Instrumental Testing of In-Car Communication Systems

Udo Muesch, Frank Kettler
HEAD acoustics GmbH, 52134 Herzogenrath, E-Mail: Udo.Muesch@head-acoustics.de, Frank.Kettler@head-acoustics.de

Introduction

In-car-communication systems (ICC) shall support and ease the communication in vehicles at higher noise levels especially between driver and passengers in the first or second row. The aim is to increase intelligibility and decrease listening effort respectively but preserve natural and high speech quality. Consequently, the speech level at the passenger's position needs to be increased by ICC systems without amplification of the noise level. Thus, the signal to noise ratio (SNR) at the passenger's position shall be increased. However, considering the acoustic coupling between loudspeaker and microphone in the vehicle cabin (see **figure 1**), the gain factors, equalization and delay introduced by the ICC system are underlying certain restrictions. The adjustment of technical parameters for tuning purposes is challenging. Consequently, requirements can only be derived from auditory testing. Following former investigations on evaluating the quality of ICC systems [1], [2] this contribution discusses auditory tests suggesting a simplified ICC model to generate listening examples. An appropriate test setup including background noise simulation technology is introduced and measurable parameters are discussed.

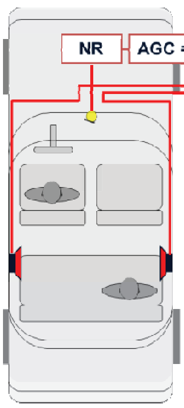

Figure 1: ICC signal processing

Generation of listening examples for auditory testing

Listening examples for auditory testing can be generated using ICC systems in vehicles and recording appropriate speech material. However, the parameters are underlying certain restrictions to guarantee system stability. Therefore a simplified ICC model from **figure 2a** is used for the generation of listening examples (S_{LOT}).

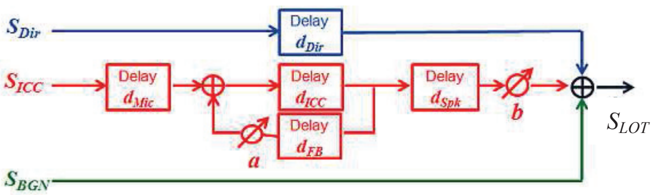

Figure 2a: Simplified ICC model for generation of sound samples

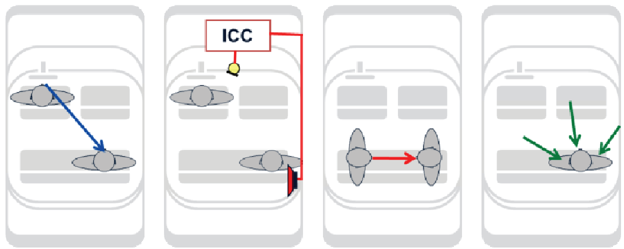

Fig. 2b **Fig. 2c** **Fig. 2d** **Fig. 2e**

Figure 2b – 2e: Individual contributions to S_{LOT}

The sounds presented in the listening tests (S_{LOT}) consist of the direct sound between talker and listener (S_{Dir}), the signal that is processed and transmitted by the ICC system (S_{ICC}) and the driving noise (S_{BGN}). The direct sound is recorded in a real car cabin between two artificial heads (Head and Torso Simulator, HATS) positioned on the driver's seat and backseat passenger's position in the first row (S_{Dir}, see **figure 2b**). It already includes the room acoustics (reverberation) of the vehicle cabin. The delay d_{Dir} is determined by the transmission path in the car cabin. The transmission path of the ICC system (**figure 2c**) is modelled by a delayed Dirac Impulse with the system delay d_{ICC}. The acoustic feedback path is approximated by the delay d_{FB} (defined by the distance between the speaker and the microphone) and the damping constant a. The input signal S_{ICC} is simulated by two artificial HATS facing each other on the passenger's seats (see **figure 2d**). The recorded signal is delayed by the acoustic propagation from the artificial mouth to the ICC microphone (d_{Mic}) and by the delay d_{Spk} considering the propagation between the ICC loudspeakers and the artificial ear on the passenger's position. The gain factor b is freely adjustable in this simulation. The third component is the recorded background noise in the driving car, which is added as S_{BGN} (**figure 2e**). This model combines the acoustics of the car cabin and the typical feedback path inherent in ICC systems with full control over amplification and ICC system delay beyond the physical limitations of existing implementations. For the generation of the listening examples S_{LOT} the ICC processing delay d_{ICC} is adjusted to 5 ms in this evaluation, which can be seen as a recommended limit for ICC signal processing.

Recordings were carried out in four different vehicles. The driving noise covered a wide range between 68 dB_{SPL}(A) and 78 dB_{SPL}(A). Simulated SNR were adjusted between -9 dB and 10 dB. Test persons were asked to give their opinion about the preferred loudness of speech with the explicit hint to consider the understanding of the speech material. A five point, centric preference scale (too loud (-2), slightly loud (-1), appropriate (0), slightly weak (1), too weak (2)) was used. The speech material consists of single meaningful words (native American English), the speech material was recorded in a driving simulation including the Lombard Effect. 11 subjects participated, the test corpus covered 79 stimuli leading to an overall amount of 869 single ratings.

Preferred SNR

The MOS for the test conditions spread over the whole range between -2 up to 2 MOS. The preferred SNR is represented by a MOS of 0 (appropriate speech level). The confidence interval did not exceed 0.5 MOS for all test conditions. It can be regarded as low considering the limited number of 11 test subjects.

The following four figures show the preferred signal-to-noise ratio (MOS on the y-axes) as a function of the SNR. The background noise conditions are subdivided into four categories. For the low background noise scenarios between 68 and 70.5 dB(A) the relationship between MOS rating and the preferred SNR is relatively vague. A linear interpolation indicates a preferred SNR of approximately 5 dB (4.9 dB, see red dot in **figure 3a**) for these noise scenarios. In principle, the SNR in this range can also be confirmed by test specifications for hands-free communication in motor vehicles (see ITU-T Recommendation P.1100/1110 [3], [4]). A signal-to-noise ratio of 6 dB is given as an appropriate playback level in the presence of background noise here.

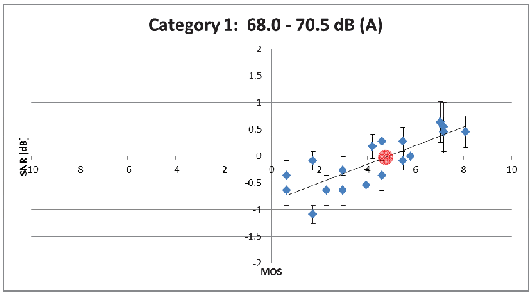

Figure 3a

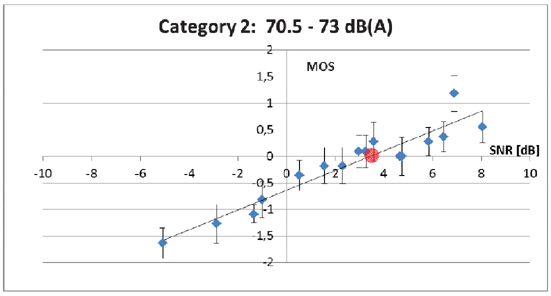

Figure 3b

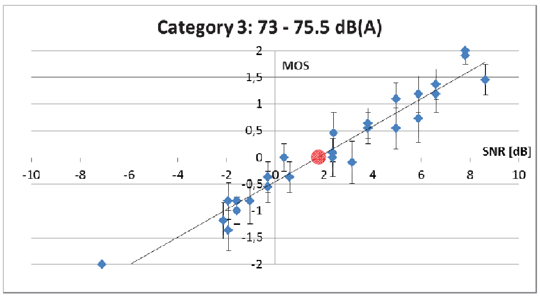

Figure 3c

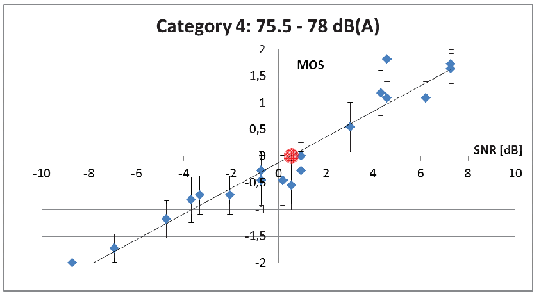

Figure 3d

Figure 3a – 3d: Preferred SNR (MOS, x-axis) for different noise categories (SNR, y-axis)

For higher noise levels (category 2, 70.5 – 73 dB(A), **figure 3b**) a more systematic relation between preferred SNR and the SNR test condition can be derived. Surprisingly, the preferred SNR (MOS 0) decreases to 3.5 dB for the higher noise scenario. The same tendency pursues for the two higher noise categories (category 3, 73 – 75.5 dB(A), **figure 3d** and category 4, 75.5 – 78 dB(A), **figure 3d**). The preferred SNR decreases to 1.8 dB (category 3) and 0.5 dB (category 4). These results clearly indicate that the preferred SNR can be expressed by a function of the noise level itself and the overall level respectively.

The preferred SNR limits are summarized in **figure 4** for the four categories. The red bars represent the MOS range of 0 ± 0.5, the blue bars represent the wider MOS range of 0 ± 1. The numerical values are given in **table 1**.

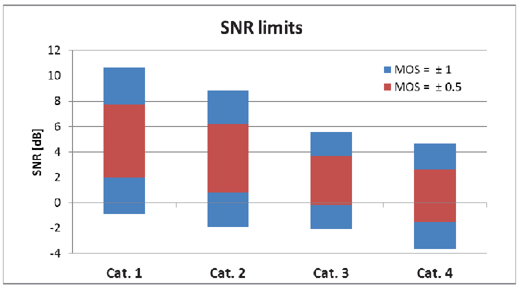

Figure 4: Preferred SNR (x-axis) for 4 noise categories (y-axis)

	Cat. 1	Cat. 2	Cat. 3	Cat. 4
preferred SNR	4.9 dB	3.5 dB	1.8 dB	0.5 dB
0 ± 0.5 MOS range	2.0 – 7.7 dB	0.8 – 6.2 dB	-0.2 – 3.7 dB	-1.6 – 2.6 dB

Table 1: Preferred SNR values and 0 ± 0.5 MOS range

These results can be used for ICC parameter setting, in particular the noise dependant gain control. They can further be used for instrumental testing to characterize ICC systems.

Instrumental tests - black box approach

Testing integrated ICC systems in vehicles without access to the microphone signal, the loudspeaker connections or the IVS signal processing itself requires the acoustic stimulation of the ICC system by acoustic noise playback in the vehicle cabin. For this purpose typically a background noise simulation system according to ETSI EG 202 396-1 [5] can be used. Such a test setup is also described in ITU-T P.1100 [3] and P.1110 [4]. The playback system is typically equalized on the ICC microphone position in order to provide the correct spectral and level characteristics, indicated by the green arrows in **figure 5**.

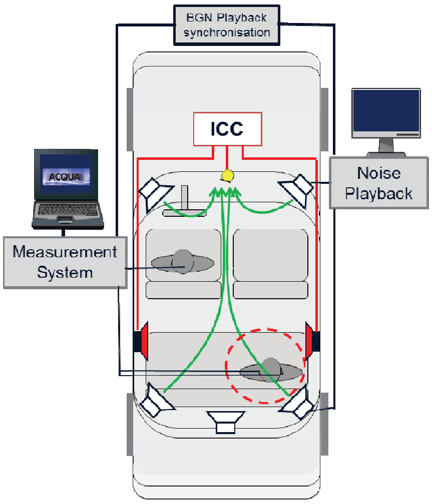

Figure 5: Measurement setup with background noise simulation system (noise playback)

Vice versa, this implicitly leads to wrong noise equalization for all other positions in the car, in particular for the measurement position on the back seat (indicated by the red dotted line in **figure 5**). Consequently, it is necessary to compensate this driving noise at the measurement positions to be able to analyze the ICC processed speech (played back via the ICC loudspeakers) without background noise or to add the correct recorded driving noise for this vehicle position before further analyses (see also [2]).

If the noise playback system and the measurement system (recording and analyzing the signals at the passenger's position) are accurately synchronized in the time domain (see *BGN playback synchronization* in **figure 5**), the noise playback can be compensated in the time domain by subtracting the noise signal (noise-only) from the speech and noise recording. This is denominated as Time-synchronized Noise Compensation (TNC) in the following.

The effectivity of TNC is demonstrated by the spectrographic analyses shown in **figure 6a** to **6d** (**6a**: speech and driving noise, **6b**: driving noise only). The Time-synchronized Noise Compensation leads to the residual speech signal shown in **figure 6c**. For comparison the speech-only signal recorded without coincident playback of noise in the vehicle is represented by **figure 6d**.

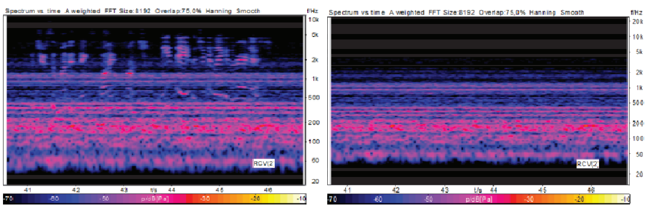

Fig. 6a: Speech and Noise **Fig. 6b:** Noise only

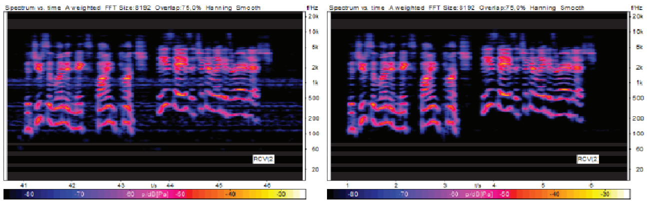

Fig. 6c: Speech after TNC **Fig. 6d:** Speech w/o noise

Figure 6a – 6d: Spectral analyses of speech and noise before and after TNC

The speech characteristics are nearly unchanged which can also be proven by acoustic verification. The active speech level [6] and noise level is given in **table 2** for the four scenarios. The TNC leads to a noise level reduction of more than 30 dB, the active speech level ASL can be accurately determined after noise compensation.

	Speech and Noise	Noise only	Speech after TNC	Speech w/o noise
Noise Level	67.5 dB(A)	67.5 dB(A)	36.5 dB(A)	---
Act. Speech Level	---	---	65.9 dB	65.9 dB

Table 2: Noise levels and ASL before and after TNC

Measurable parameters

The application of TNC for noise compensation leads to very accurate analysis possibilities for characteristic parameters of an ICC system, if applied for the binaural background noise recordings of the HATS positioned on the passenger's seat. **Figure 7** shows the analysis of the (undesired) amplification of noise by an ICC system installed in different vehicles. The system was configured with different parameter settings. The three analyses show the noise amplification vs. frequency together with a tolerance of 3 dB. The two systems A and C do not lead to a significant amplification of noise, whereas system B is configured with a high amplification which also affects the noise. The red colored curve, representing the right ear of the HATS, is above the tolerance. The ICC system amplifies the noise, this is more audible on the right ear signal of the HATS in the setup used due to the closer distance of the right ear to the ICC loudspeaker.

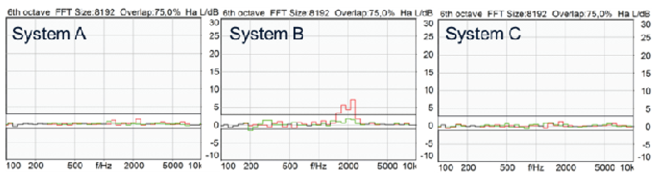

Figure 7a **Figure 7b** **Figure 7c**

Figure 7a – 7c: Analysis of background noise amplification vs. frequency

Vice versa, very accurate analyses can also be carried out for the processed speech. **Figure 8** shows the speech amplification factor vs. frequency again for three different configurations of an ICC test system. System A, which is optimized for the vehicle used for testing, shows an expected gain over the entire frequency range with stronger emphasis on higher frequencies above 1 kHz (**figure 8a**). System C leads to a more uniform gain distribution over the entire frequency range in **figure 8c**, which leads to a muffled sound. Again system B was configured for demonstration purposes with an aggressive gain setting. The red curve indicates a significantly too strong amplification which is especially audible and annoying for the right ear signal of the HATS on the passengers position (**figure 8b**).

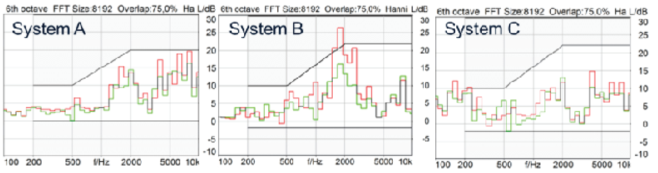

Figure 8a **Figure 8b** **Figure 8c**

Figure 8a – 8c: Analysis of speech amplification vs. frequency

TNC also provides the possibility to analyze parameters in the time domain like the activation time, i.e. the duration until the full amount of amplification of the ICC system is established. System A and C in **figure 9a** and **9c** show the complete activation after approximately 25 s whereas the full amplification of the transmitted signal can be analyzed after approximately 30 s for system setting B (**figure 9b**).

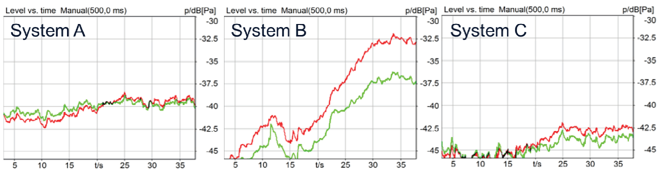

Figure 9a **Figure 9b** **Figure 9c**

Figure 9a – 9c: Analysis of activation time

Further parameters like the absolute speech signal level and the resulting signal-to-noise ratio, the calculation of speech intelligibility index (SII) or the ICC processing delay can be determined in such a test setup. Beside the absolute analysis of these parameters, relative results can also be derived by comparing the ICC performance to corresponding measurements without activating the ICC system. **Table 3** shows some parameters for the three ICC system settings A, B and C operated in two different vehicles with different background noise levels (BGN). According to the SNR category derived from the auditory tests a target SNR of 3.5 dB (range between 1 and 6 dB considering the MOS variation of 0 ± 0.5) is derived for system A and C operated in the vehicle with a background noise level of 71.4 dB(A). A recommended SNR of 2 dB (range between 0 and 3.8 dB) is derived from the auditory tests for system B operating in a background noise environment of 73.6 dB(A). The SNR is calculated together with the SNR improvement and is given in **table 3**. The result of 6.3 dB and 6.8 dB respectively for the two systems A and C are in a reasonable and recommended range for well-tuned ICC systems. The 12.7 dB SNR measured with system B is approximately 6 dB too high which can also be verified by a significantly degraded speech quality.

Similar results can be calculated for the SII and SII improvement. Again system A and C with SII improvements between 13 and 14 % represent reasonable parameter settings of these ICC systems. Vice versa, the SII result for system B is high but does of course not take into account the significant degradation in speech quality.

	System A	System B	System C
Noise level	71.4 dB(A)	73.6 dB(A)	71.4 dB(A)
SNR Category	2	3	2
preferred SNR	**3.5 dB**	**2 dB**	**3.5 dB**
SNR (0 ± 0.5 MOS)	1 – 6 dB	0 – 3.8 dB	1 – 6 dB dB
SNR	**+3.2 dB**	**+12.7 dB**	**+3.3 dB**
SNR improvement	6.3 dB	14.3 dB	6.8 dB
SII	55.6 %	61.4 %	56.6 %
SII improvement	13.0 %	16.4 %	13.6 %

Table 3: Measurement results

Conclusion

Auditory tests are carried out to derive reasonable limits for recommended parameter settings for ICC systems. The results also motivate the requirements for instrumental analyses. An acoustic background noise simulation provides the testing capability to activate ICC systems in laboratory tests using a “black box” approach. The background noise equalization at the ICC microphone position guarantees the correct activation of the ICC system. However, this is accompanied by a wrong equalization and wrong noise playback for the different measurement positions in that vehicle. This inaccuracy can be solved by the Time-synchronized Noise Compensation in the time domain when using a highly accurate synchronized noise playback and measurement system. The whole setup then provides the possibility to analyze and characterize ICC systems accurately and finally provides, in combination with the auditory test results, hints and guidelines for ICC tuning and verification.

References

[1] A. Theiss, G. Schmidt, J. Withopf, C. Lüke, “Instrumental Evaluation of In-Car Communication Systems”, Proc. ITG Conference on Speech Communication, Erlangen, Germany, 2014

[2] U. Muesch, F. Kettler and M. Lepage, “Speech Quality Tests of In-Car Communication Systems”, Proc. ITG Conference on Speech Communication, Braunschweig, Germany, 2012

[3] ITU-T Recommendation P.1100, Narrow-band hands-free communication in motor vehicles

[4] ITU-T Recommendation P.1110, Wideband hands-free communication in motor vehicles

[5] ETSI EG 202 396-1, Speech Processing, Transmission and Quality Aspects (STQ); Speech quality performance in the presence of background noise; Part 1: Background noise simulation technique and background noise database

[6] ITU-T Recommendation P.56, Objective measurement of active speech level.