Implementation of a new metric for assessing and optimising speech intelligibility inside cars

M. Viktorovitch, Rieter Automotive AG

F. Bozzoli and A. Farina, University of Parma

Introduction

Obtaining optimal speech intelligibility in the interior of a vehicle is of paramount importance for car manufacturers, as it is one of the actual features contributing to the acoustical comfort of the vehicle.

The current descriptor used in the automotive industry for assessing the intelligibility is the Articulation Index [1] (AI). Its evaluation is based on the knowledge of the background noise solely, by taking into account its disturbing effect on the speech signal in function of the frequency range. The main drawback of the AI is that it doesn't account for the influence of the enclosure on the actual transmission of the speech signal, such as absorption and temporal distortions (echoes…).

In the early 80ies, Hootgast and Steeneken [2,3] elaborated a new metric named the Speech Transmission Index (STI). The goal of the STI is to characterize the loss of intelligibility during the transmission of a speech signal between a talker and a listener, within a given enclosure or an acoustical channel. The STI aims at gathering in one number the effects on the speech transmissibility of both the frequency distortions (e.g. masking of the speech signal by background noise, absorption of speech signal…) and the temporal distortions (e.g. reverberation effects…).

An STI measurement requires the use of an artificial talker (so-called the artificial mouth), emitting a speech like equalized signal and a receiver (generally a binaural head).

Thus, contrary to the AI, the STI takes into account the actual speech signal and its modifications due to the acoustic properties of the transmission chain.

The initial research results obtained by Hootgast and Steneken were rapidly transferred to the industry, principally in the telecommunication sector, for assessing the quality of equipment such as telephones. This activity lead to the editing of a European Standard [4], which exhaustively describes the measurement process.

Despite its success in the telecommunication industry, this metric has never been implemented in the automotive world. Rieter Automotive AG, in co-operation with the University of Parma has worked for several years on its adaptation for automotive applications [5,6,7].

In this context, this paper aims at presenting the latest developments, in terms of experimental procedures and equipment for a sustainable implementation of the STI in the automotive field. The first part of the article deals with the presentation of the fundamentals of STI. Then, a second section reports an important experimental campaign carried out for collecting data about the 3D directivity of the emission of speech.

This information not available to date, is of major importance for the prototyping of an artificial mouth with properties similar to those of humans..

Finally, two experimental campaigns dealing with vehicle interior noise measurement and aiming at evaluating the STI and the AI in different configurations are reported. These campaigns are described and analysed in order to highlight the similarities, the differences as well as the potential strengths of the two metrics.

1 Fundamentals of STI

1.1 Generalities

The STI measurement protocol is standardised in the framework of a European Standard (Sound Equipment System – Part 16: Objective rating of speech intelligibility by speech transmission index, EN 60268-16:2003). The STI is a physical quantity used to predict the absolute rating of the speech transmission quality with respect to intelligibility, of a listening space under specific speech conditions. The fundamental formulation elaborated by Hootgast and Steeneken [2,3] is summarized in this section.

The notion of modulation transfer function (MTF)

For any speaker-listener pair, the enclosure in which they are located is an acoustic transmission system. One of the characteristics of such a system is its ability to transfer the signal with a certain level of faithfulness. Speech may be regarded as a flow of sound with a spectrum varying continuously over time. A faithful transmission means that these spectrum variations are preserved.

In principle, the effects of the enclosure on the transfer of the speech signal are of two kinds:

1. The transfer may take place via different transmission paths, the direct path and the numerous indirect ones in which the sound is once or more times reflected within the enclosure, constituting the reverberation.
2. The transmission from one point to another in the enclosure is sensitive to acoustic interference by ambient noise penetrating the enclosure or generated within.

Since neither of these two disturbances transfer energy from one frequency to another, it is possible to replace the generally formulated requirement for faithful transmission by a more specific one: the temporal variations in the intensity, or temporal envelope, at any audio frequency should be preserved as well as possible. Preservation of the temporal envelope implies preservation of the individual sinusoidal components of which this envelope may be considered to be composed.

This reasoning strongly suggests that the acoustical merits of an enclosure with respect to speech intelligibility are determined by the extent to which sinusoidal intensity modulations produced at the speaker's location are still present at the listener's position. As a function of modulation frequency, this is the Modulation Transfer Function (MTF). For an input signal with a varying intensity $I_i\left(1+\cos 2\pi Ft\right)$, with F modulation frequency, the output signal is $I_o\left(1+m\cos 2\pi F\left(t-\tau\right)\right)$, with *m* the modulation index and τ the time lag due to transmission. The function *m(F)* is defined as the MTF (see Figure 1).

The MTF is only of interest for the range of F relevant for speech. Thus, one calculates these MTF for 14 modulation frequencies at one-third octave intervals ranging from 0.63 Hz to 12.5 Hz.

In general conditions, the disturbances are not constant on the whole frequency range (non constant S/N ratio, and reverberation time). Hence, it is of common rule to evaluate the *m(F)* function on octave bands. Seven octave bands were defined ranging from 125 Hz up to 8 kHz. The notation becomes m_{k,F_i}, *k* accounting for the k^th^ octave band and F_i the modulation frequency third-octave band.

The STI method is finally based on the determination of the modulation transfer function m_{k,F_i} for 98 data points (14 modulation frequencies for 7 octave bands).

From the MTF to the Speech Transmission Index

The aim is now to convert these 98 MTF values into one single index, which will be the STI-value. The most important step in converting the MTF into STI is its interpretation in terms of an apparent S/N ratio. In the actual case of only continuous interfering noise (no reverberation), m_k is simply determined by the S/N ratio and is independent of the modulation frequencies:

$$m_k = \frac{1}{1+10^{-\frac{1}{10}\left(\frac{S}{N}\right)_k}}$$

In the general case, and by extension, each individual $m_k(F)$ value is interpreted as follows:

$$SNR_{k,F_i,App} = 10\log\left(\frac{m_{k,F_i}}{1-m_{k,F_i}}\right)$$

where $SNR_{k,F_i,App}$ denotes the apparent signal to noise ratio, for the octave band *k* and the third-octave modulation frequency band F_i.

The next step in the STI evaluation procedure is to account for the effect of the auditory masking from a lower frequency band upon a higher frequency band, occurring in the hearing organ. This correction affects the values of the 98 $m_{k,Fi}$ and hence, the apparent signal to noise ratios.

Then, the $SNR_{k,F_{i},App}$ is normalized to vary in a range comprised between 0 and 1. The 98 resulting values are called the transmission index ($TI_{k.Fi}$).

The contribution of each octave band is then summarised in one index, the Modulation Transfer Index (MTI_k), obtained by considering the average of the $TI_{k.Fi}$ with respect to the modulation frequency:

$$MTI_k = \frac{1}{14}\sum_{i=1}^{14} TI_{k,F_i}$$

Finally, the speech transmission index is obtained as a weighted combination of the MTI's for the seven octave bands:

$$STI = \sum_{k=1}^{7}(\alpha_k .MTI_k) - \sum_{k=1}^{6}\beta_k \sqrt{MTI_k .MTI_{k+1}}$$

The factors α_k and β_k are defined by the European standard; they were derived from actual intelligibility tests.

Octave band Hz		125	250	500	1000	2000	4000	8000
Male	α	0.085	0.127	0.23	0.233	0.309	0.224	0.173
	β	0.085	0.078	0.065	0.011	0.047	0.095	-

Table 1 – coefficients used for male STI evaluation

1.2 Equipment and different Techniques available for measuring the MTFs

The experimental set-up required for the measurement

To calculate the MTFs, two equipments are required: a sound source, and a receiver.

The sound source:

To be as close as possible to the reality, the sound source is an artificial torso (so-called the artificial mouth) shaped like the upper part of a human body, with a loudspeaker positioned at the location of the mouth. The aim is to obtain a sound directivity diagram matching with the actual directivity of a human being.

The receiver:

The receiver is a binaural head, in order to measure signals close to those that would hit the eardrums of a real listener.

The techniques of measurement

The methods for determining the MTFs can be divided into two distinct groups: the methods using test signals sinusoidally modulated in intensity, and the methods based on the measurement of the impulse response of the system.

Group 1: The methods using sinusoidal test signals

These methods follow exactly the definition of MTF, using a test signal with a modulation index of one, at each of the 7 frequencies of the octave-band-filtered noise carrier, and at each of the 14 modulation frequencies.

They are used seldom because of long time needed for the 98 measurements and because of problems generated by unsteady background noise.

Group 2: The Noise Free Impulse Response method

The MTF can be derived from the knowledge of the squared impulse response of the enclosure, and the signal to noise ratio at the listener's position, using the formula:

$$m(F) = \frac{\int_0^\infty h_f^2(\tau).\exp(-j2\pi Ft)d\tau}{\int_0^\infty h_f^2(\tau)d\tau.} \frac{1}{1+10^{\left(\frac{L_{noise}-L_{signal}}{10}\right)}} \;,$$

where $h_f(\tau)$ is the impulse response at carrier frequency f, obtained in absence of background noise. L_{noise} and L_{signal} denote respectively the background noise and the signal at the listener's position.

For obtaining L_{signal}, the artificial mouth emits a signal which is a Maximum Length Sequence (MLS). This signal is octave filtered and amplitude calibrated in order to adhere strictly to the normalized spectrum of the human speech (according to EN 60268-16:2003).

In parallel, the impulse response $h_i(\tau)$ is derived from the equalized and amplitude calibrated MLS response mentioned above. This is possible due to the intrinsic characteristics of the MLS (its circular autocorrelation sequence is an impulse).

2 Validating a commercial and an internal artificial mouth for the measurement of the STI in the framework of vehicle applications

For carrying out the STI measurement, an artificial mouth generating a speech like equalized signal is required. The artificial mouth shall emit a signal whose directivity is close to the one of an average human. In the telecommunication sector, a standard defining the directivity of the mouth was elaborated [8]. The specifications stressed in this standard concern principally the speech directivity in a cone located in front of the mouth.

For automotive applications, these specifications are not sufficient. The directivity of the mouth shall be similar in all directions to the one of an average human. For instance, when a STI measurement is performed with the artificial mouth located at the driver's position and the receiver on the seat behind the driver, the directivity of the mouth in the back is an important property.

To date, the literature does not count many studies dealing with the measurement of the sound field surrounding a human talker. Nevertheless, two different experimental campaigns carried out recently should be mentioned. One is due to Bozzoli et al. [9] who compared the directivity of one human with the Bruel & Kjaer mouth simulator. The directivity was evaluated on a circle centered on the human talker and located in the horizontal plane of the mouth. The comparisons between the artificial mouth and the test person highlighted mismatches up to five dBs, notably in specific areas such as the back of the talker. In this study, one can highlight two important weaknesses:

1. Measuring the directivity of only one human is not sufficient for obtaining information about the dispersion of directivity in the human population
2. Obtaining data only in the horizontal plane containing the mouth is not sufficient. The directivity shall be investigated on a 3D sphere surrounding the artificial mouth.

Another experimental campaign was carried out by Chu et al. [10]. This directivity study also aimed at comparing the directivity of the Bruel & Kjaer mouth simulator to real human beings. A total of 40 persons (20 males and 20 females) were recorded and the sound field was evaluated on a complete sphere centered on the subjects.

Comparing and analysing the results obtained by the two teams lead to the following remarks:

1. The results may only be compared in the horizontal plane since the study of Bozzoli was restricted to this area.
2. The directivity of the real human measured by Bozzoli differs with the average directivity evaluated by Chu.

3. On the other hand, the results obtained by the two teams for the Bruel and Kjaer mouth are comparable.
4. Chu showed that the directivity of the artificial mouth of Bruel and Kjaer was in good agreement with the average directivity measured for the test persons.
5. The directivity of the human talker measured by Bozzoli was sensitively differing from the directivity of the Bruel and Kjaer artificial mouth.

The analysis of these two studies leads to the conclusion that they cannot be compared, and hence validated. The three following activities were therefore decided:

1. An experimental campaign for evaluating the sound field distribution around real talkers should be performed internally, in order to validate the results of Chu, and to possess reliable data for future validations of artificial mouths.
2. The good agreement obtained by Chu et al., between the directivity graphs of the Bruel and Kjaer artificial mouth and the test persons should be confirmed thanks to the internal experimental campaign aiming at evaluating the average directivity of test persons.
3. An internally developed artificial mouth will be evaluated by measuring its directivity and comparing it to the data obtained from the experimental campaign performed on the test persons.

In the next section, the experimental campaign performed on test persons is first reported. Then, some results concerning the directivity of the prototyped artificial mouth are presented, and these are compared with the ones obtained by Chu with the Bruel and Kjaer artificial mouth.

2.1 Description of the experimental campaign for evaluating the speech directivity of humans

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 measurement positions

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

1. **Microphone 1** remains in the horizontal plane of the mouth in front of the speaker during all the measurements. This is the reference microphone.
2. 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.
3. Horizontally, the measurements are carried out every 15 degrees (by shifting the microphone stand).
4. A measurement is also performed with a microphone located vertically above the mouth and one meter away from the mouth (**microphone 5**).

The measurement protocol

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

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 1. This procedure enables to get rid of the unavoidable differences occurring in the spectra of each repetition of the speaker.

Results obtained from the experimental campaign

The results are given in Figures 5 to 7, for three different octave bands, 250 Hz, 1000 Hz and 4000 Hz. The x-axis corresponds to the successive positions of microphones 2, 3 and 4 around the speaker (expressed in degrees). The y-axis represents for each octave band, the difference between microphone 1 and microphones 2, 3 and 4 respectively. These values are homogeneous to a sound pressure level expressed in dB.

The figures contain:

1. Thick curves which represent respectively the average of the results, the average plus the standard deviation (s.d.) and the average minus the standard deviation
2. One thick curve which represents the average directivity obtained by Chu for the male speakers.

Directivity of humans: analysis of the results

The results are displayed in octave band levels versus a position around the speaker given in degrees. Each line represents the subtraction between the level measured by the reference microphone and the levels measured by microphones 2, 3 or 4.

The lower the result, the less speech is emitted towards the considered location. When the result is low, the expression low directivity is used. On the contrary, when the obtained value is close to 0 or higher, the expression high directivity is used. From the curves, one can deduce the following remarks:

1. The directivity is always higher for position 4 than for positions 2 and 3.
2. The lowest directivity is always observed at position 3.
3. When the frequency increases, the speech directivity focuses to the front of the speaker.
4. The results obtained for the different subjects are homogeneous, even though the deviation increases with the frequency. The deviations between the two curves “average + s.d.”, “average – s.d.” in function of frequency, vary between 2 dB at 250 Hz (which is very low) to a maximum of 5 dB at 4000 Hz.
5. The results obtained in the framework of this study are comparable to those obtained by Chu.

2.2 Validation of directivity properties of the internal designed artificial mouth

The prototyped artificial mouth consists in a torso with a head in which a loudspeaker was implemented. The head is filled with a viscous material to minimize vibrations and the torso is damped with glued heavy layer and stuffed with absorbing material (Figure 8).

Its directivity was measured and compared to:

1. The average directivity experimentally evaluated for human talkers,
2. The directivity of the Bruel and Kjaer mouth simulator evaluated by Chu.

<u>Directivity of artificial mouth: analysis of the results</u>

The directivity of the artificial mouth is compared with the average, average + s.d. and average – s.d. directivity curves of the human talkers and the results of Chu concerning the directivity of the Bruel and Kjaer artificial mouth. The comparisons are proposed for two octave bands, 500 Hz and 2000 Hz, in Figures 9 and 10.

1. These results show an overall good agreement in all directions between the directivity of the artificial mouth and the one obtained from the subjects.
2. Despite some small local differences, the overall directivities of the Bruel and Kjaer artificial mouth and the prototype are in good agreement.

The major conclusions of this experimental evaluation are:

1. The internal design and fabrication of an artificial mouth with relevant properties in terms of 3D directivity was achieved, and is available for internal use. This activity also lead to gaining know-how and better understanding on the functioning principles of this very specific equipment.
2. The response in terms of directivity of a commercial artificial mouth and currently available on the market, was validated thanks to data obtained from both internal and external experimental campaigns.

3 Application of the STI: reporting of two measurement campaigns on vehicle

Two experimental applications are proposed in this section. The first one aims at highlighting the potentiality of the STI to discriminate the speech transmission properties of a vehicle, in function of the successive positions of the talkers and the listeners in a vehicle. The second application presents a STI measurement campaign carried out on a vehicle successively equipped with five different headliners.

3.1 Comparison of STI and AI measurements for three different configurations of talker and listener within a vehicle

The STI is measured on a vehicle for three different positions of the binaural head and the artificial mouth (see Figures 11 and 12).

The driving conditions for which the STI is assessed were constant speed on dry smooth road, 50 kph, 80 kph and 120 kph.

comparison of the STI and the AI for the studied configurations

The STI values obtained for the different configurations are displayed in Tables 1 and 2. These results are compared to the Articulation Index (AI) calculated from the background noise measured with the binaural head and averaged over time. The AI was calculated for the binaural head placed on the left back seat and right front seat respectively.

Binaural head on left back seat			50 kph	80 kph	120 kph
STI	Mouth driver position	Left	80	66	43
		Right	84	71	47
	Mouth right back seat	Left	86	73	50
		Right	93	86	67
AI		Left	90	68	40
		Right	93	75	48

Table 2 – STI and AI values obtained for configuration 1 and 2 (see Figure 11)

Binaural head on passenger front seat			50 kph	80 kph	120 kph
STI	Mouth right back seat	Left	86	77	58
		Right	87	79	60
AI		Left	93	72	44

	Right	89	67	38

Table 3 - STI and AI values obtained for configuration 3 (see Figure 12)

<u>Description of the AI results:</u>

For the two positions of the binaural head it is observed that:

1. The AI values are comparable.
2. As expected, the AI is always lower for the ear close to the window (left ear when the binaural head is positioned on the left back passenger seat – Figure 11, and right ear when the binaural head is positioned on the front passenger right seat – Figure 12).
3. The AI decreases with the speed of the vehicle, due to the increase of the background noise.

<u>Description and analysis of the STI results when the binaural head is positioned on the passenger left back seat (Figure 11):</u>

1. The STI decreases when the speed increases.
2. The STI is significantly higher when the artificial mouth is positioned on the right back seat than when it is located at the driver’s position. Moreover, the differences in the STI values observed for the two positions of the mouth increase when the speed of the vehicle increases.
3. When the mouth is located at the driver’s position, the STI obtained for the left ear (which is the closest to the window) is lower than for the right ear. In this configuration, it was observed as expected, that the background noise recorded by the right ear is lower than the one recorded by the left ear, whatever the speed. On the other hand, the opposite trend was observed for the recorded speech signal. Nevertheless, for this configuration the differences measured between the left and right levels of the background noise have a predominant effect on the STI.
4. When the mouth is located on the right passenger back seat, the STI obtained for both ears is significantly higher than for the case when the mouth is located on the driver’s seat.
5. When the mouth is located on the right passenger back seat, the STI for the right ear is significantly higher, and as expected the differences between left and right are higher than for the case when the mouth is on the driver’s seat.

<u>Description and analysis of the STI results when the binaural head is positioned on the passenger right front seat (Figure 12):</u>

1. As already observed for the configurations described previously, the STI decreases when the speed rises.
2. Contrary to the AI, the ear with the highest STI is the one obtained for the right ear which is the closest to the window. This result, not straightforward at first sight is explained by the fact that even though the recorded background noise is higher for the right ear, the speech signal is sufficiently higher to be predominant for driving the behaviour of the STI.
3. The differences observed when comparing this configuration to the inverse one (binaural head positioned at the back and mouth in the front) are important (a difference of round 15 points on a scale of 100 are observed at 120 kph for both ears).

Comparison of STI and AI: final remarks

From this example, some noticeable differences of the behaviour of the AI and the STI in function of the measurement conditions (position of the mouth and binaural head, speed of the vehicle), may be pointed out:

1. Taking into account different positions of the talker and listener leads to important variations of the STI, while the AI is not much modified when accounting for different positions of the binaural head.
2. When observing the variations of the STI for the different measured configurations it appears that a relevant assessment of the speech intelligibility inside a vehicle must take into account both, the actual path of the speech between the talker and the listener, and the masking effect of the background noise.
3. The STI is able to reveal interesting phenomena such as a better speech intelligibility for the ear close to the window due to the beneficial reflections of the speech signal on the window panels, counter-balancing the increase of background noise emitted by the same windows.

3.2 Measurement of the STI on a vehicle successively equipped with different headliners

In this study, the aim was to evaluate the influence of the absorption of a headliner on the STI. For this purpose, a vehicle was successively equipped with five different headliners made of different materials. The study comprised three stages:

1. Measurement of the absorption characteristics of the different headliners in alfa-cabin,
2. Evaluation of the STI on roller bench for two driving conditions: third gear slow run-up and coast down engine off, and two talker / listener configurations (artificial mouth back right – binaural head front right and opposite),

Nota Bene: In the framework of this study, a new STI evaluation procedure was developed to automatically calculate the STI versus the RPM. This new procedure widens the range of utilisation of the STI metric to non-constant speed driving conditions.

3. Calculation of the AI for the same driving conditions.

Multi-layer description of the five headliners and absorption curves

Table 4 describes the composition of the different layers of the five headliners.

Sample	Headliner description	AFR meas. (Nsm-3)
S1	Non-woven decor, PES Felt, Non woven backing	300-400
S2	Non-woven decor, non-woven, PP+GF felt, non-woven	2000

S3	Knit with soft foam, thin PES-felt, thick GM-PUR-GM carrier, NF-web, impervious backing	Only backing impervious
S4	Knit with soft foam, impervious NW, GM-PUR-GM sandwich carrier, impervious backing	Impervious
S5	NW-decor, PES felt, NW-backing	1000-1300

Table 4 – Description of the five headliners

An alfa-cabin measurement for each headliner was carried out. The samples utilised for these measurements were all of the same size (0.5 m$^{2)}$. The curves are displayed in Figure 13.

These results lead to the following remarks:

1. The headliner S3 is more absorptive than the four other headliners up to 2000 Hz. Then its absorption capabilities decrease comparatively to the other samples.
2. The headliner S4 exhibits lower absorptive characteristics from 2000 Hz upwards.

The aim is now to verify if these two singular behaviours do affect the STI and/or the AI.

Evaluation of the STI for slow run-up and coast down engine switched-off

The STI was evaluated on a roller-bench for a slow run-up in third gear (Figure 14) and a coast-down engine off (Figure 15). For these two cases, the artificial mouth (talker) is positioned on the front right passenger seat and the binaural head is located on the back right passenger seat.

The results obtained for the two driving conditions highlight the same trends, which may be summarized as follows:

1. As expected, STI is inversely proportional to the speed
2. In the configuration mouth in the front and binaural head in the back, the STI obtained for the left ear is always higher than the STI obtained for the right ear (the right ear is the one close to the window). This result is not universal, especially in a configuration where the artificial mouth is in the back and the binaural head in the front, this trend maybe reversed (see previous application).
3. The calculated STI is very similar for all the headliners, except S3 for which the STI is always sensitively lower.

Evaluation of the AI for slow run-up and coast down engine switched-off

The AI was evaluated from the background noise recordings used for the STI calculations. The AI was calculated for respectively slow run-up and coast down engine switched-off. For the two driving conditions, the same behaviour was observed. Thus, only the results for the coast down engine off are reported in this paper, they are displayed on Figure 16.

Comparison of the AI and the STI in relation with the absorption properties of the headliners

First, the general trends predicted by AI and STI are coherent. Especially, AI as STI is always higher for the left ear than for the right ear, and naturally AI and STI decrease smoothly with the speed or RPM of the vehicle. Second, the overall results obtained for the different headliners are quite comparable.

On the other hand, a detailed analysis of the results proves that the higher absorption characteristics up to 2000 Hz of the headliner S3 is influencing the STI (inducing a decrease of the index), while this influence is not clearly noticeable when analyzing the AI curves. An explanation of this fact is that the strong absorption of this headliner does not affect the background noise spectrum but brings an increase of absorption on the speech signal.

In conclusion, it appears that at least for this application study, the STI is more sensitive than the AI to fine acoustic tuning of the headliner.

Conclusions

In this paper, were presented the different activities leaded by Rieter Automotive AG for transferring the utilisation of the Speech Transmission Index (STI), a speech intelligibility metric used to date in various industrial sectors such as room acoustics design and telecommunications, to automotive applications.

First, was reported a significant experimental campaign dealing with:

1. The statistical evaluation of the directivity of speech emission of an average human talker,
2. The evaluation of the speech directivity of two artificial mouths, the Bruel and Kjaer Head and torso simulator and an internally developed artificial mouth. The aim was to compare their directivities with the average directivity of human talkers.

The conclusions of this experimental study were:

1. The directivity of the response of the Bruel and Kjaer mouth simulator is in good agreement with the actual directivity of an average human talker,
2. The artificial mouth prototype developed internally and the commercial torso simulator have similar directivity diagrams.
3. Therefore, the prototype is currently used internally for STI measurements.
4. Furthermore, this work lead to gaining valuable know-how and data within the company in this very specific field dealing with speech emission.

In the last section of this paper, two applications of the STI measurement were presented. The results obtained and compared with the classical AI show that the STI brings a more refined description of the speech transmissibility properties of the vehicle. Especially, it was shown that the STI is a valuable mean of information for effectively evaluating separately the different speech transmission channels in the passenger cabin (left/right, back/front…). The second example specifically highlighted that the STI is a more sensitive metric than the AI to the modification of the acoustic properties of the passenger cabin trim, such as the headliner. Thus, for this application the STI turned out to be a more accurate descriptor than the AI for a fine tuning of the interior acoustics.

References

[1]	K.D. Kryter	Methods for the calculation and use of the Articulation Index. J. Acoust. Soc. Am., 34, 1962.Year
[2]	H.J.M. Steeneken and T. Hootgast	A physical method for measuring speech-transmission quality. J. Acoust. Soc. Am., pp. 318-326, 67(1), 1980.
[3]	T. Hootgast, H.J.M. Steeneken and R. Plomp	Predicting Speech Intelligibility in Rooms from the Modulation Transfer Functions. I. General Room Acoustics. Acustica, pp 60-72, 46, 1980.
[4]	European Standard – EN 60268-16:2003	Sound system equipment – Part 16: Objective rating of speech intelligibility by speech transmission index
[5]	F. Bozzoli, A. Farina, P. Strasser, C. Castagnetti and C. Scaffidi	Design of a Sound Quality Assessment Method for Automotive Interiors: Development of a Jury Test and Adaptation of the Metrics PR, STI to Automotive Sounds Rieter Automotive Conference 2003, Luzern, 4-6 June 2003
[6]	A. Farina, F. Bozzoli and P. Strasser	Comparative Study of Speech Intelligibility Inside Cars Euronoise 2003 Conference, Naples, 19-21 May 2003
[7]	F. Bozzoli and A. Farina	Measurement of Speech Transmission Index inside cars using throat-activated microphone and its correlation with drivers' impression 117th AES Convention, San Francisco, USA, 28-31 October 2004
[8]	International Telecommunication Union	ITU-T Recommendation P.51. Series P: Telephone Transmission Quality – Objective measuring apparatus, artificial mouth
[9]	F. Bozzoli and A. Farina	Directivity Balloons of Real and Artificial Mouth Simulators for Measurement of the Speech Transmission Index 115th AES Convention, New York, 10-13 October 2003
[10]	W.T. Chu and A.C.C. Warnock	Detailed Directivity of Sound Fields Around Human Talkers. National Research Council Canada, Institute for Research in Construction, Internal report available at http://irc.nrc-cnrc.gc.ca/ircpubs. IRC-RR-104, December 2002.

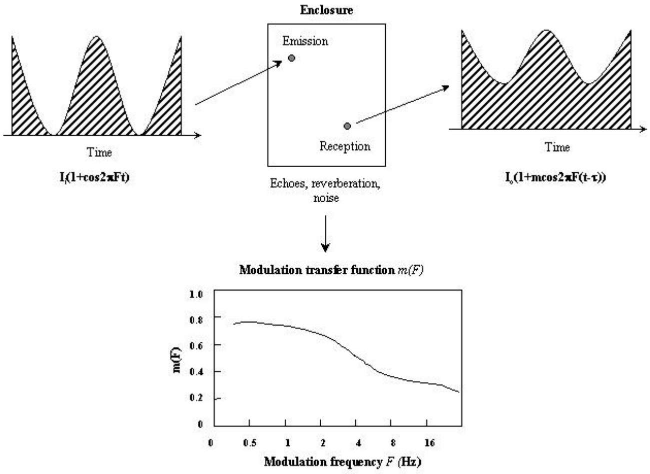


Figure 1– Definition of the modulation transfer function

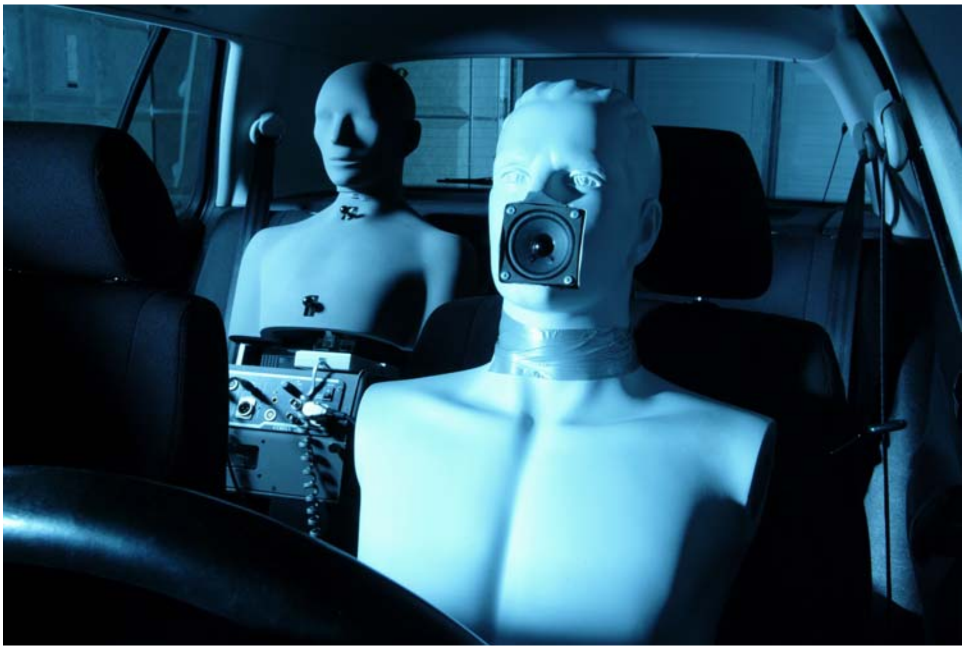


Figure 2 – Set-up for STI measurement. In foreground a Rieter artificial mouth prototype (the talker), in background, the artificial mouth (the receiver).

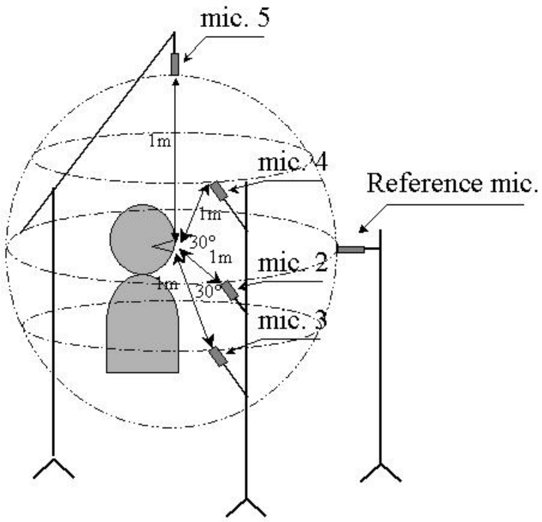


Figure 3 – Positioning of the microphones

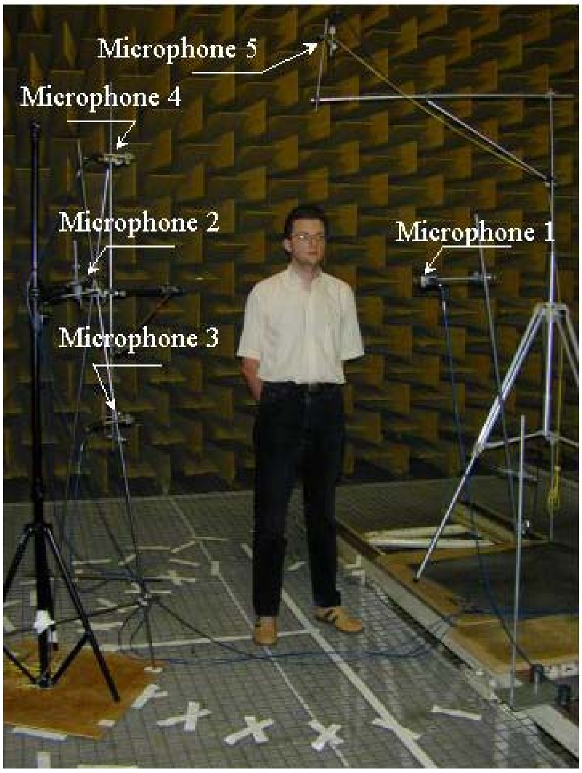


Figure 4 – Illustration of the set-up for the directivity measurements

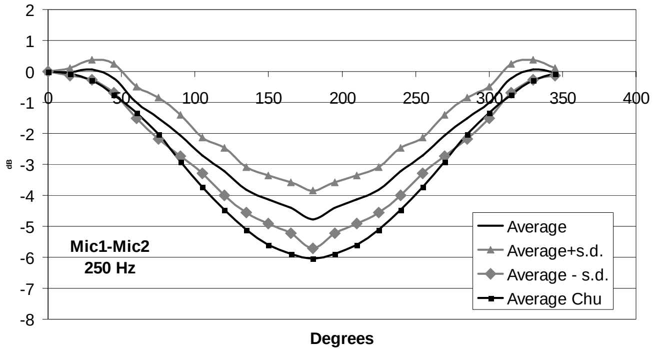
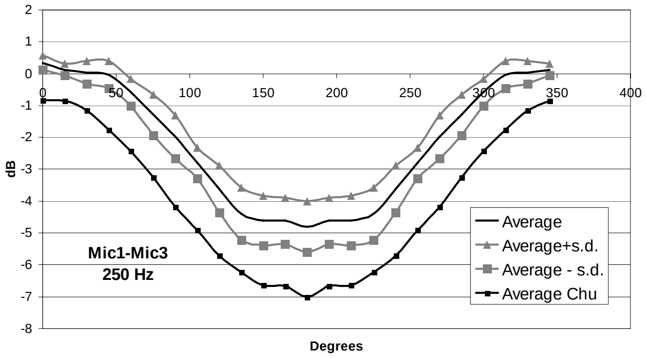


Figure 5a



Figre 5b

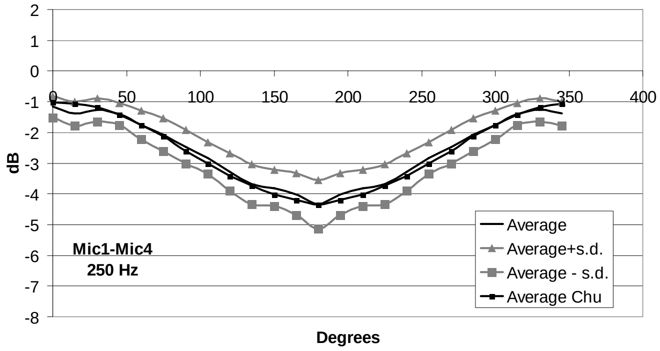


Figure 5c

Figure 5 – Directivity diagrams of the subjects for the octave band 250 Hz for microphones 2 (Figure 5a), 3 (Figure 5b) and 4 (Figure 5c)

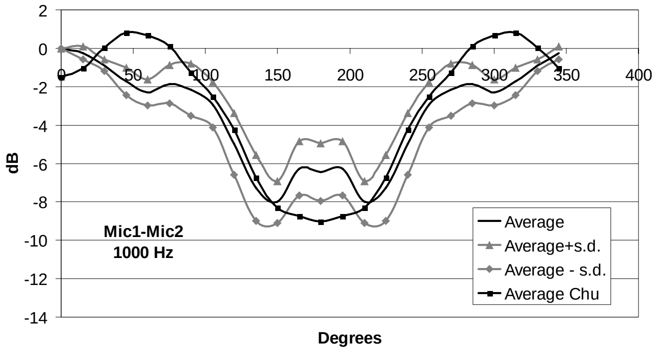


Figure 6a

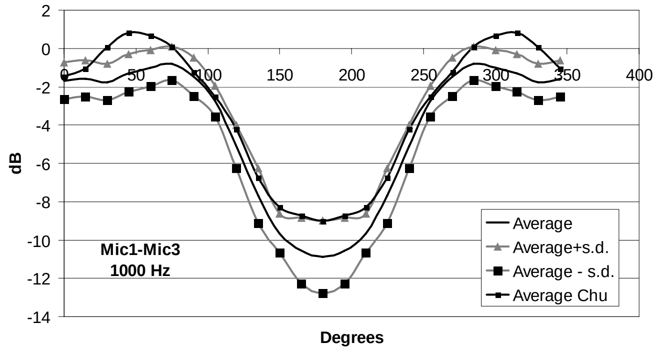


Figure 6b

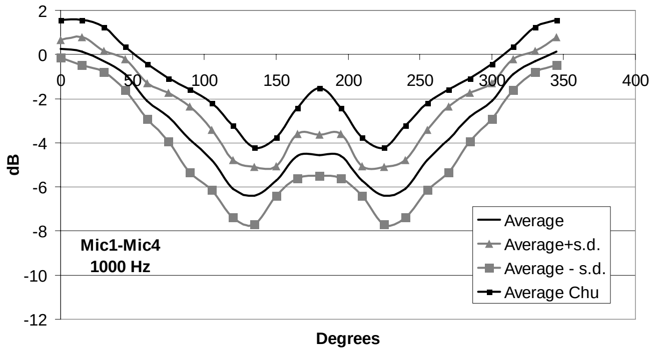


Figure 6c

Figure 2 - Directivity diagrams of the subjects for the octave band 1000 Hz for microphones 2 (Figure 6a), 3 (Figure 6b) and 4 (Figure 6c)

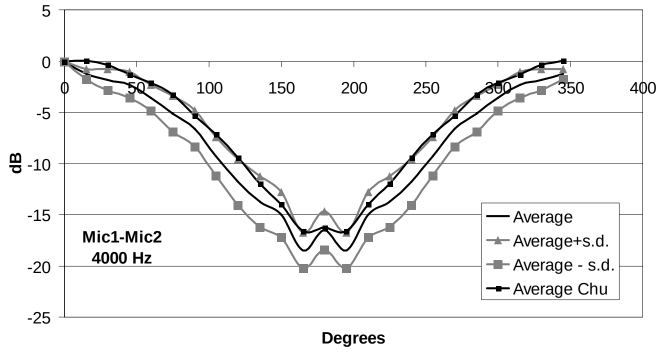


Figure 7a

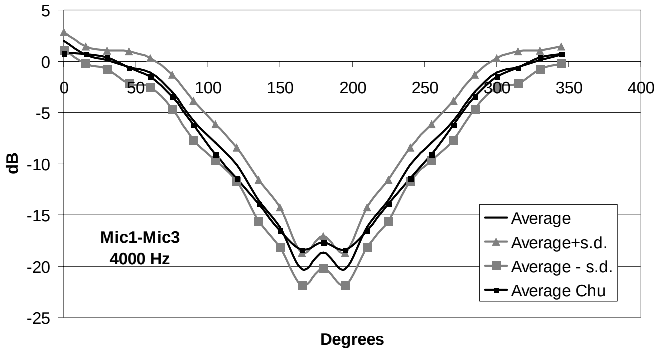


Figure 7b

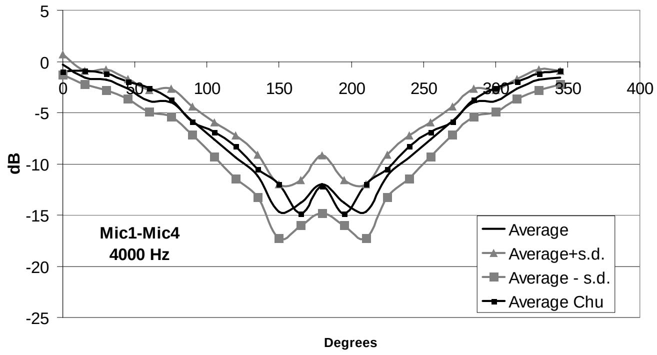


Figure 7c

Figure 3 - Directivity diagrams of the subjects for the octave band 4000 Hz for microphones 2 (Figure 7a), 3 (Figure 7b) and 4 (Figure 7c)

Figure 8 – The internal artificial mouth

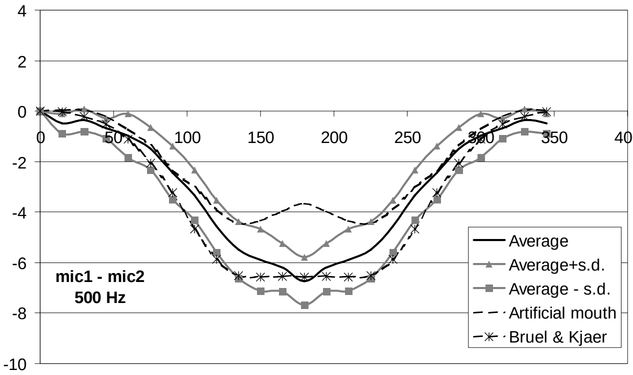


Figure 9a

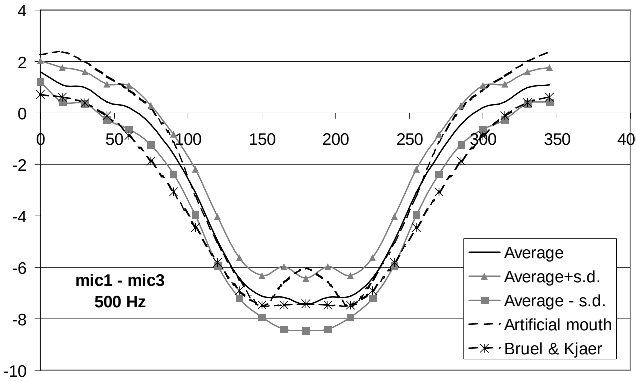


Figure 9b

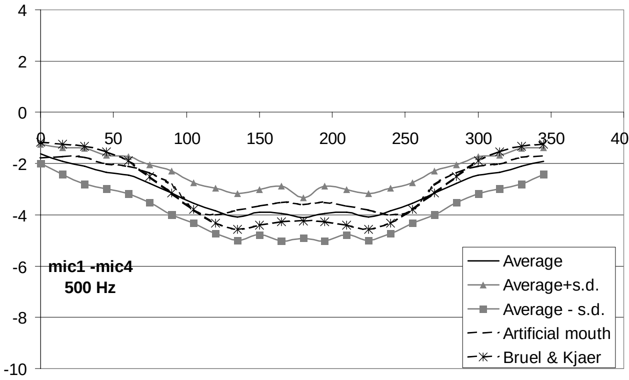


Figure 9c

Figure 9 a, b, c - Directivity diagrams of the prototyped artificial mouth and the Bruel and Kjaer mouth; comparison with average human directivity. Results for octave band 500 Hz and microphones 2, 3 and 4

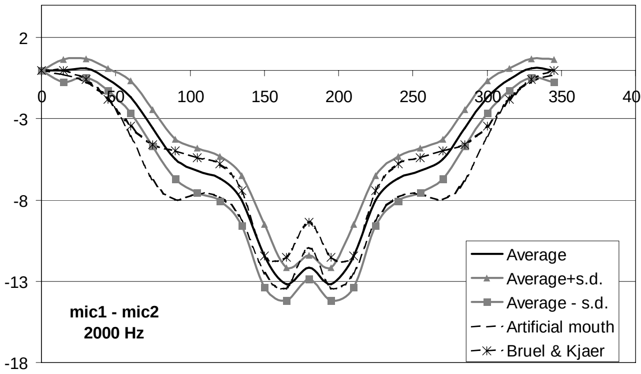


Figure 10a

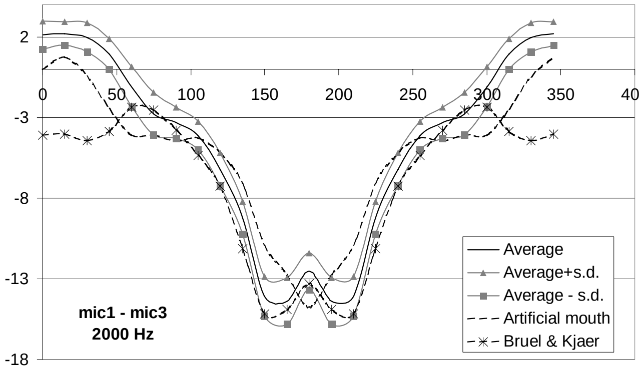


Figure 10b

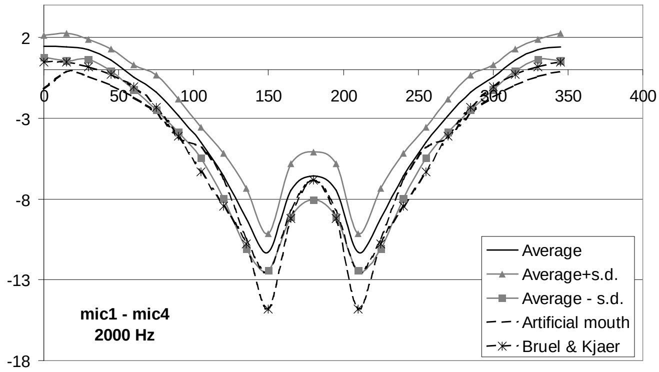


Figure 10c

Figure 4 a, b, c - Directivity diagrams of the prototyped artificial mouth and the Bruel and Kjaer mouth; comparison with average human directivity. Results for octave band 2000 Hz and microphones 2, 3 and 4

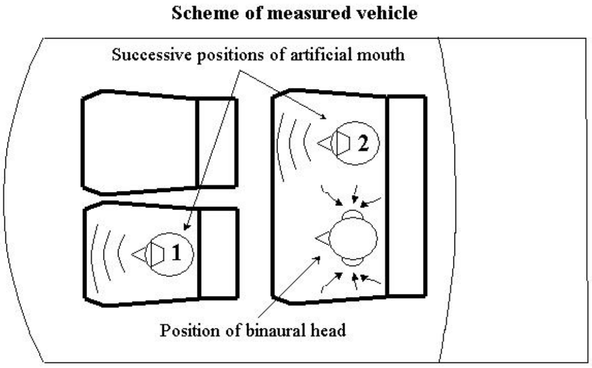


Figure 11 - Listener positioned on the back passenger left seat: talker successively positioned at the driver's location and the back passenger right seat.

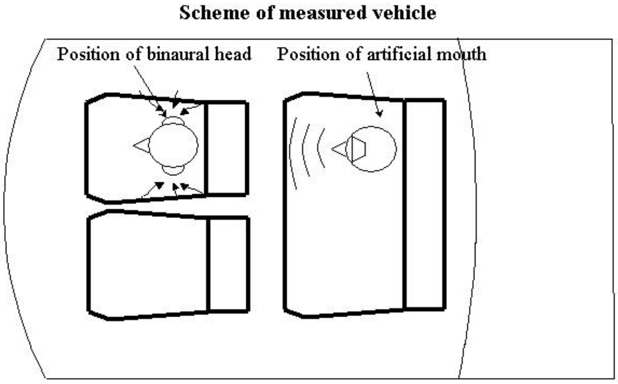


Figure 12 - Listener (binaural head) positioned on the right passenger seat and talker located behind the listener.

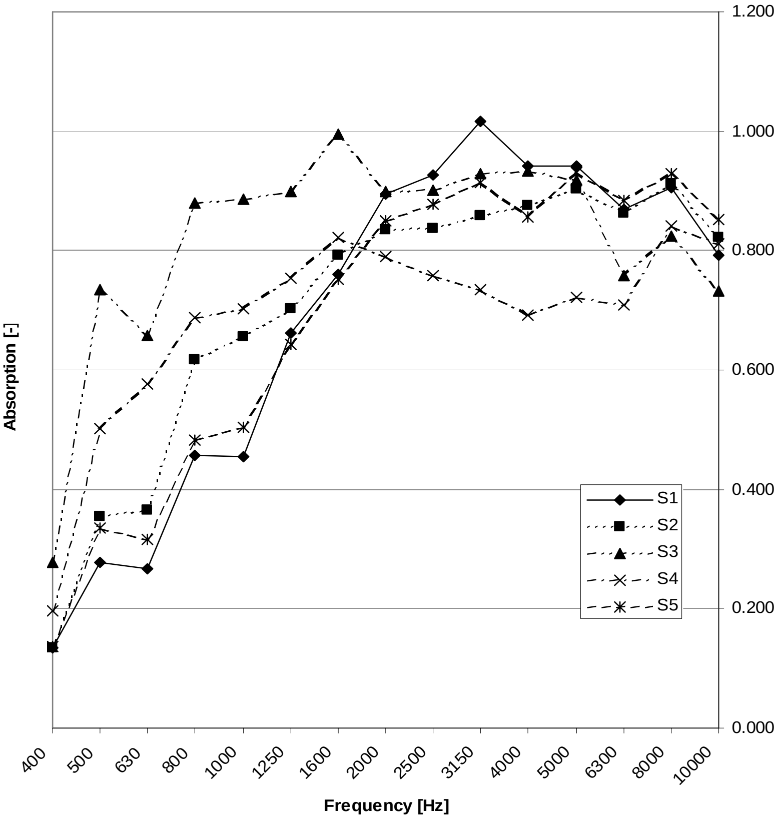


Figure 13 – Absorption curves obtained for the five different headliners

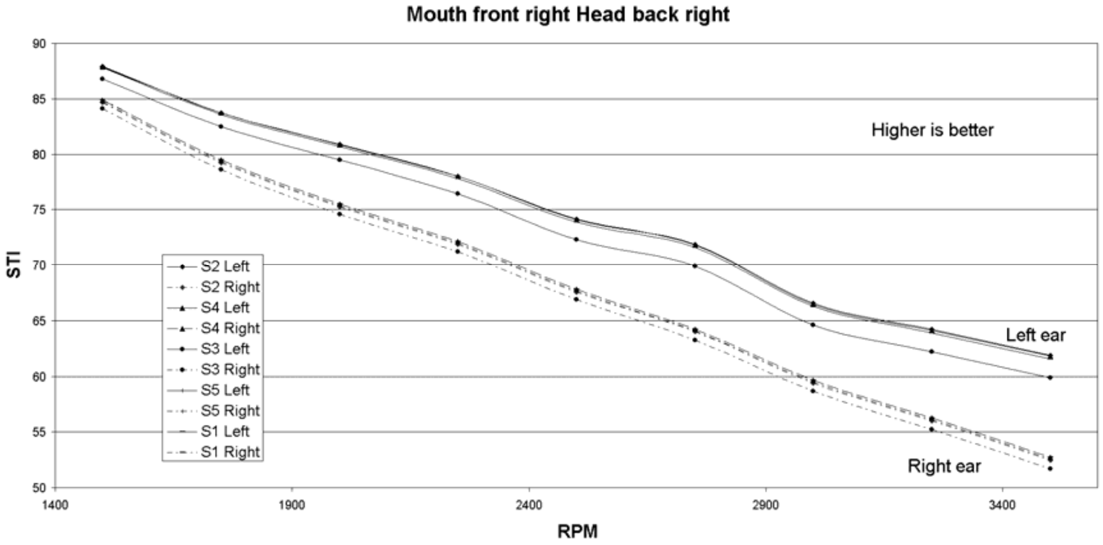


Figure 14 – Evaluation of STI for a third gear slow run-up. Results for left and right ears and for each headliner

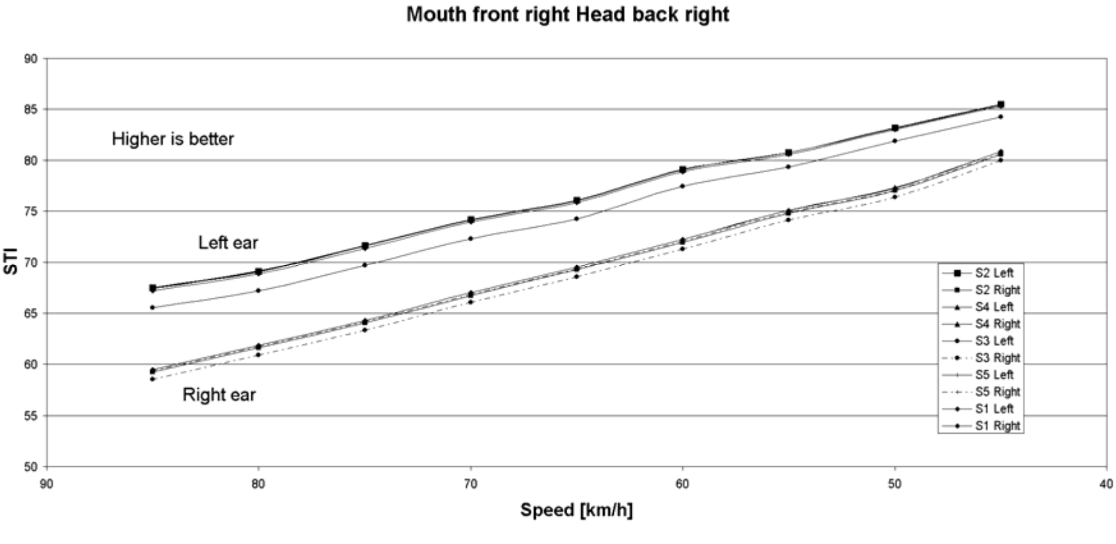


Figure 15 - Evaluation of STI for a coast-down engine-off. Results for left and right ears and for each headliner

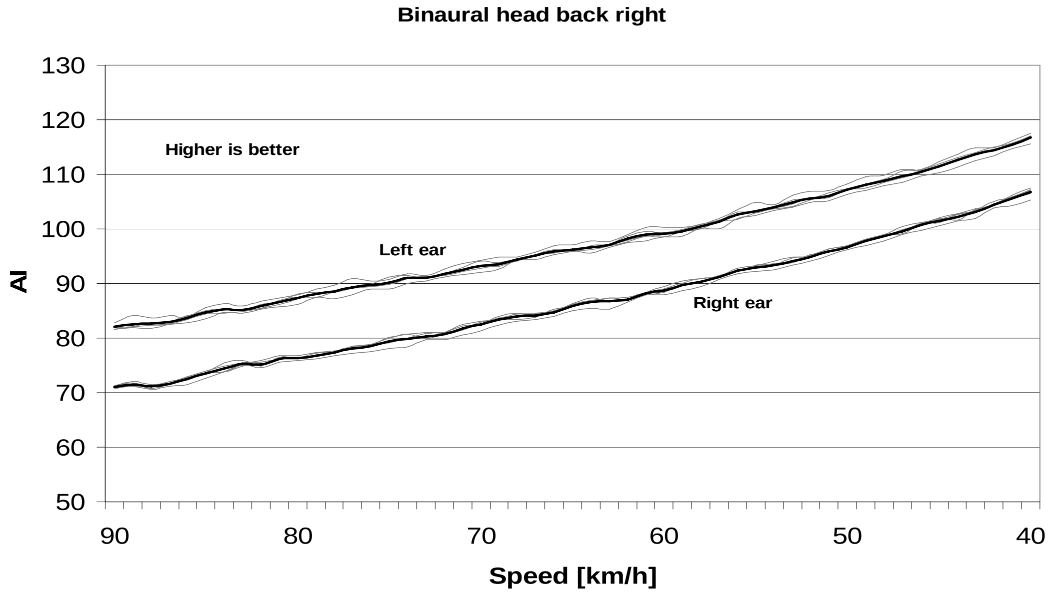


Figure 16 – Variation of the AI during a coast down engine off. For each ear, is represented the AI curve for each headliner (thin grey lines) and their average (black thick line).