inside cars. The measurement system is composed by head and torso simulators: one group is composed by techniques based on measurement of the impulse response that directly includes background noise; on the contrary, the other group contains techniques based on a noise-free measurement of the impulse response and a subsequent correction for the effect of background noise. Inside cars, where the signal to noise ratio is low, the technique based on noise-free impulse response shows, compared to the other, its strength in terms of systematic error and standard deviation.

1. INTRODUCTION

The optimal listening conditions inside a car compartment are of paramount importance for carmakers, as this is one of the most relevant points in assessing the "comfort" of the car. Typically, “sound quality” methods were used for assessing the perceived noisiness and harshness of the background noise without taking into account the effects of internal reflections, echoes and resonances inside the cavity. The parameter that is able to consider all these effects is the Speech Transmission Index: the methods for determining it, exposed in IEC standard n.60268-16 [1], are based on the reduction of the modulation index m_i of a test signal simulating the speech characteristic of a real talker, when emitted in an acoustic environment. The test signal is transmitted by a sound source situated at the talker’s position to a microphone (or better, to a binaural dummy head) at any listener’s position and it consists of a noise carrier with a speech-spread frequency spectrum and a sinusoidal intensity modulation at frequency F (see Figure 1).

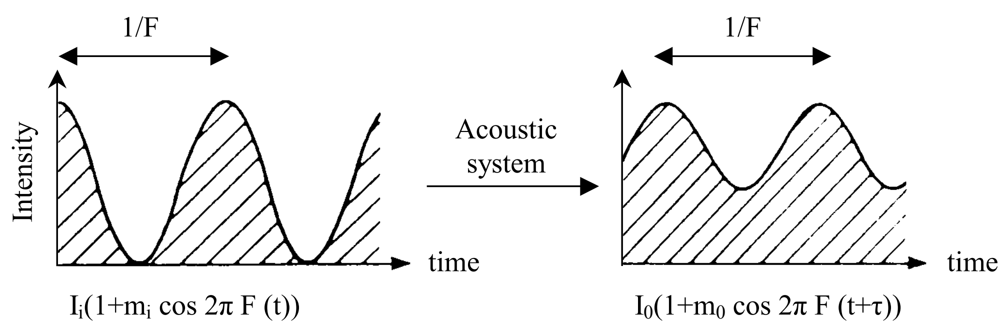

Figure 1. *Modulated signal emitted by the artificial mouth (left) and received at the listener position (right), showing a smaller modulation at the receiver .*

The reduction in the modulation index is quantified by the modulation transfer function m(F) which is determined by :

$$m(F) = \frac{m_0}{m_i} \tag{1}$$

The STI is got from these modulation transfer functions, taking in account auditory masking and absolute hearing threshold, and with the octave weighting factors given in [1]. STI goes from 1.0, when the intelligibility is optimal, to 0.0 when it's not possible to understand anything .

2. MODULATION TRANSFER FUNCTION

The methods for determining MTF can be divided by two big groups: methods based on test signals sinusoidally modulated in intensity and methods based on measurement of the impulse response of the system.

The firsts follow exactly the definition of MTF using a test signal with a modulation index of one, at each of 7 frequencies of the octave-band-filtered noise carrier and at each of 12 modulation frequencies (one at time) recommended by [1]: they are used seldom because of long time needed for 84 measurements and because of problems generated by unsteady background noise.

The seconds derives the MTF values from a single impulse response measurement h(τ), as initially suggested by M. Schroeder and refined by D. Rife [2] , using the formula:

$$m(F) = \frac{\int_0^\infty h_f^{\,2}(\tau)\cdot \exp(-j\cdot 2\cdot \pi \cdot F\cdot \tau)\cdot d\tau}{\int_0^\infty h_f^{\,2}(\tau)\cdot d\tau} \tag{2}$$

where $h_f(\tau)$ is impulse response octave-band filtered at carrier frequency f.

Another division can be made inside this group: techniques based on "noisy Impulse Response" and techniques based on "noise free Impulse Response".

"Noisy IR " measures Impulse Response directly using a single repetition of a MLS sequence, filtered and amplitude calibrated so that it adheres strictly to the normalized spectrum of the human speech, in presence of background noise; in this way the background noise is superposed to the impulse response and the m(F) values are measured correctly using (2).

On the contrary "Noise free IR" methods measure Impulse Response in absence of background noise, making use of special techniques (for example multiple MLS or Sweep signal) for maximizing signal to noise ratio. The real m(F) can be derived calculating m'(F), obtained applying (2) to noise free IR, and taking into account for the effect of background noise with the following expression:

$$m(F) = m'(F) \cdot \frac{1}{1+10^{\left(\frac{L_{noise}-L_{signal}}{10}\right)}} \tag{3}$$

3. RESULTS

In collaboration with Rieter Automotive Systems (Winterthur, Switzerland), a measurement campaign of STI has been performed on a D-segment five-door vehicle in the standing car with engine off (neutral configuration) and on roller bench in different constant speed driving conditions. RPM was maintained constant, in order to get a relatively steady background noise. However not steady effects like modulations can be observed in the background noise at certain speeds.

The measurement system, fully explained in [3], 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 “listener” was in the driver position while the “speaker” was in the real seat exactly behind the driver.

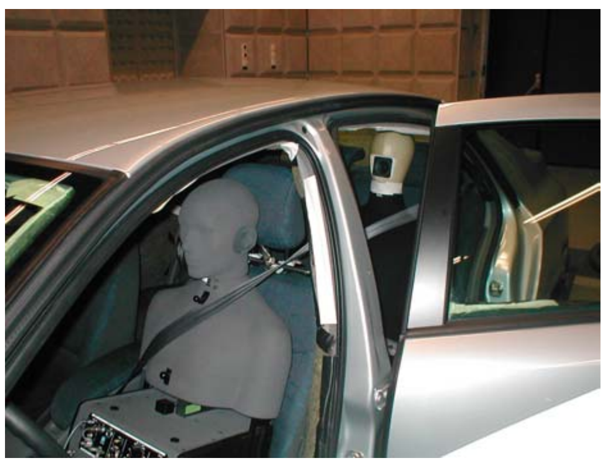

Figure 2. *Head and torso simulators used.*

In this configuration, impulse response (with MLS technique) and background noise have been measured at each speed and STI values deduced by postprocessing with two commercial softwares that employ “noisy IR” and a “noise free IR” technique developed at University Parma. The softwares used are MLSSA (DRA Laboratories) and Dirac (Acoustics Engineering), they both include the subroutine for computing the STI value according to IEC standard n.60268-16 [1].

In Figure 3 the comparison between the three methods is shown: it can be noticed that, with small background noise (condition found only when the car is in neutral), the results are exactly the same; when the signal-to-noise ratio is small, the results are different, but they all follow the same trend of going down when speed increases; finally, we have applied the techniques at a speed of 110 km/h without playing the signal by the mouth (no signal): incredibly only “noise free IR" gives a STI of 0.0, the others gives higher results because of they confuse the fluctuations of the background noise with fluctuations of the carrier (which is not present at all).

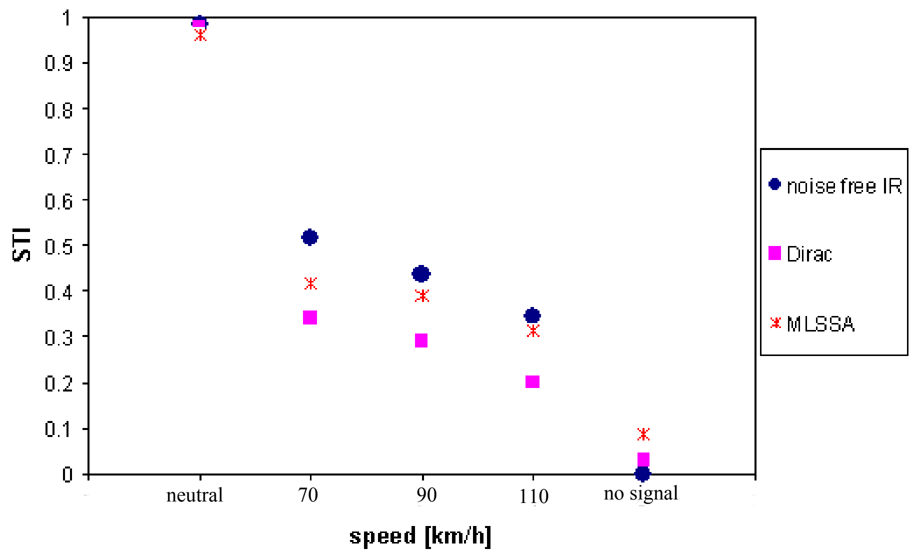

Figure 3. *Comparison between different techniques for computing STI.*

Later we have tested how large is the standard deviation using the tree different processing techniques: we have set the speed at 70 km/h and repeated in sequence the same measure, without moving anything.
It's distinctly noticeable that "noise free IR" gives the smallest deviation whereas "noisy IR" gives a standard deviation of about 0,03 that is about 10 % of the value of STI.
These different results are mainly connected with two aspects of "no noise IR techniques": first of all, these techniques, when the signal to noise ratio is not high enough, are not able to distinguish perfectly signal from noise. Secondarily because they use an MLS sequence (in this case a 16A at 48.000 Hz) of about 1.3 seconds that, probably, is not long enough compared with fluctuations of background noise inside vehicles.

Table 1: *Averages and standard deviations of STI computed with different techniques.*

Techniques	Average	Standard Deviation
Dirac	0,345	0,033
MLSSA	0,460	0,030
noise free IR	0,518	0,003

4. CONCLUSIONS

We have compared different methods for determining Speech Transmission Index inside car: in particular we have compared “noisy IR techniques” with “noise free IR techniques”. We have noticed that, because of the noise to signal ratio inside cars is quite low, “noisy IR techniques” give a bigger standard deviation and, in same cases, they tend to generate systematic errors: for example, when the signal is absent, the STI found is not zero but, at intermediate speeds, they tend to underestimate STI.

On the contrary “ Noise free IR techniques” give more precise results tanks to the use of a clean IR measurement, and averaging the background noise on long periods; moreover they are quicker and more practical to execute: they make it possible to measure the impulse response in the laboratory, and then to perform separately a car noise measurement under different driving conditions, including on-road measurements.

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

3. A. Farina, F. Bozzoli, “Measurement of the speech intelligibility inside cars”, 113th AES Convention, Los Angeles, 5-8 October 2002.