Audio Engineering Society

Convention Paper

Presented at the 117th Convention
2004 October 28–31 San Francisco, CA, USA

Measurement of Speech Transmission Index inside cars using throat-activated microphone and its correlation with drivers' impression

Fabio Bozzoli[1] and Angelo Farina[2]

[1] Industrial Eng. Dept., Università di Parma, Via delle Scienze - 43100 Parma – Italy
bozzoli@ied.eng.unipr.it

[2] Industrial Eng. Dept., Università di Parma, Via delle Scienze - 43100 Parma – Italy
farina@unipr.it

ABSTRACT

One of the most used intelligibility's parameters is the Speech Transmission Index: the techniques for determining it employs artificial speaker and listener. Inside cars, where signal to noise ratio is particularly low, the value of STI is mainly influenced by this ratio and determining the sound power of real speakers is the only way for piloting correctly the artificial mouth. We have implemented a technique that is based on throat-activated microphone and it is able to find the level of real speaker's voice inside noisy spaces in the effective conditions. Especially, we have studied the speech inside cars and we have discovered how the value defined by typical configuration may be extremely different from real one and, in this way, we have been able to produce more reliable excitation signals. Using this "raised" signal we have tested one car and we have tried to find a good correlation between drivers' impression and objective values.

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 the IEC standard n.60268-16:2003 [1], are based on the reduction of the modulation index of a test signal simulating the speech characteristic of a real talker, when emitted in an acoustic environment.

The test signal consists of a noise carrier with a speech-spread frequency spectrum and a sinusoidal intensity modulation at frequency F (see Figure 1), it is transmitted by a sound source situated at the talker's position to a binaural dummy head at any listener's position.

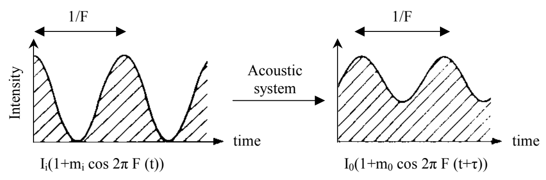

Figure 1 - Modulated signal emitted by the artificial mouth (left) and received at the listener position (right), showing a smaller modulation at the receiver.

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

$$m(F) = \frac{m_o}{m_i} \qquad (1)$$

The STI is derived from these modulation transfer functions, taking in account auditory masking, absolute hearing threshold, and 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.

To obtain the actual STI value, we have developed a method, fully explained in [2], based on measuring the Impulse Response in absence of background noise, making use of special techniques (for example MLS or Sweep signal) for maximizing signal to noise ratio.

Using the "noise free IR" technique obviously brings some practical advantages: it is possible to measure the impulse response in the laboratory the engine switched off, and then to perform separately a car noise measurement under different driving conditions, including on-road measurements.

Finally, we have developed a plug-in inside "Adobe Audition" to calculate the STI by means of this noise free technique.

The problem with the standard is that the level of the speech-like test signal is not specified. it is simply mentioned that : *"…for a determination representative of the signal to noise ratio , the mean intensity of the test signal should be equivalent to the normal speech level at the test position, i.e. the Leq of the test signal is adjusted to the typical Leq of on going speech in that position."*

The previous standards and many commercial systems suggest a reference value of 68 dB(A) at 1 meter from the lips of the artificial mouth. Using this level of the test signal we have found really low values of STI inside cars.

These results are badly correlated with the subjective experience of the driver, who doesn't find so hard to listen while the car is moving.

We explained this fact with the hypothesis that the real emission level of a talker inside a car is much higher than 68 dB(A) at 1m. Of consequence, we decided to study the real level of the speech in the operative conditions and to use this as the proper level of the test signal for determining a new "raised-voice" value of STI.

2. THE MEASUREMENTS INSIDE CAR

We have implemented a technique based on a throat-activated microphone, that is able to determine speech level at 1 m from the mouths in every kind of situation. It is useful because we can use spontaneous speech in the real environments without problems of reverberation or background noise. This technique has been deeply explained in [3].

In this study we have considered at the same time the Average Level (Leq) of on going speech and Active Speech Level (ASL) because it seams more connected to "running" speech as it is requested in the norm and overall it is less influenced by the fluency of the speaker.

Active Speech Level is defined by the ITU – T Recommendation P.56 [3]: it is measured by integrating a quantity proportional to instantaneous power over the aggregate of time during which the actual speech is present (called the active time), and then expressing the quotient, proportional to total energy divided by active time, in decibels relative to appropriate reference. Ideally, the criterion should indicate the presence of the speech for the same proportion of time as it appears to be present to a human listener, including those brief period of low or zero power that are not perceived as interruptions in the flow of speech.

Here we show the result of calibration procedure.

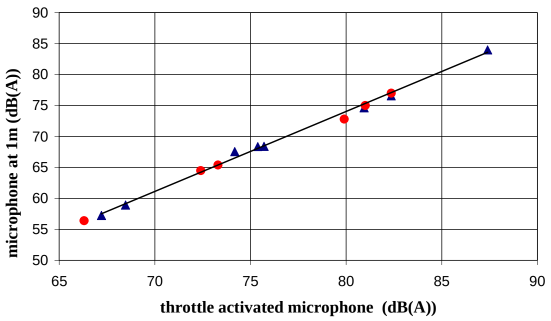

Figure 2 - Active Speech Level of the microphone in function of Active Speech Level of throttle-activated microphone is marked with triangles while Average Level of the microphone in function of Average Level of throttle-activated microphone is marked with circles.

It's clear from figure 2 that there is a perfectly linear dependence and that this chain of measurement has an error of about ±1 dB(A) .

The correlation doesn't change if we consider Active Speech Level or Average Level.

The next step consisted in testing a car at different speeds, for finding how much the speech level varies, and checking if the values of STI, determined with the raised signal, is able to better describe the car compartment.

Employing the throat microphone, we have measured the speech level inside a D-segment three-door vehicle car at different speed. The same subject already employed for the anechoic tests was employed, and he was asked to speak "normally" with the driver, while being seated behind him. We have done more than one test for each speed for checking also the dispersion of the results.

Here we report the values obtained (table 1).

Speed (km/h)	Background noise (dB(A))	Speech Leq (dB(A))	Speech A.S.L. (dB(A))
90 - a	69,8	73.0	73.1
110 - a	72.6	75.2	75.6
110 - b	73,0	75.1	75.5
130 - a	76.1	75.4	75.5
130 - b	76,7	74.2	74.5
130 - c	76,2	72.1	72.7

Table 1 - Background noise, Averege Level of the speech and Active Speech Level at different speeds.

We can see that, in this case, using Active Speech Level or Average Level doesn't change too much if we only consider "running speech". Active Spech Level is usually higher of not more than 0.4 dB.

From the results, these two behaviors are quite clear:

- loudness of speaker's voice can change of 3.0 dB cam by case with the same background noise, it depends on what he is saying ad how much he is emotionally involved in the conversation;

- increasing background noise doesn't mean necessary increasing speech level; i.e. at 130 km/h with a background noise of 76,2 dB(A) Active Speech Level is 72,7 (case c) while at 90 km/h with a background noise of 69,8 dB(A) we have measured an Active Speech Level of 73,1.

At the end, for this range of speed, seems reasonable taking an average of 74.0 dB(A) for the test signal.

Finally we have measured STI inside the same car using two different signal levels: 68 dB(A) , as traditionally is used, and 74 dB(A) as found in this study. The listener was in the driving position and the speaker in the rear seat exactly behind the driver (Figure 2).

Speed (km/h)	STI - 68 dB(A)	STI - 74 dB(A)
90	0,52	0,69
110	0,44	0,61
130	0,35	0,52

Table 2 - STI inside the car determined using two different signal levels: 68 dB(A), as traditionally is used, and 74 dB(A) as found in this study

Figure 2 - Head and torso simulator used.

3. SUBJECTIVE TEST

In the automotive industry the problem of intelligibility raised up because some customers complained of not been able to speak easily in some models of car.

In a car is not requested a perfect conversation, it is not a class room, but the driver has to be able to listen to what back passengers say without turning the head.

For this reason we have asked to five different drivers, having a conversation with the same speaker, to give a rank to the intelligibility at each speed. The possible choices were only three : bad, poor, good.

	Rank		
Speed (km/h)	Bad	Poor	Good
90			five people
110			five people
130		four people	one person

Table 3 - drivers' impression at each speed.

Our analysis of data is relatively straightforward. The subject selection is converted to a number. The scale is labeled with linear numerical values: we have associated 0 with "bad", 5 with "poor" and 10 with "good". The ratings can be used to calculate an average: in Figure 3 average rating is plotted in function of STI determined using two different signal levels.

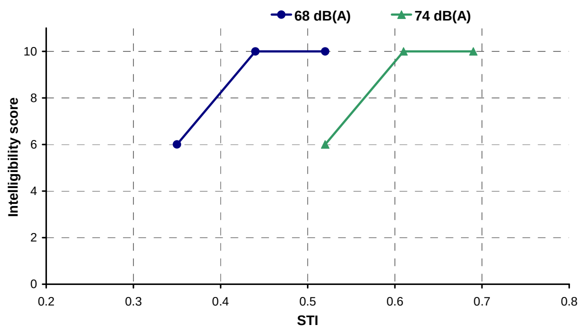

Figure 3 - average rating plotted in function of STI determined using two different signal levels.

4. RESULTS

IEC standard [1] gives a qualification of STI and it also reports relation with some subjective intelligibility measurements (Figure 4).

The nonsense word score for equally balanced CVC is obtained from [5]. The relation with PB words of the so-called "Harvard list" is according to [6].The relation with sentence intelligibility is based on Speech Reception Threshold results.

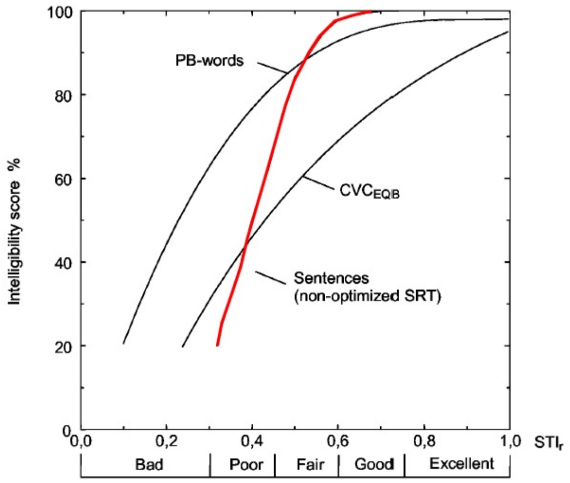

Figure 4 - average rating plotted in function of STI determined using two different signal levels.

This preliminary subjective test shows that also small changes of STI may generate problems of intelligibility.

It can be explained connecting this result with the sentences intelligibility: as it is shown in figure it has a fast fall under a STI of 0.6. According to this, intelligibility inside car compartment is good when STI is over 0,6 (intelligibility score over 90 per cent), it is bad when STI is under 0,4 (intelligibility score under 50 per cent) and it is poor when STI is between 0,4 and 0,6.

If we compare this qualification of STI with the values reported in Figure 3 , we see a good correlation with the value obtained using a s signal "raised" and a bad one using traditional signal.

5. CONCLUSION AND FURTHER WORK

We have improved a technique based on a throat activated microphone, that is able to determine active speech level at 1 m from the mouths in every kind of situation.

We have tested a car at different speeds, for finding how much the speech level varies. We found, for a speed between 90 km/h and 130 km/h, a speech level of 74 dB(A) with negligible differences between Active Speech Level and Average Level.

Using a source with this strength we have measured STI inside car and we have also collected drivers' judgment at various speeds. At the end we have found a good correlation between drivers' judgment and Speech Transmission Index values calculated with a "raised" source. We have been able to find that 0.6 is the value under witch conversation results difficult inside car.

The next step will be testing more cars and more speakers, for finding how much the speech level varies, and checking if the values of STI, determined with a source with these new levels, is able to better describe the car compartment, considering also a larger number of subjective tests.

6. ACKNOWLEDGEMENTS

This work was supported by Rieter Automotive (Winterthur, Switzerland). We also thank Mr Andrea Capra for his "speaking" and Mr Paolo Bilzi for his support .

7. REFERENCES

[1] CEI-IEC standard n. 60268-16, "Sound System Equipments – Objective rating of speech intelligibility by speech transmission index", July 2003.

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

[3] F.Bozzoli, A. Farina, "Measurement of Active Speech Level inside cars using throat-activated microphone", Pre-prints of the 116th AES Convention, Berlin, 8-11 May 2004.

[4] ITU-T Recommendation P56, "Telephone Transmission Quality – Objective Measuring Apparatus", March 1996.

[5] Steeneken H.J.M. and Houtgast T., "Validation of STIr method with the revised model", Speech Communication 38, 2002, p.413-425.

[6] Anderson B.W. and Kalb J.T.,"English verification of the STI method for estimating speech intelligibility of communication channel", J. Acoust. Soc. Am. 81,1987.