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

F. Bozzoli, A. Farina, University of Parma; P. Strasser, C. Castagnetti, Rieter Automotive Switzerland; C. Scaffidi, Rieter Automotive Italy

Summary

Besides measurement of physical quantities like sound pressure level, the assessment of perceived sound quality inside a passenger compartment and its link to the interior sound package is an important issue.

A jury test focused specifically on the requirements of subjective sound package assessment and improvement is being designed at Rieter, in co-operation with the University of Parma.

First, the choice and implementation of an optimal binaural recording and replay chain is needed, in order to play back to the jurors' sound samples which reproduce in the most accurate way the spatiality, frequency content and natural feeling of automotive interior sound. This part of the work includes the comparative assessment of various binaural recording systems (different dummy heads or binaural microphones) and replay possibilities (by headphones or stereodipole). The implementation of a replay environment (sound quality lab), meeting the requirements of optimal jury test conditions, has to be foreseen. The design of a paired comparison jury test, based on these binaural recordings, and first applications of the methodology are described in this paper.

In parallel, the applicability to automotive sound of metrics used in other fields has been investigated. Thus the Prominence Ratio (PR), developed by the Computer Industry can be applied to tonal components of vehicle sound. This metric is measuring emergent sound peaks as compared to neighbouring frequencies. Thus the PR can have high values, denoting disturbance, independently of the dBA level. The PR concept has been extended to low frequencies, in order to make it applicable to automotive booming noise. First correlations with results from a jury test are shown.

Another interesting sound quality metric is the Speech Transmission Index (STI), mainly used in building acoustics, which takes into account both background noise and propagation of the acoustic signal. A feasibility study of its applicability during sound package improvement studies is on-going.

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

F. Bozzoli, A. Farina, University of Parma; P. Strasser, C. Castagnetti, Rieter Automotive Switzerland; C. Scaffidi, Rieter Automotive Italy

Introduction

A whole panel of objective measurements and simulations are currently used for acoustic optimisation. These methods are mainly aiming at minimising the sound pressure level at driver's ear and at the different passenger positions and show to be indispensable tools for sound package design. However, they are sometimes limited in giving a detailed picture of comfort or disturbances as perceived by the users of the vehicle. In co-operation with the University of Parma, Rieter is investigating jury test methods and a series of metrics suited to characterise the perceived annoyances or communication comfort.

Rieter's aim is to draw a link between vehicle characteristics - especially the sound package - and the "comfort feeling" of a representative end user population. Specifications for the "Sound Quality Assessment Method" are:

- precision and representativeness of people's perception
- affordable measurement / recording effort and implementation of jury tests
- as much as possible, analytical link to sound package characteristics

A variety of difficulties specific to automotive sound have to be overcome, such as for example non-acoustic cues contributing to the perception of comfort, the non-steady aspect of vehicle noise and rpm dependence, low signal to noise ratios for intelligibility as compared to room acoustics.

For future applicability of the technique in daily development work, the measurement effort should be kept reasonable, especially for on-road recordings.

Figure 1 shows the sound quality assessment method positioned in relation to the other experimental and numerical methodologies used for the assessment of vehicle noise.

1 Rieter's activities in the subjective acoustics field

1.1 Optimal use of binaural technology

Besides on-road subjective assessment by experts, which is the assessment closest to the real use of the vehicle, Rieter is using binaural recordings. This method offers in particular the possibility to compare recordings from different cars or different treatment configurations in a short time interval, thus making direct comparisons possible, despite the short memory of hearing.

Using binaural technology means trying to reproduce with a recording/replay system exactly the same sound pressure at the listener's eardrum, as the one in real conditions (listener sitting in the driving car). In order to reproduce as close as possible the recorded vehicle sound, an optimal recording and replay system (headphones, listening room) has to be used. In particular the spatial reproduction of the sound should be realistic. The spatial characteristics of 7 binaural recording systems (binaural dummy heads or binaural microphones placed on a dummy head) available either at Rieter or on the market, have been recorded and compared. The MLS technique was used to obtain the impulse responses of the binaural devices. Measurements were performed in a fully anechoic room. The dummy heads were placed on a rotating table, allowing one measurement every 5° in the horizontal plane. The loudspeaker was hung by an overhead crane at the azimuth angles 0° (in the ear plane), + 30°, -15°, -30°, -45°. The measurement set-up is shown in **Figure 2**. For all measurements, the distance between loudspeaker and ear-plane was kept constant at 2.5m. Since automotive noise sources (except aerodynamic noise) are mainly below the driver's or passenger's ears, a more detailed study was done for the lower half space (more loudspeaker positions).

Figure 3 shows the results of the spatial characterisation for 2 different recording systems. Plot a) shows the impulse responses over rotation angle of the table for a good recording device: the contribution to the left ear decreases smoothly while the right ear increases. For comparison, plot b) shows the characterisation of another recording device, which will lead to worse spatial accuracy of the recorded noise. Jury tests to correlate perceived and actual spatial localisation in the measurement are foreseen for different recording/replay chains.

Moreover, a set of inverse filters for different pairs of recording and replay systems has been determined through measurements and inverse filter computation in AURORA, as for example "recording with dummy head 1 and replay with headphones 1", "recording with dummy head 1 and replay with headphones 2", "recording with dummy head 2 and replay with headphones 1", "recording with microphones 3 and replay with headphones 2 " etc...This approach makes it possible to exchange recorded files easily and to listen to them in best conditions, independently on the locally available replay hardware.

1.2 Jury tests

Recordings are used by experts together with the results of objective measurements for diagnostics of perceived annoyance and assessment of the improvement brought by an optimised sound package. In addition to this evaluation, a designed jury test with either an expert or a naïve jury may give a statistically valid evaluation of the comfort perception of a given population.

Pair comparison tests are performed at Rieter Italy, especially for truck noise. Recordings are presented pairwise to the juror in different combination of pairs. For each pair, the juror is asked which sound sample is better according to a given criterion, like "acoustic comfort" or "sportiness" (in the case of passenger cars). The evaluation of the overall responses allows a ranking of the presented sound samples, together with an evaluation of the coherence of each juror's responses through triangular checks.

A link to sound package properties can be drawn by the introduction of experimentally modified sound package (recording 1: original sound package; recording 2: modified "prototype package" in the same driving conditions), by filtering the measurement result (band-stop for removing potentially disturbing sound) or superimposing simulation (more absorption). In a test such as semantic differential, the influence of the sound package could be analysed by an adapted wording.

Jury tests need effort and time. In order to assess more quickly some characteristics of the vehicle sound closely linked to human perception, some metrics, i.e. objective quantities that can be derived from measurements and related to perceived comfort are investigated.

1.3 Metrics to draw a link between subjective perception and measurements

Currently interior vehicle sound is characterised experimentally by 2 important objective quantities, the A-weighted spectrum and the Articulation Index (AI). Both quantities take into account some characteristics linked to human perception, such as the frequency dependent perception of loudness (A-weighting) or the disturbance to intelligibility of speech brought by the background noise level (AI).

In order to quantify more precisely specific sound characteristics linked to the subjective comfort perception inside a vehicle, two specific metrics, used in other fields, have been investigated and their applicability to the automotive sound package design evaluated:

- the Prominence Ratio PR, originally used in computer industry, can help targeting tonal components in the spectrum, whose disturbing effect is not due to high SPL, but to their prominence compared to neighbouring frequency bands.
- the Speech Transmission Index STI, used for telecommunication and in building acoustics should give a more detailed picture of the intelligibility inside a car than the commonly used Articulation Index. Whereas AI quantifies intelligibility only from the recording of the background noise (car in running conditions), the Speech Transmission Index considers both the transfer of the acoustic signal from the speaker to the listener and the background noise. Therefore, the influence of the trim characteristics of the passenger compartment (reflections, absorption) on both, the attenuation of noise and the transmission of speech, are taken into account. Moreover the position of the speaker compared to the listener is also considered.

Possible modifications or extensions as compared to the generic use of this metrics are proposed.

2 Prominence ratio

The quality of a car's interior acoustics can be influenced by the presence of discrete tones or narrow-band sounds in the spectrum. Even if the measured overall level and the averaged third octave spectrum is low for a car of a certain segment, this car can be subjectively judged as poor because of a particular tone that affects our perception and thus our judgement. A strong 2nd gear whine for example, might result to be annoying when driving in town where the stops at the traffic lights are often and the use of low gears is more intensive. When driving on the highway at a fixed speed, the contact between the road and the tires could generate strong tonal components in the spectrum. Other examples could be the noise emitted by ancillaries such as the injection pump, the fan etc. Tonal components catch the attention of the driver even if their energy is low. This phenomenon is explained by the nature of human's auditory system, which is particularly sensitive to pure tones and narrow-band sounds.

Prominence Ratio metric has proved to be a "reliable" tool to identify and quantify (to certain extent) the annoyance caused by tonal components. This metric allows a fast identification of possible annoyances inside the spectrum, which have a direct correlation with the subjective impression. As the auditory system is particularly sensitive to discrete tones, it is sometimes difficult to identify a problem by looking at a FFT spectrogram. In this kind of analysis, the more visible features in the spectrum are the ones at higher levels. Therefore it can occur that a tone with low level would be clearly audible but not visible. The PR metric is a *level independent* quantity that evaluates the *emergence* or the *prominence* of tones and narrow band noises. When discrete tones appear in a broadband noise, the signal is perceived as being more annoying than the broadband itself in absence of the tone. The annoyance is linked to the amount by which the tone "sticks out" above the noise or, in other words, it is linked to the amount of *prominence.*

In previous years, a standard was developed to measure objectively the prominence of a tone. This is described in the ANSI S12.10, ISO 7779-1999 and in the ECMA-74 -1997 standards. The method is based on the metric *tone-to-noise ratio.* The tone-to-noise ratio is defined as the ratio of the power in the tone to the power contained in the critical band centered on the tone. A tone would be called prominent if the tone-to-noise ratio has a value greater or equal to 6 dB. If the spectrum contains multiple tones close together in frequency, than this metric would fail to measure their prominence. For example, the electronics of computers and business equipment results in a noise emission spectrum rich in discrete tones. The need to overcome this problem, pushed Matthew Nobile (IBM Acoustics Laboratory) and Gordon Bienvenue (State University of New York) to develop the metric called *prominence ratio* [1]. The prominence ratio should be able to indicate the prominence of these multiple tones that would be judged as non-prominent if the metric tone-to-noise ratio were used. Furthermore, this objective metric should have a good correlation with the subjective response. Some psychoacoustical studies [2], [3], [4] show this correspondence. The prominence ratio quantifies the ratio of the total power of the critical band containing the tone, to the average power contained in the two surrounding bands, one below and one above the middle band. This calculation computes how much one *critical band* 'sticks out' compared to the neighboring ones.

2.1 Definition of prominence ratio PR

The prominence ratio sums the energy of the tones in the critical band, similar to what the ear does subjectively. The width Δf_c of the critical band centered at any frequency f_0, can be calculated from the following equation:

$$\Delta f_c = 25 + 75 \cdot \left[1 + 1.4 \cdot (f_0 / 1000)^2\right]^{0.69} Hz$$

where $f_0 = \sqrt{f_1 \cdot f_2}$ and $f_2 - f_1 = \Delta f_c$

f_0 = centre frequency, f_1 = lower band edge frequency and f_2 = upper band edge frequency.

The critical band noise power in the middle band W_M is defined as the total power (or mean-square value) contained in the critical band centered on the tone under investigation (f_0 is set equal to f_t = frequency of the tone). The critical band noise power in the lower band W_L is defined as the total power contained in the band immediately below the one centered on f_0. The critical band noise power in the upper band W_U is defined as the total power contained in the band immediately above and contiguous to the one centered on f_0. The prominence ratio is then calculated as follows:

$$\Delta LP = \left[\frac{WM}{1/2 \cdot (WL + WU)}\right] dB$$

It is important to point out that, according to the PR method, for the computation of the prominence of a tone, it is necessary to previously know where the tone is situated in its spectrum in order to center the critical band around it. A more useful approach would be to have a procedure where this step is not necessary. For example, if a certain car has a whistling problem, and the whistle tone is not clearly visible and identifiable by performing the FFT spectrogram analysis, then the PR method can not be performed as it is defined above. To solve this, the PR calculation can be performed in a different way. In the software ArtemiS, by Head Acoustics, the PR calculation is defined as follows.

$$R(n) = \frac{Eg(n)}{\left[Eg(n-1) + Eg(n+1)\right]/2}$$

where n = 1,2,3,........,22. $R(n)$ is the result of group n, $Eg(n)$ is the energy of group n. The energy of a group of frequencies $Eg(n)$, is calculated by adding the energy contained in that group. This is done trough an FFT calculation with 4096 samples (sample number $N_s = f_s * \tau$). Because the position of the tonal components in their spectrum is not known, the software ArtemiS starts the calculation by taking the frequency of 150 Hz as the first center frequency of the group number 2 and shifts then the center frequency in steps of 1/24 octave. So the first center frequency for the first group, is 150 Hz, the second center frequency is 156,25 Hz, the third is 162.76 Hz and so on. For each center frequency, the energy contained in a critical-band wide group centered around the center frequency is summed resulting in the energy $Eg(n)$ that can be compared with the energies of $Eg(n\text{-}1)$ and $Eg(n+1)$, which are the energies of the adjacent critical bands.

The prominence ratio metric can be applied as it is defined, in the frequency range between 150 and 10500 Hz. Outside of this range, the calculation is not possible. In the above formula the energy contained in the group $Eg(n)$ is compared to the energies $Eg(n\text{-}1)$ and $Eg(n+1)$, therefore the first critical band that can be used for the PR calculation is the band for which $n = 1$. The frequency range corresponding to the bandwidth is from 100 to 200 Hz (centre frequency is 150 Hz). The last band is the one numbered with n=22 and it ranges from 9500 to 12000 Hz (centre frequency 10500 Hz).

2.2 Application to tonal components in vehicle noise

The PR analysis is a powerful tool to identify annoyances due to emergent sounds. This metric when applied on binaural recordings of road measurements allows a fast detection of the noises that regardless of their level catch the attention of the driver or of the passengers. **Figures 4** and **5** show the comparison between a FFT spectrogram and a PR ratio analysis of the same recording, respectively in a case with gear whine noise and a case of diesel knock. In the PR analysis it is possible to identify immediately the frequency, the time when it occurs and the amount of emergence of single narrow-band sounds. Other techniques could be used, like performing various filtering on the original signal while listening at the recording and, at the same time, tracking this information along the FFT spectrogram. Still the spotting and the quantification of some features could be difficult and it would result in a time consuming process. In fact, in the mentioned pictures we can observe that the visible features in the PR colormap are few in comparisons with the amount present in the FFT colormap. Because of the definition of PR there is a direct link between the prominent sounds and our perception. This ensures time saving in the identification of disturbing noises. Furthermore, the level of disturbance is linked to the amount of prominence and therefore the PR value in dB can be used as a quality indicator.

2.3 Extension of PR to lower frequencies - Application to booming noise

The calculation of Prominence Ratio can also be performed with a software like Matlab using 1/nth octave analysis data acquired with any acquisition software. Because of the limitations in the frequency range that can be analysed (150-10500 Hz) the PR cannot be used to evaluate low frequency emergent phenomena such as booms. For this reason, a modified version of this metric was designed in Rieter to explore its potential in identifying and quantifying booms [5]. This modified version is defined as follows:

$$EPR = \left[\frac{10\wedge^{(lm/10)}}{(10\wedge^{(ll/10)} + 10\wedge^{(lu/10)})/2} \right] dB$$

where lm = level of the middle band, ll = level of the lower band and lu = level of the upper band.

The width of the bands can be fixed as any fraction of octave bands for an easier use with the common softwares (which usually include the 1/nth octave band analysis). In the case of booming noise, the width is set to 1/3 octave. The use of 1/3-ocatve bands was chosen because the booming noise doesn't appear as a pure tone. Finer band resolutions, like 1/12-ocatve or 1/24-ocatve, would lean to judge as non-prominent, noises that are prominent. This is because if we analyze a boom in its frequency spectrum, this appears as a 'broad' peak. By using the 1/24-octave band resolution, for example, it could happen that more than one band is contained in the peak. In other words when the level of two adjacent 1/12 octave bands for example, is high and similar, then the calculated prominence with the above expression would return a low value for each of the two bands.

The human hearing system is less sensitive at the lower frequencies in comparison with the higher ones. For this reason the booming noise becomes annoying when the low frequencies have high levels. In this case, a boom can be spotted also from a FFT spectrogram or by comparing the overall level and the main orders level using a B-weighting curve. Therefore the modified version of the prominence ratio can be used more as a quality indicator for booming noise than as an identifier. In this case a boom can be rated more or less strong not only according to its level but also by taking into account the amount of prominence in regards with the frequencies around the resonance which is creating the sensation of the boom.

As a conclusion, the Prominence Ratio is a powerful and reliable tool that is used in Rieter Automotive as a fast detector of possible annoyances caused by middle-high frequency pure tones and narrow-band noises. The PR is used not only to quantify the annoyance but also to target specific noise features and to design the treatment to improve the sound quality of a car.

3 Speech transmission index STI

3.1 Speech transmission index - Definition and advantages compared to AI

The Speech Transmission Index (STI) as exposed in IEC standard n.60268-16 [6], is 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 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 6**). The intensity of the emitted signal for modulation frequency F is given by:

$$I_i \cdot (1 + m_i \cos(2\pi F(t)))$$

The intensity of the signal at the receiver is:

$$I_0 \cdot (1 + m_0 \cos(2\pi F(t + \tau)))$$

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

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

According to the standard, the modulation transfer function can be derived either from test signals sinusoidally modulated in intensity or from the measurement of the impulse response (IR) of the system (usually done in commercial software like MLSSA or DIRAC). This impulse response is obtained by an MLS (maximum length sequence) measurement technique, using a MLS sequence, filtered with a human speech shaped frequency spectrum.

The STI is got from 94 modulation transfer functions, taking in account auditory masking and absolute hearing threshold, and with the octave weighting factors given in [6]. STI goes from 1.0, when the intelligibility is optimal, to 0.0 when understanding is impossible.

3.2 "Noisy" and "Noise free" Impulse Response Techniques for STI measurement

Commercial softwares measure the impulse response in presence of a background noise, i.e. the artificial mouth is emitting in a "noisy" environment and both the room characteristics (reverberation, echoes, absorption) and the background noise may simultaneously deteriorate the impulse response.

On the contrary the technique developed by University Parma [7], [8] computes the modulation transfer function m'(f) from an IR in absence of background noise using equation (1) and then the real *m(f)* can be derived taking in account the effect of background noise with the following expression:

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

where the L_{noise} and L_{signal} are the noise and signal sound pressure levels (in dB) in the considered octave band.

Measurements inside a vehicle have been performed with the binaural microphone on the driver's position and the artificial mouth in the rear left passenger position (see positions in Figure 7). Based on the "noisy" impulse reponses, STI evaluated with 2 commercial softwares (MLSSA, DIRAC). Based on the impulse response in the standing car ("neutral") and the background noise, STI is computed with the "noise free" technique. Comparisons are shown in **Figure 8** or various driving conditions. Much lower standard deviations are found for the "noise free" IR technique, which appears to be better suited for STI computation in the case of a relatively low Signal to Noise ratio, which is the general case for human speech in a driving car [8]. Moreover the "noise free" IR technique is more precise in the lower STI range [8].

Using the “noise free IR” technique, there are also some practical advantages: it is possible to measure the impulse response in the laboratory with engine off, and then to perform separately a car noise measurement under different driving conditions, including on-road measurements. Similarly, modifications affecting only the background noise, such as changes of the engine compartment trim or underbody, will not require an additional impulse response measurement for the determination of STI.

3.3 Measurements

A measurement campaign of STI has been performed on a D-segment five-door vehicle. University Parma has built the artificial mouth that was used and checked that its directivity and level are compliant with standards [9]. The artificial listener used is the Cortex MK1 head with linear equalisation (LO). Time domain responses both for background noise (car alone) and the mouth emitting the standardised male speech MLS signal in the running car have been measured. The transfer function from mouth to listener has also been recorded without background noise. The impulse response with and without background noise has been obtained by deconvolution of the MLS responses with the AURORA software.

Measurements have been performed in the standing car with the engine off ("neutral" condition without background noise) and for following driving conditions on the roller bench:

- idling
- constant speed 70km/h; 3rd gear
- constant speed 90km/h; 5th gear
- constant speed 110km/h; 5th gear

All the measurements were done on the roller bench, i.e. without aerodynamic noise. On-road measurements in the same driving conditions would probably lead to different STI values.

With increasing speed, i.e. with higher background noise, STI values decrease. All STI values obtained for different speeds in the running car are below 0.65. At the speed of 110km/h, STI values between 0.2 and 0.3 are found. This shows the need of using an STI evaluation method that is precise enough in the low STI range. The "noise free" impulse response technique responds to this criteria.

STI gives an indication on how the intelligibility varies depending on the positions of speaker and listener. Two different STI values are obtained at the driver position depending whether the speaker is seated in the front passenger position (F) or in the rear position (I). **Figure 9** shows the corresponding results measured with a single microphone: STI is 0.503 with speaker in position F (speaking towards the windscreen) and 0.567 with the speaker in position I, i.e. just behind the driver position. Comparatively, an analysis through Articulation Index would provide a single value for the driver position, taking into account the effect of background noise on this position.

Moreover, the STI of rear passenger (I) speaking to the driver (STI=0.567) is much higher than the STI of 0.451 of the front passenger speaking to the rear passenger (L). This high difference is probably due to the fact that in this last case, the artificial mouth is seated on the front passenger position and directed towards the windscreen, so the speech arrives only indirectly after some reflections to the listener. In this position, the directivity characteristics of the artificial mouth are possibly of importance on 360° azimuth angle, and not only on the 30 degrees prescribed by the standard ITU P.51 [9].

3.4 A tool to analyse the influence sound package modifications

The sensitivity of STI to treatment changes of the passenger compartment, has been evaluated through a series of measurement in various trim configurations, as shown in **table 1**:

	Configuration	**Description**
1	original package	Serial vehicle state
2	rigid	Impervious thick foil on top of floor, seats (visible surfaces), headliner, parcel shelf
3	absorbing	20mm felt on top of: floor, doors, headliner, IP, parcel shelf, pillars
4	absorbing with reflective patches	Absorbing configuration with 2 A4 formats of aluminium 400μm and no felt on the B-pillar driver side - see **figure 10**
5	absorbing with more reflective patches	Absorbing configuration with 4 A4 formats of aluminium 400μm and no felt on the B-pillar driver side
6	RUL materials	Headliner, floor and parcel shelf replaced by prototype parts in RIETER ULTRA LIGHT™ technology

The "rigid" and "absorbing" state are non-realistic treatment configurations that allow big changes in the vehicle reverberation time. The realistic configurations "original" and "RIETER ULTRA LIGHT™ materials" represent a conventional sound package as compared to an ultralight package with higher absorption and less insulation.

In **Figure 11**, STI values computed for configurations 1 to 5 are plotted. Clearly the configurations with higher absorption "absorbing", "absorbing + aluminium" and "absorbing more aluminium" have higher STI values than the original and rigid configuration, i.e. intelligibility in a car treated accordingly would be better. This is due to the fact that the background noise is lower in the "absorbing" configurations, while the transmission from speaker to listener is not so much deteriorated. In "absorbing + aluminium" (Figure 10) and "absorbing more aluminium" configurations the aim was to increase the speech signal reverberation in order to get a better STI, while keeping the overall absorption high. STI was found higher in the left channel, probably because the aluminium panels were placed only in the left part of the headliner, so that the increased reflections influenced mainly the left ear of the mannequin. In the right ear channel, the addition al aluminium reflectors lead even to a slight deterioration in STI, however with STI values still significantly higher than for the left ear.

We tried also to create a virtual configuration with features of two different packages: a STI value with test signal from "original" configuration and background noise from "absorbing" package was computed (only for the right channel) and we got a better STI than with all the other packages. Results for all driving conditions are shown in **Figure 12**. This numerical example shows that best values for STI are obtained when using a package with good absorbing properties (to have low background noise) and simultaneously providing good reflections of the speech signal inside the car. It also shows how this approach can separate the different effects of the vehicle treatment on intelligibility and thus shows the potential of STI to be used as analytical tool for sound package improvement.

A comparison between results for the original sound package and a package on the basis of RIETER ULTRA LIGHT™ parts (except for the dash) has been made. In the lightweight package, prototype parts replace the headliner, the floor and the parcelshelf, with a weight saving of 5.8kg (see **Figure 13**). This package is more absorptive than the original trim, while having a lower transmission loss (especially in the floor area). By definition, STI represents both the reduction of speech signal transmission due to the increase in absorption and the effect of the modified background noise, due to modifications both on absorption and on insulation. Results of the STI computations are shown in **Figure 14**. For the driving condition 70km/h, the two packages are comparable according to STI. Improvements in STI with the lightweight package are obtained at high speeds (90km/h and 110km/h), especially for the left driver ear position, as shown on Figure 14a.

Finally the results of STI calculations have been compared to the values of Articulation Index calculated for the same background noises. The computations have been performed with the Artemis software using the non-extended Articulation Index with values between 0 and 1 (as for STI). Calculation results are shown in **Figure 15**. Ultra Light gives improved AI results for all driving conditions (70km/h, 90km/h and 110km/h). The sensitivity to the influence of the change in sound package, depending on the ear position (Left or Right) seems less important for AI than for STI. Thus STI gives a slightly more detailed analysis of the influence of the sound package on intelligibility, by taking into its influence on speech transmission and by showing a higher spatial sensitivity. Moreover, except the initial MTF measurement in the standing car, STI can be computed based on similar measurements than AI, when using the "noise free IR technique", developed at University Parma. Thus the Speech Transmission Index, is a promising tool (complementary to AI) for the assessment of intelligibility of lightweight absorptive packages.

4 Future work

The next step in the evaluation and improvement of binaural recording and replay facilities is the construction of a listening room.

Activities for coupling the replay of recorded noise with sound package simulation, allowing listening to a known car with a modified numerical treatment are foreseen, in particular an auralisation tool in the internal DIAMONDS software is planned. Moreover a jury test design taking into account specifically sound package modifications is foreseen.

Finally, Rieter will pursue the integration of subjective evaluation into the standard vehicle noise assessment procedures. A key topic here is minimising the effort:

- in measurements (for example using the noise free STI measurement procedure)
- in post-processing (by stream-lining and automating signal treatment)
- in jury tests (design of the test and multiple replay facilities in the replay room) .

while keeping the best possible level of precision.

Subjective data will also be integrated into Rieter's database structure together with objective measurement values for materials, parts and vehicles. Together with SPL measurements or simulations, the tools presented in this article will be used as a diagnostics one to help reducing perceived discomfort or as evaluation and comparison tool to guide sound package design and increase the perceived comfort.

Conclusions

A series of tools from binaural recording to adapted metrics can be used for the assessment of sound quality in automotive interiors. Available tools as well as research activities have been presented with special emphasis on their applicability to the automotive field and link to sound package.

References

[1] M.A. Nobile, G.R. Bienvenue " A procedure for determining the prominence ratio of tones in noise emissions", Noise-Con, Tarrytown1991

[2] A.C. Balant, K. Barringer, R.D. Hellweg " A comparison of two methods for the evaluation of prominent discrete tones - Psychoacoustic Results", Internoise 2000

[3] R.D. Hellweg, M.A. Nobile, A.C. Balant "A comparison of two methods for the evaluation of prominent discrete tones - Phase II", Internoise 2000

[4] R.D. Hellweg, M.A. Nobile "Modification to procedures for determining prominent discrete tones", Internoise, Dearborn 2002

[5] C. Castagnetti "Sound Quality of Emergent Sounds in the Automotive Field", Tesi di Laurea, Università di Roma 2001

[6] CEI / IEC Standard Nr. 60268-16 "Sound System Equipments - Objective Rating of Speech Intelligibility by Speech Transmission Index", Second Edition, March 1998

[7] A. Farina, F. Bozzoli "Measurement of speech intelligibility inside cars", 113th AES Convention, Los Angeles, October 2002

[8] A. Farina, F. Bozzoli, P. Strasser "Comparative Study of Speech intelligibility inside Cars", Euronoise 2003, Naples

[9] ITU-T Recommendation Nr. P.51 "Telephone transmission quality - Objective measuring apparatus- Artificial Mouth", August 1996

Fig. 1 + 2
Titel

Fig. 1 Sound Quality Assessment Method

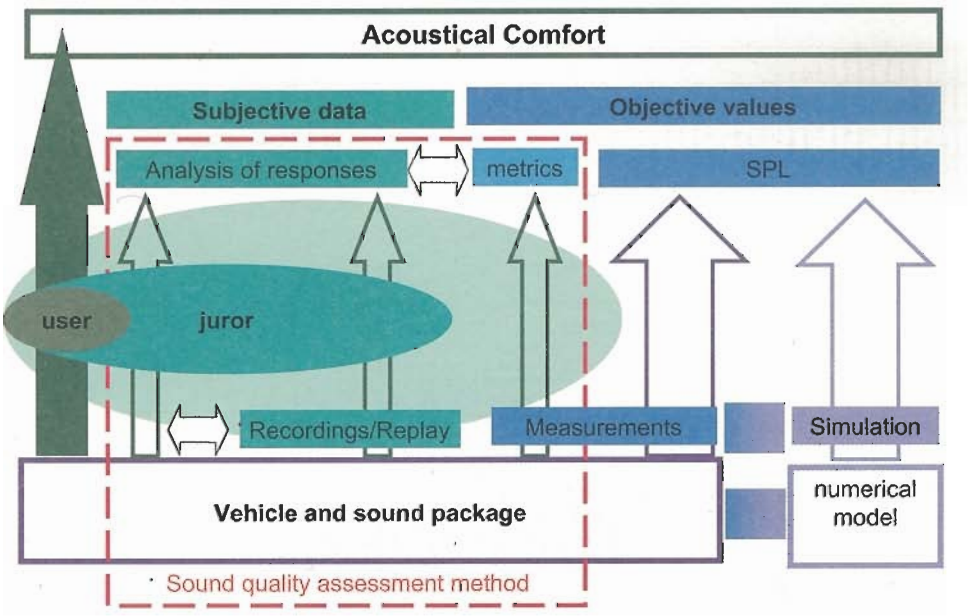


Fig. 2 Spatial characterisation of binaural recording devices

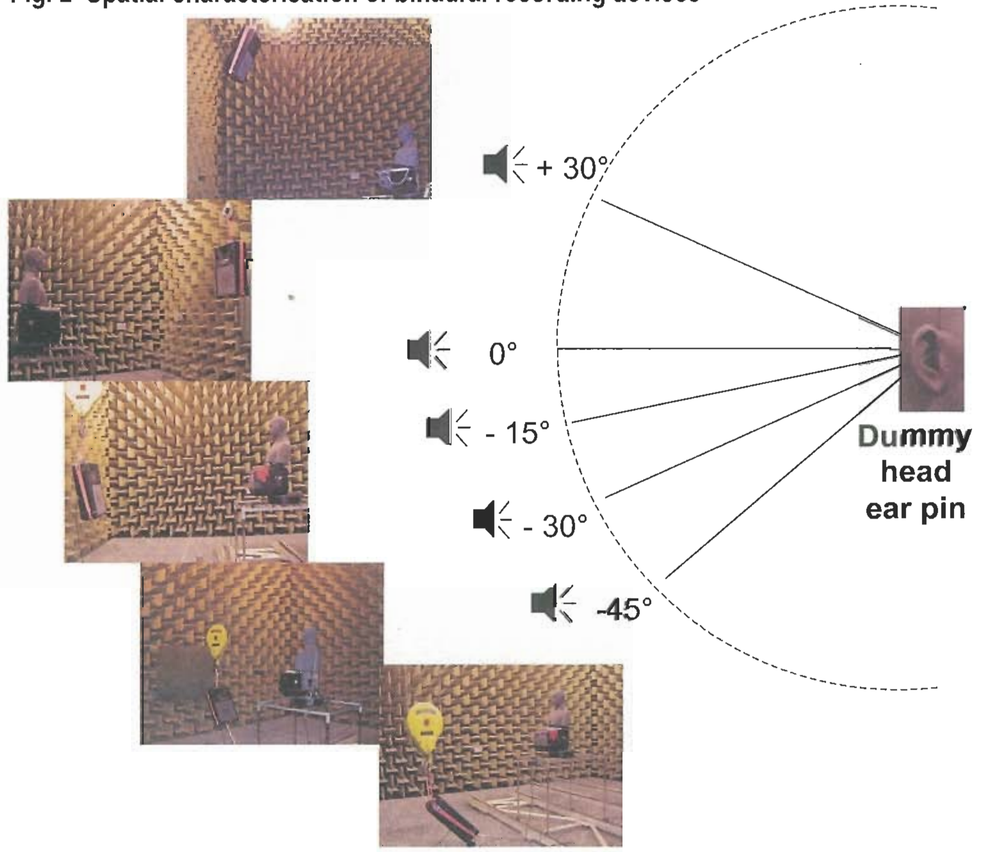
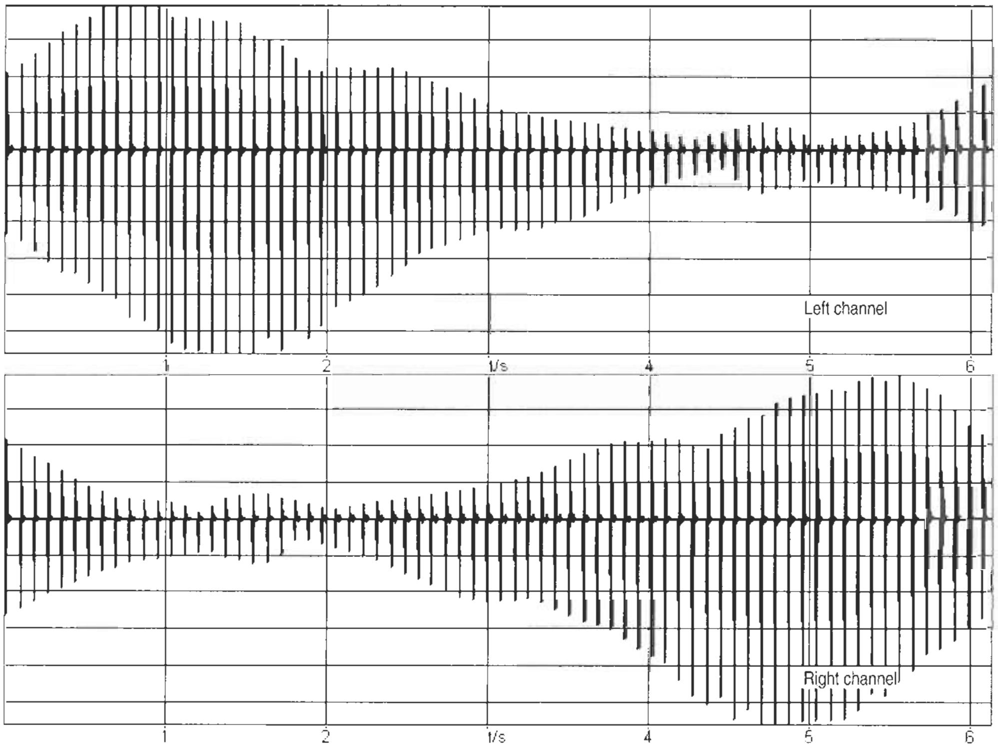


Fig. 3
Impulse response of binaural recording devices

a) Good spatial response



b) Device will lead to errors in the spatial perception of recorded sound

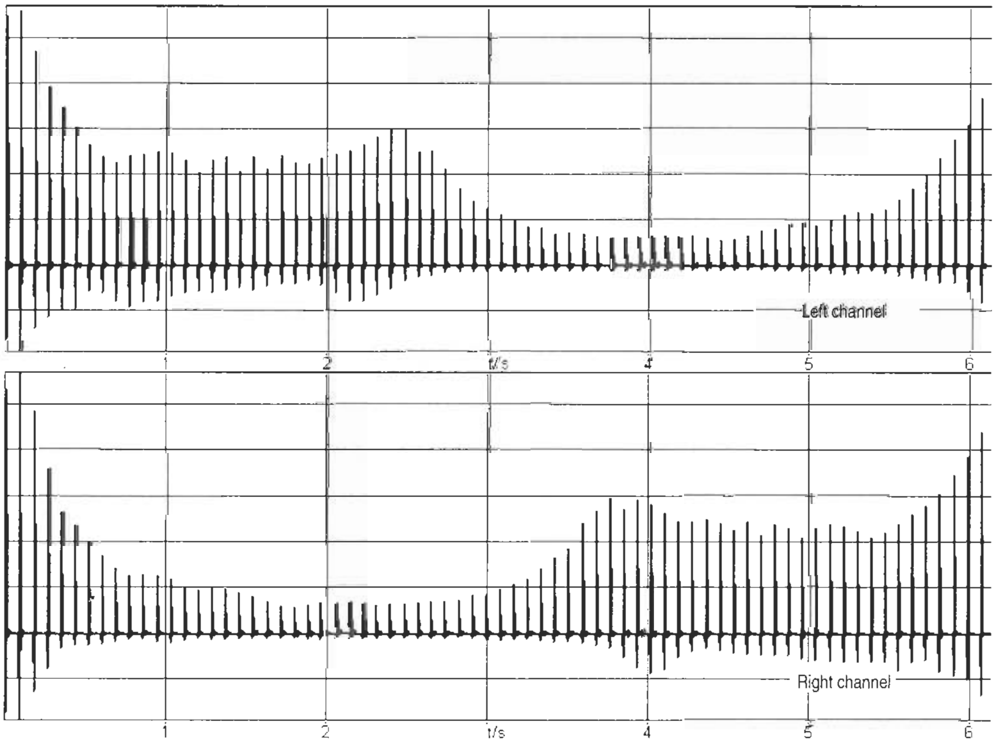


Fig. 4

Comparison of PR analysis with standard FFT spectrogramm for gear whine noise

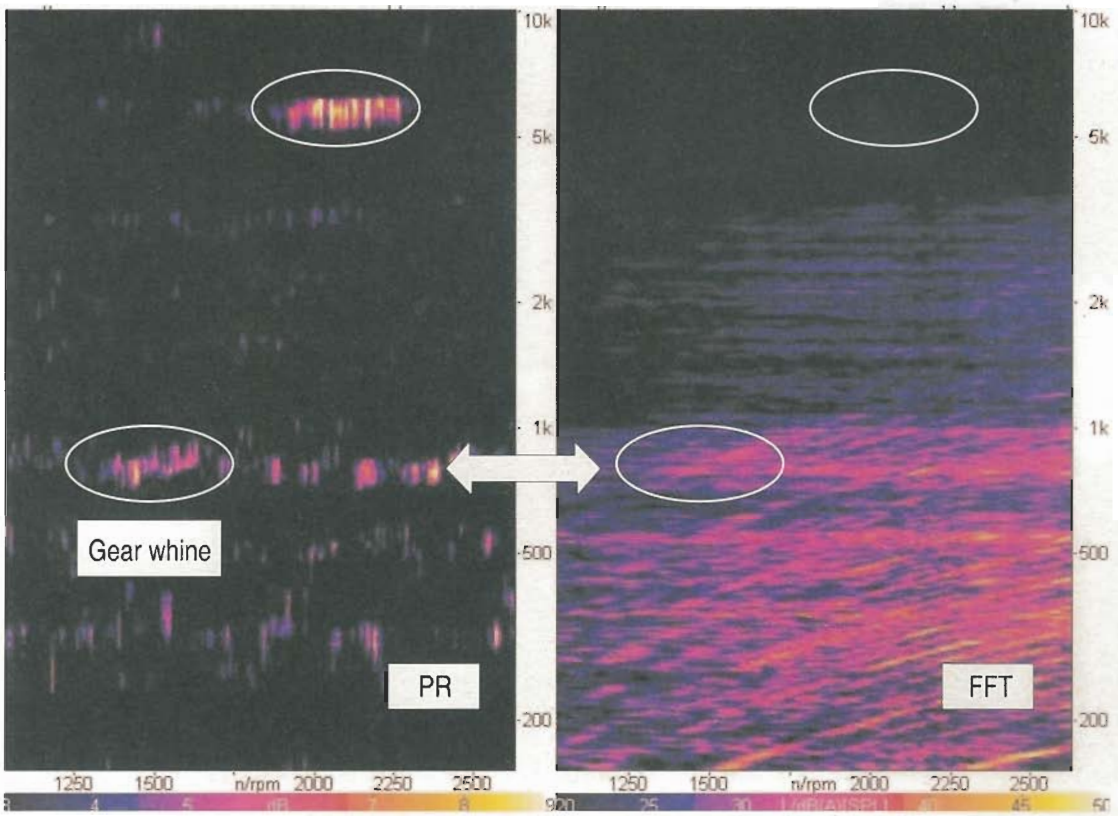


Fig. 5

Comparison of PR analysis with standard FFT spectrogram for Diesel nock

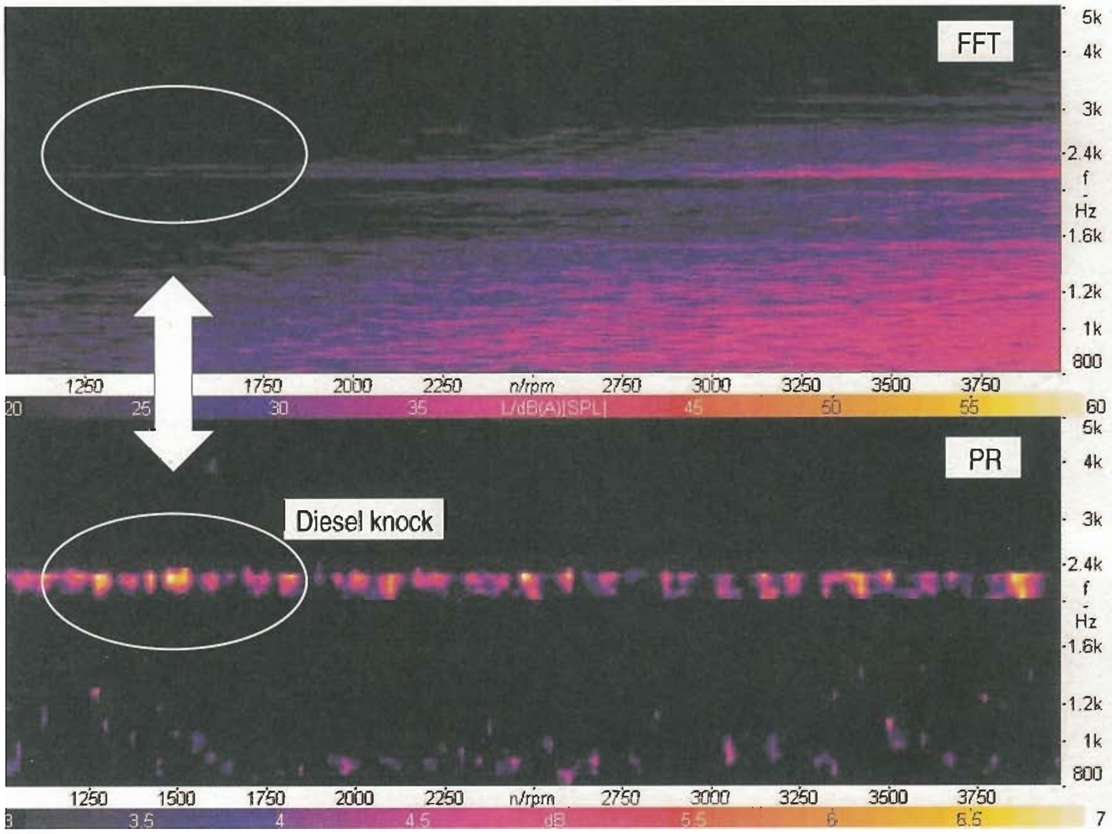


Fig. 6 + 7

Fig. 6 STI - Definition of modulation transfer function m(F)

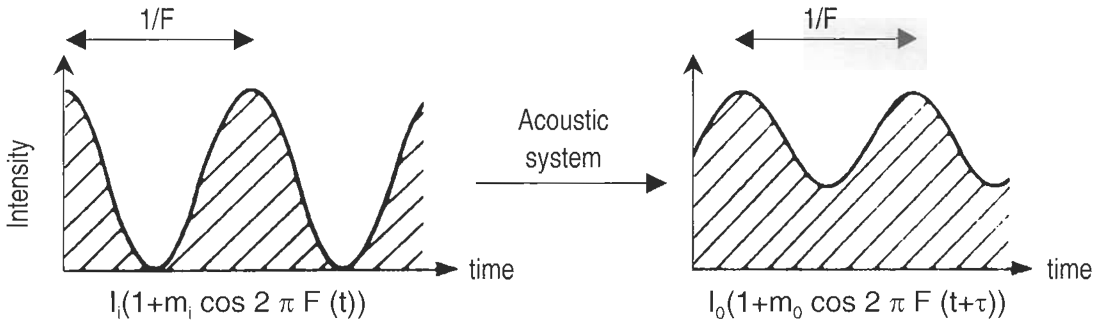


Fig. 7 STI Measurement set-up

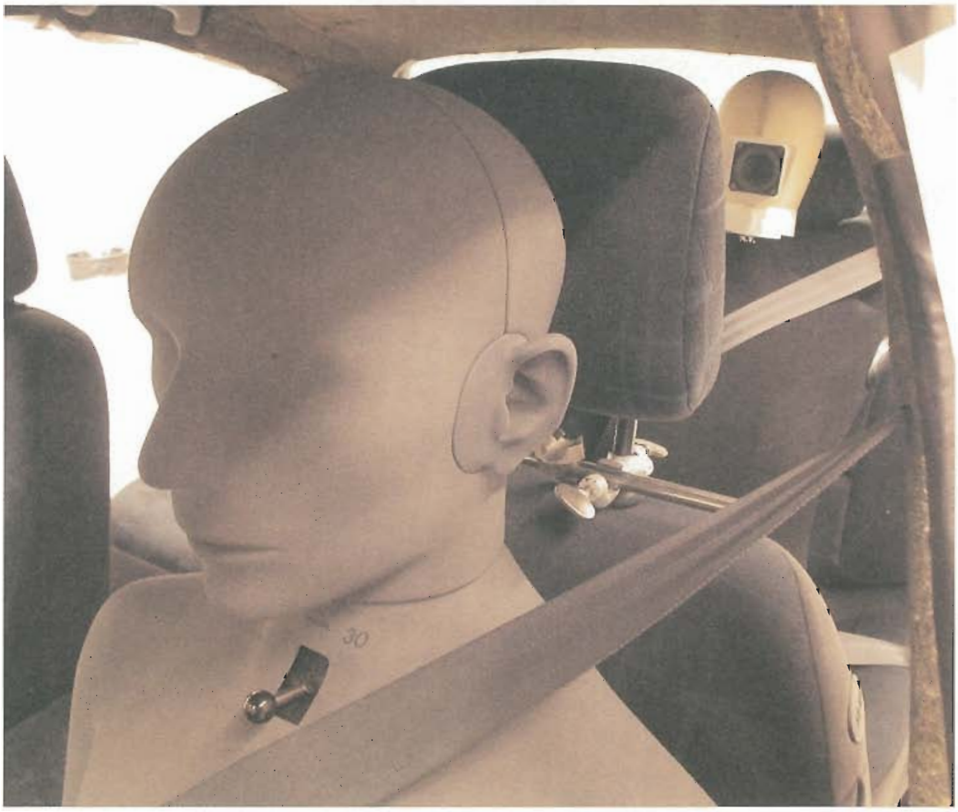


Fig. 8 + 9

Fig. 8 Results STI "noisy" / "noise free"

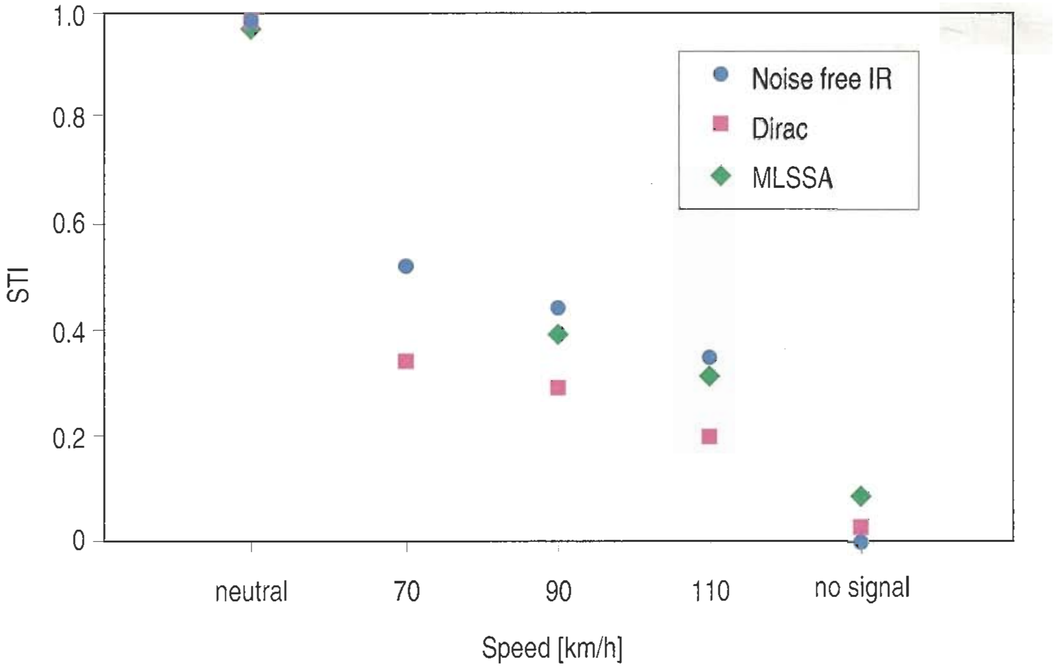


Fig. 9 Spatial sensitivity of STI

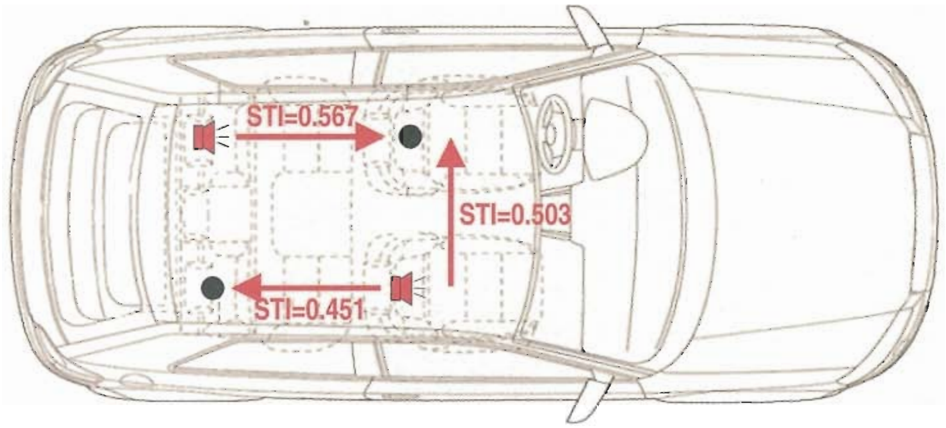


Fig. 10 + 11

Fig. 10 Configuration with high absorption and aluminium reflectors

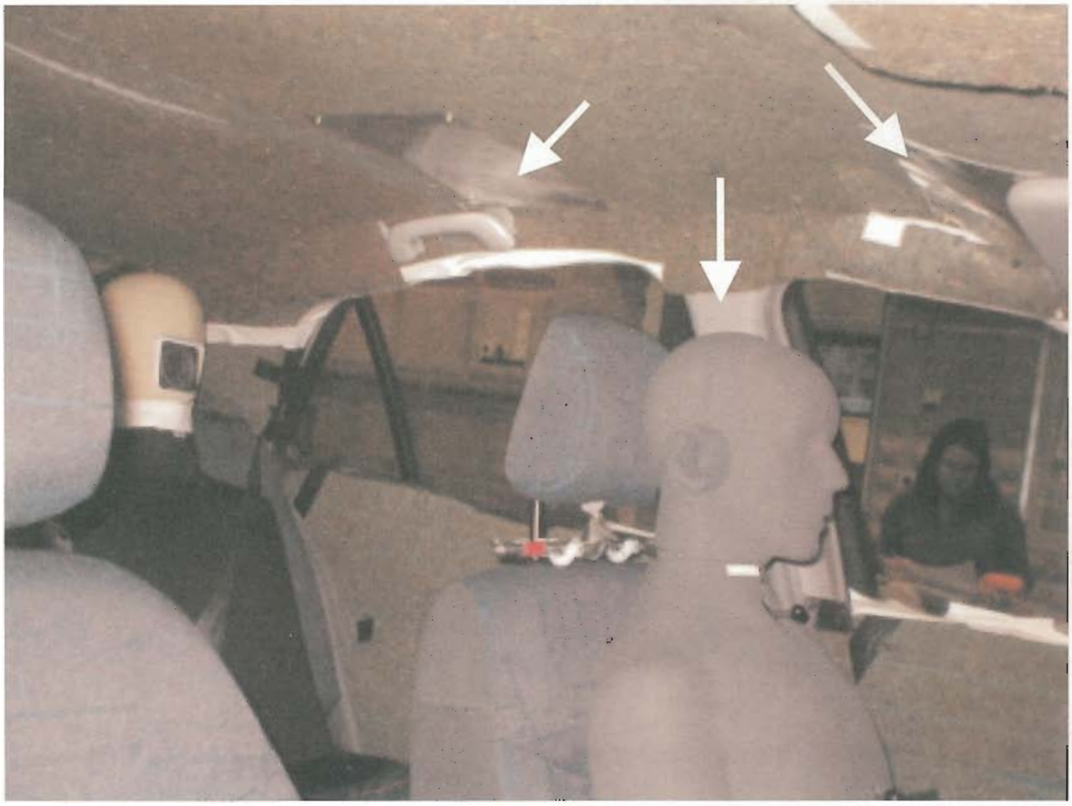


Fig. 11 STI sensitivity to treatment changes and to additional "reflectors"

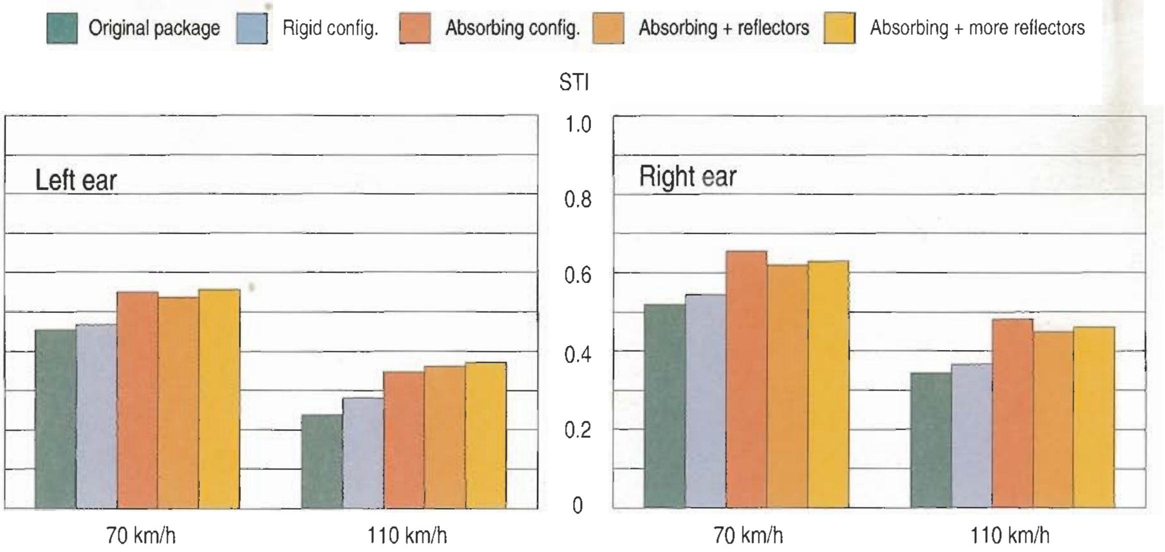


Fig. 12 + 13

Fig. 12 STI for measured "original" and "absorbing" configurations compared to the virtual test case (Cortex D, Mouth I)

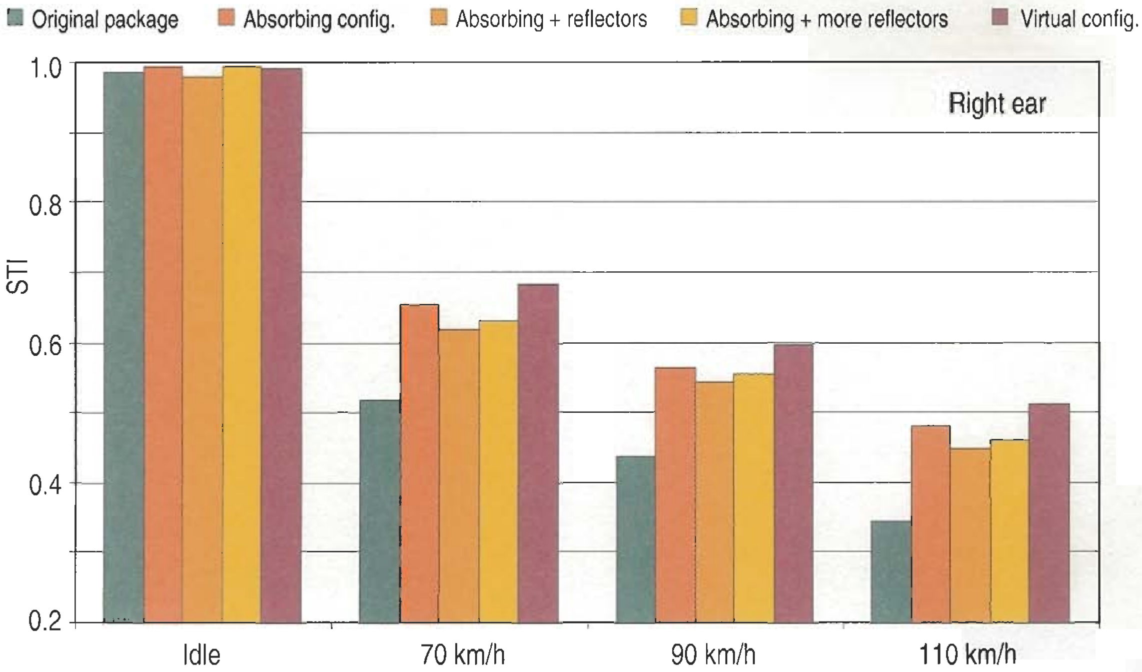


Fig. 13 Vehicle configuration with Rieter Ultralight materials

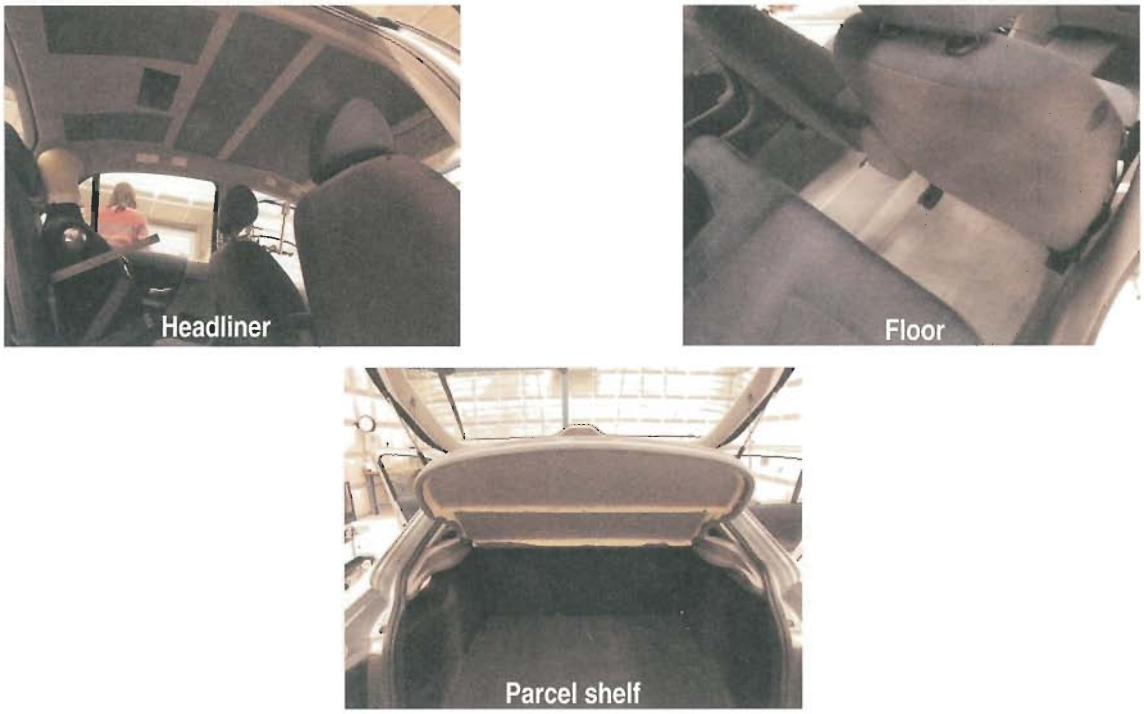


Fig. 14 + 15

Fig. 14 Comparison of Speech Transmission Index for Original and Ultra Light package

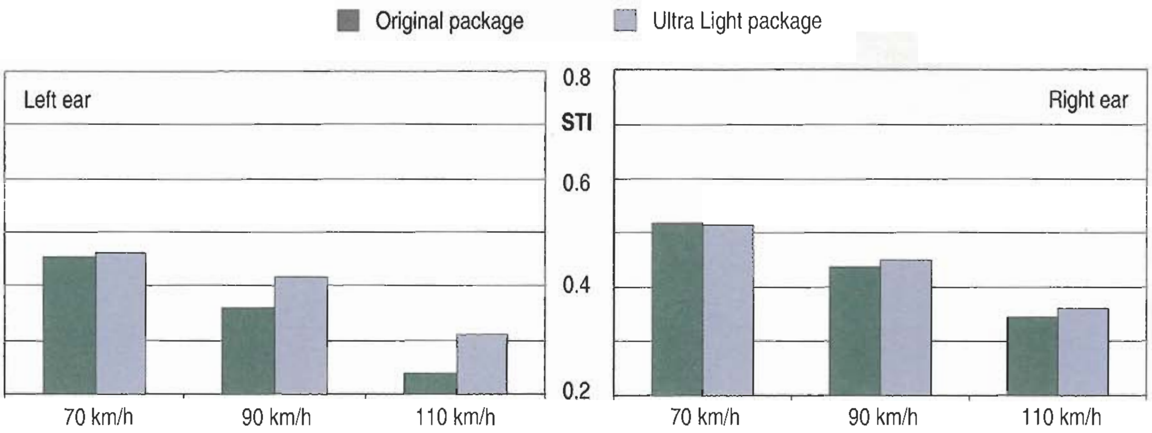


Fig. 15 Comparison of Articulation Index (non extended) for Original and Ultra Light package