Detection of Silent Calls in Emergency Call Scenarios

Radi Serafimov, Frank Kettler, Silvia Poschen
HEAD acoustics GmbH, 52134 Herzogenrath, E-Mail: Radi.Serafimov@head-acoustics.de
Frank.Kettler@head-acoustics.de, Silvia.Poschen@head-acoustics.de

Introduction

According to the EU commission vehicles need to be equipped with "eCall" (emergency call) systems in the future. In case of accident, a minimum set of data (MSD) including GPS position is transmitted and a voice call via "112" is established to the Public Safety Answering Point (PSAP). Hands-free functionality is used in the vehicle. The audio connection is the only communication channel between PSAP and driver or passengers in the vehicle. Thus, transmission quality is of highest importance.

These systems require new tests and limits compared to conventional hands-free communication [1], [2], [3]. This contribution presents the results of a listening test for the "silent call" problem. A "silent call" in this context designates an emergency call (real emergency case or erroneously generated call, e.g. from a mobile phone) where no one is actively communicating with the PSAP side. Thus, the transmitted noise scenario is the only information for the PSAP operator to decide about the relevance of this call. It is important to conduct auditory tests to define minimum requirements for the transmitted ambient noise from a vehicle involved in an accident.

Motivation

Silent calls, sometimes also designated as "call without connection", are emergency calls which are in most cases today erroneously generated from mobile phones. These calls lead to a significant workload at the PSAP side even today; up to 20 % of emergency calls are reported as silent calls. The PSAP operator of course answers the calls, tries to establish a communication with the dialing subscriber and judges the transmitted background noise in order to verify the severity of the call. If neither voice communication nor any acoustic characteristics in the transmitted background noise indicate an emergency case, the call is terminated. However, for eCalls originating from vehicles, the silent call problem is more severe because such a call is typically generated when sensors in the car (like airbag sensors) indicate an accident. The MSD is transmitted and received on the PSAP side, thus, it is clear that an incident has happened. However, the severity is not clear; the PSAP personnel still needs to decide about the rescue units to send.

Silent calls occur in eCall scenarios, when the call is manually generated by the driver of a car, which is not involved in the accident and the driver leaves the car in order to provide first aid. Automatically generated eCalls may also lead to silent calls, if driver and passengers are still able to leave the car after the accident, as it is generally recommended by the police. Of course silent calls may also occur, if persons in the car are seriously injured and unable to communicate. People from outside the vehicle (e.g. first aiders) may communicate with the PSAP side or among each other. Communication in eCall scenarios is incomparable to a regular hands-free communication from the driver's position. Thus, the transmitted background noise, with or without additional speech ("noise-only" case or "noise and speech" case), plays an important role.

The typical signal processing in eCall systems is indicated in **figure 1**. In the microphone path (lower transmission path in **figure 1**) the noise reduction signal processing ("NR") is typically tuned to provide a strong and significant noise reduction. This results from technical solutions available today, as the algorithms in use today for in-vehicle systems (IVS) are typically tuned for regular hands-free communication. These algorithms need to be configured differently compared to regular hands-free communication [1], [2], [3] in order to avoid the problem of silent calls. Silent call tests are not covered in eCall specifications today [4].

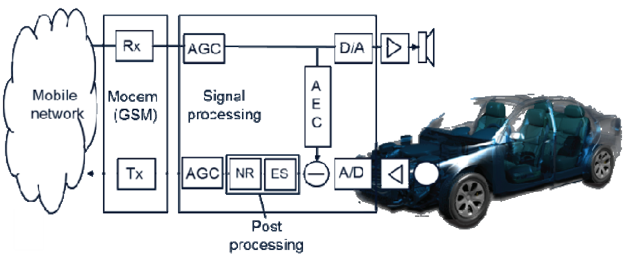


Figure 1: Typical hands-free signal processing in IVS

A listening test (LOT) was carried out in order to identify the important and necessary acoustic characteristics in a transmitted background noise scenario for certain identification. The results are needed to derive appropriate analyses and limits for IVS laboratory tests and tuning and as suggestion for standardization [5].

Listening test design

Various background noise scenarios were used including everyday live situations (recordings from train station, pubs, …), eCall related scenarios with a vehicle with open windows parked on a motorway or on a quiet street with passing vehicles. These traffic scenarios were partly combined with everyday live conversations or with conversations about emergency cases ("call an ambulance").

The background noise scenarios were played back in a driving simulator (Daimler E-type vehicle) equipped with a noise simulation system. Different eCall modules (ECM), after market hands-free devices in conjunction with mobile phones (HFT-M) and mobile phones operated in handheld hands-free mode were installed in the vehicle. The devices were connected to a mobile network simulator (3G mode, 12.2 kbit/s EFR). The setup is shown in **figure 2**.

The uplink signal was then level modified for some devices in order to cover a wide level range. The listening situation on the PSAP side (typical use of monaural headsets) was simulated by applying IRS filter and adjusting the Receiving Loudness Rating (RLR) to 2 dB. The recordings led to a final set of 62 noise files to be judged in the listening test.

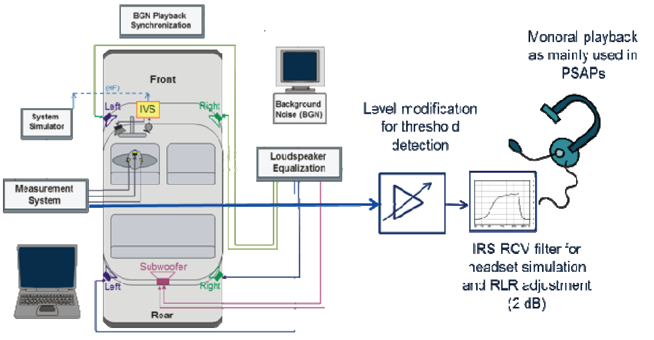


Figure 2: Recording procedure for LOT

24 naïve test persons participated in the test. They were informed to anticipate the situation in a PSAP workplace. Three parameters were assessed for each noise file:

The recognizability of the noise environment (designated as "listening effort" in the following analysis, being aware that the parameter does not represent the listening effort in the traditional sense)

- certainly recognizable (5)
- in general recognizable (4)
- recognizable only with difficulties (3)
- hardly recognizable (2)
- not recognizable (1)

The decision about the noise environment: The present call is
- an emergency call
- a false alarm
- I am not sure

In the noise scenario verification, test subjects were asked to try to identify the noise environment which they assumed they were called from. This was later verified by the test operator as

- correctly identified
- situation identified
- not identified
- misidentified

Analysis of LOT results

The test results for the recognizability (in **figure 3** designated as "listening effort") are sorted with decreasing effort from left to right. The listening examples cover a wide range from 1.2 MOS up to 4.5 MOS. The confidence interval on a 95 % level is in the range of 0.5 MOS for all examples.

The results for the decision about the noise environment (emergency case ▲, faults alarm **x**, "not sure" **-**) are given in **figure 4**. The results are sorted in the same way from left to right with decreasing listening effort. The results also indicate that a low listening effort (high MOS scores in **figure 3**) led to unambiguous decisions and, in particular, a lower uncertainty for the test persons. Vice versa the uncertainty ("not sure" **-**) is high for high listening effort to recognize the scenario.

The verification of correctly identified scenarios is analyzed in **figure 5**. Again the results correlate to the recognizability scores (**figure 3**) and the decision scores (**figure 4**). The correctly identified scenarios (•) increase from approximately 20 % up to 100 % with decreasing listening effort.

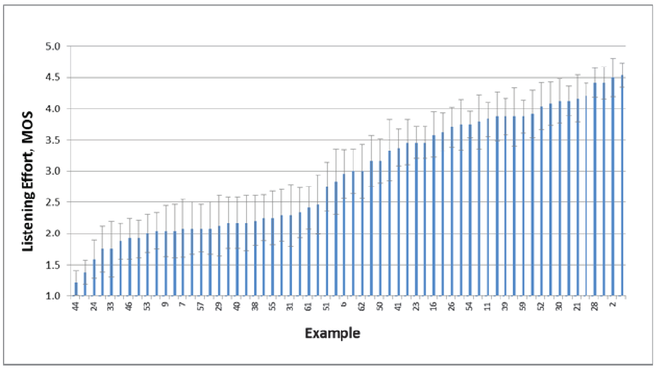


Figure 3: Recognizability ("Listening effort")

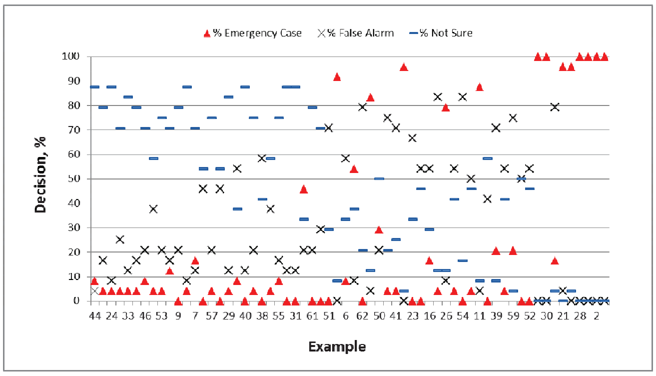


Figure 4: Decision rating

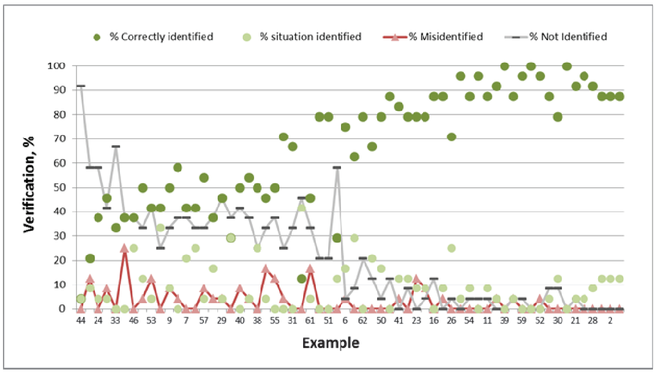


Figure 5: Verification of correctly identified scenarios

The auditory results also serve as a basis to derive instrumental tests to verify the transparency of uplink transmission. In a first step, a simplified test method can be derived from the listening test for the "noise-only" scenarios. The suggested limits shall ensure that the transmitted background noise signal in uplink, picked up by the microphone of the IVS, processed through the implemented signal processing (like noise reduction, see **fig. 1**), and GSM coded and decoded can certainly be identified on the far end side.

In order to provide an overview about the ratings for "noise-only" scenarios in the LOT, **figure 6** compares the LOT parameters in one diagram. The bars again represent the listening effort; the percentage of correctly identified scenarios is indicated by the green dots. In general, high listening effort ratings (corresponding to low effort) correlate to a high percentage of correctly identified scenarios.

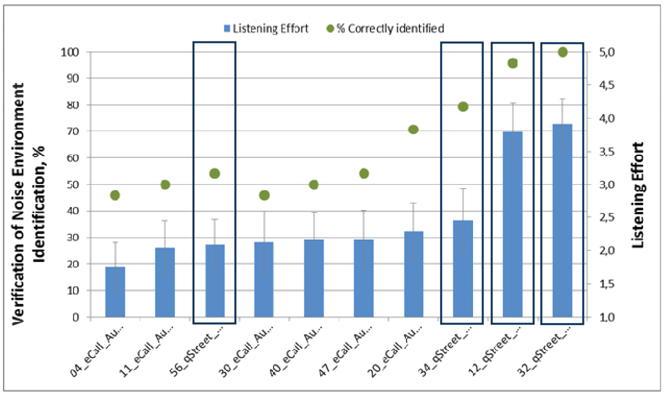


Figure 6: Results of listening test (noise-only examples, "eCall" and "quiet street" scenario)

The "quiet street" noise samples, which are used for the simplified test method, are indicated as example 56, 34, 12 and 32 on the x-axis (see rectangles). This noise sequence consists of a recording in a vehicle with four open windows on a relatively quiet street with one car passing-by. Consequently, the level vs. time of the sound event is relatively low, then increasing during the passing-by sequence and decreasing again. This is an appropriate test sequence for noise reduction algorithms as they typically more strongly attenuate lower noise parts but may erroneously detect an increasing noise (caused by a passing-by vehicle) as speech like and transmit this part transparently. **Table 1** gives the numerical numbers for these examples.

LOT example	56	34	12	32
Listening effort [MOS]	2.1	2.5	3.8	3.9
Decision ["not sure", %]	75.0	70.8	45.8	50.0
Verification ["correct", %]	54.2	79.2	95.8	100.0

Table 1: Numerical results ("quiet street" scenarios)

The two examples 12 and 32 (right hand side in **figure 6**) represent an eCall module with disabled noise reduction, but different listening level in the LOT. Vice versa, the two bars shown on the left hand side for the higher listening effort, corresponding to a lower identification rate, represent two listening examples recorded over the same hands-free device with enabled noise reduction, but judged with different levels in the LOT. The noise reduction introduces audible distortions (level changes, attenuation of low level signal path) in the signal, which leads to a higher uncertainty to identify the nature of the noise scenario, especially at higher listening levels (see **table 2**).

Table 2 summarizes the information for the "quiet street" listening examples extracted from **figure 6**. As indicated before, the uplink signal level and the corresponding listening level in the LOT do not directly correlate to a higher verification score (correctly identified scenario). The listening level in dB(A) for example 34 is approximately 10 dB lower compared to example 56, but the verification rate shows that nearly 80% of the test subjects correctly identified the noise scenario in this example. Vice versa, only 54 % correctly identified the listening example 56.

It can be concluded that a transparent noise transmission is essential to identify the nature of the background noise scenario.

LOT example	56	34	12	32
Test device	HFT-M	HFT-M	ECM	ECM
Level in LOT [dB(A)]	54.1	44.1	33.3	43.7
Uplink level [dBm0]	-49.6	-59.6	-70.7	-60.7

Table 2: Description of listening examples (HFT-M: hands-free implementation with linked mobile phone; ECM: eCall module)

The limits to derive the requirements were set as follows:

- Estimated "listening effort" for recognizability of noise environment ≥ 3.0 MOS
- Verification result ≥ 70% correctly identified.

Instrumental testing

Figure 7 shows three analysis curves in the time domain (level vs. time analysis, time constant 125 ms). The upper black curve was analyzed within the driving simulator based on a recording in the car cabin close to the DUT microphone. The level is rather low and increases significantly by more than 20 dB during the passing-by sequence.

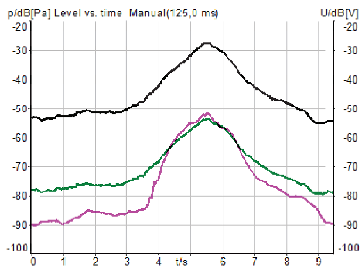
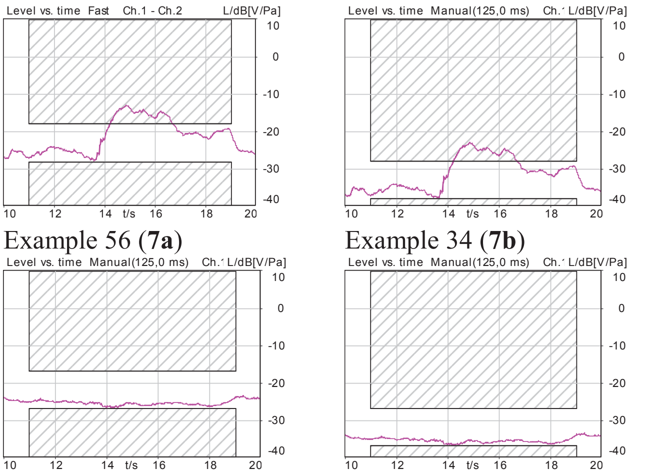


Figure 7: Level vs. time analysis (black: original; magenta: uplink signal for example 34; green: uplink signal for example 32)

The transmitted signal in sending direction of two devices under test (example 34 and 32) is given by the green respectively magenta curve. These signals are analyzed in uplink and the POI of the network simulator. The level offsets compared to the reference curve are caused by the band limitation, coding and scaling (electrical signal vs. acoustical signal for the reference analysis). The green curve indicates a transparent transmission; the level vs. time follows the reference curve recorded in the car cabin indicating that the noise reduction does not attenuate lower or higher signal parts differently. Vice versa, the magenta curve indicates an approximately 12 dB higher attenuation for the lower level parts of the transmitted signal. The high level components are transmitted with the same sensitivity. Thus the level vs. time significantly increases. This makes the sound unnatural and difficult to identify for subjects.

Figure 7 shows the uplink sensitivity analysis for the listening examples 56, 34, 12 and 32 from **figure 6** ("quiet street noise", left to right). The curves are calculated as the level differences between the uplink signal and the signal level within the car cabin acoustically recorded close to the IVS microphone position (time constant 125 ms). The four curves are given together with a 10 dB tolerance vs time.



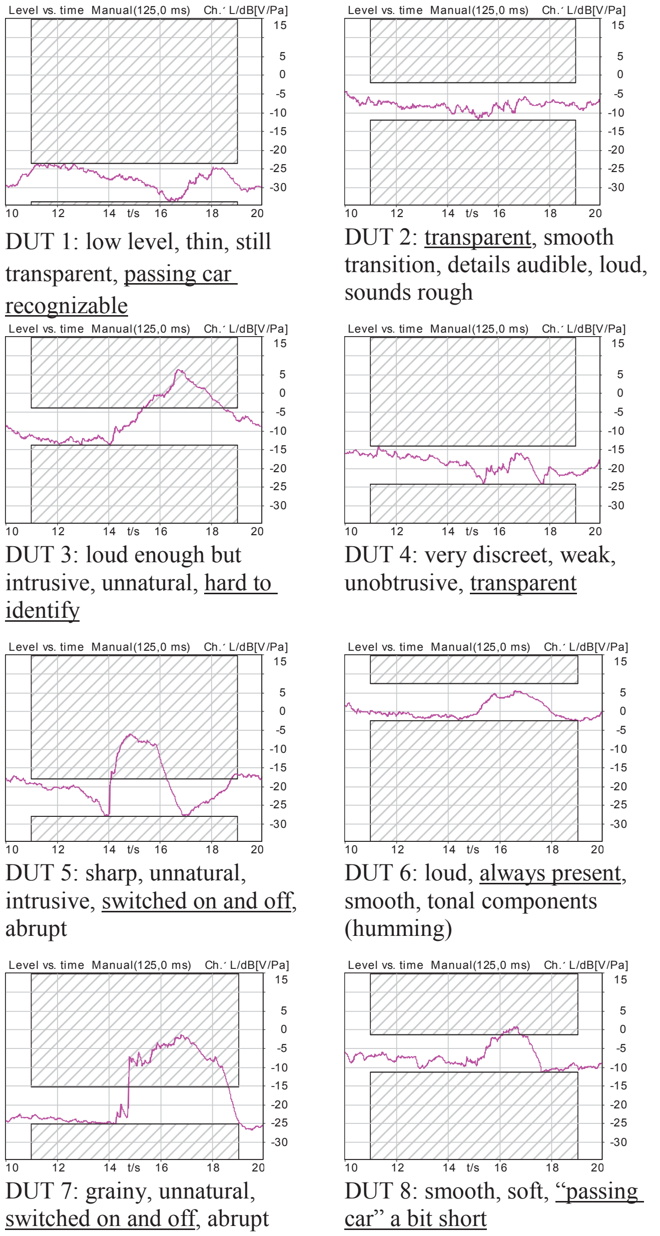
Example 56 (**7a**) — Example 34 (**7b**) — Example 12 (**7c**) — Example 32 (**7d**)

Figure 7: Uplink sensitivity analysis for "quiet street" noise examples

For constant sensitivities vs. time (see **figure 7c** and **7d**) the listening effort to recognize the nature of the background noise situation is represented by high MOS scores of 3.8 and 3.9 (see **table 1**). Nearly all test persons (100 % and 95.8 %) correctly identified the scenario. Furthermore, the decision rate (especially the "not sure" rating) is rather low around 50 %. Vice versa, for sensitivity variations exceeding the 10 dB tolerance in these analyses led to significantly lower MOS scores (2.1 and 2.5) to recognize the nature of the background noise and a higher percentage of uncertainty (75 % and 70.8 %). This corresponds to rather low percentage of correctly identified scenarios (54.2 % and 79.2 %).

As an interesting fact the result in **figure 7a** represents a higher sensitivity vs. time (corresponding to a higher measured level in sending direction and a higher level audible at the simulated PSAP side) compared to **figure 7b.** However, it leads to a lower percentage for correct identification. Strong level modulations are even more confusing for test persons if they try to identify the background noise scenario.

The suggested test method and tolerance has been verified further using a variety of additional eCall implementations. For these devices no formal LOT have been conducted. However, the analyses for further 8 devices (DUT) are shown in **figure 8** together with informal experts' comments. Their findings are listed below the individual analyses.



DUT 1: low level, thin, still transparent, passing car recognizable

DUT 2: transparent, smooth transition, details audible, loud, sounds rough

DUT 3: loud enough but intrusive, unnatural, hard to identify

DUT 4: very discreet, weak, unobtrusive, transparent

DUT 5: sharp, unnatural, intrusive, switched on and off, abrupt

DUT 6: loud, always present, smooth, tonal components (humming)

DUT 7: grainy, unnatural, switched on and off, abrupt

DUT 8: smooth, soft, "passing car" a bit short

Figure 8: Uplink sensitivity analysis and experts ratings

The experts confirm in general the validity of this simplified method to exclude background noise handling in silent calls, which may lead to unnatural and hard to identify background noise situations.

References

[1] Speech Quality and Tuning Aspects for eCall Implementations, F. Kettler, R. Serafimov, DAGA 2014, March 2014, Oldenburg, Germany

[2] ITU-T Recommendation P.1100, Narrow-band hands-free communication in motor vehicles (03/11)

[3] ITU-T Recommendation P.1110 Wideband hands-free communication in motor vehicles (12/09)

[4] GOST ERA-GLONASS Specification, In-vehicle emergency call system compliant test methods for quality of speaker phone in a vehicle, 2012

[5] "Silent Calls": Detection and evaluation of background noise scenarios in emergency call situations, Contribution to Q.4 Rapporteurs Meeting Dec. 10-12, 2014, Herzogenrath, Germany