Speech Quality “Quick Check” for VoIP Terminals

Frank Kettler, Frank Rosenberger, Hans Wilhelm Gierlich

HEAD acoustics GmbH, 52134 Herzogenrath, Germany, Email: frank.kettler@head-acoustics.de

Motivation

Standardized tests for IP phones are not suitable today to guaranty a sufficient speech quality. TIA 810A [1] defines basic requirements for parameters like frequency response, loudness rating and echo. These tests are limited neither taking into account IP specific impairments (packet loss, jitter) nor covering conversational aspects like double talk performance. ETSI, the **E**uropean **T**elecommunications **S**tandards **I**nstitute organized speech quality test events for VoIP equipment [2], [3]. These tests cover all conversational aspects and IP specific impairments providing a detailed quality overview. Due to the number of test conditions and parameters the test program fills a complete testing day. What is obviously missing are speech quality tests providing a quick quality overview but still covering all conversational aspects including the important double talk performance. A comparison test was carried out with 5 IP phones. These speech quality tests (“Quick Check“) were designed

- to provide a comprehensive overview about the current speech quality of each device and
- to measure additional parameters providing important information in order to improve quality.

Test Setup and Parameters

The tests concentrate on the acoustical quality of IP phones in handset mode using the G.711 speech coder. The following measurements were implemented in the speech quality “Quick Check” for these VoIP terminals:

- loudness rating (SLR), frequency responses and MOS-LQO (acc. to ITU-T P.800.1 [4]) in sending direction using the TOSQA2001 algorithm [5]
- receiving loudness ratings (RLR), frequency responses and MOS-LQO, pressure force dependent,
- echo and double talk performance tests and
- recordings using real speech

In order to reproduce the acoustical leakage between handset and human ear a HATS (ITU-T P.58 [4]) was used equipped with a type 3.4 artificial ear (ITU-T P.57 [4]). Note that the parameters in receiving direction were measured with pressure forces of 2 N, 8 N and 13 N. In addition to the 5 IP phones reference measurements were carried out with a standard ISDN phone and a GSM mobile. The ISDN phone mounted to the HATS is shown in figure 1.

Fig. 1: HATS HMS II.3 with mounted ISDN phone)

Test results

The MOS-LQO values in sending direction are given in table 1 together with the sending loudness ratings (SLR). The results for the 5 IP phones under test differ slightly between 3.9 and 4.2 MOS-LQO.

	A	B	C	D	E	ISDN	GSM
MOS-LQO	**3.9**	**4.2**	**4.2**	**4.2**	**4.0**	**4.0**	**3.4**
SLR [dB]	**10.3**	**6.2**	**8.6**	**7.8**	**10.2**	**8.7**	**11.4**

Table 1: MOS-LQO in sending direction and SLR

The frequency responses are comparable for the 5 IP phones. The MOS-LQO differences are mainly caused by slight differences in sending loudness ratings. In order to model a subjective listening test (ITU-T P.800 [4]) TOSQA2001 equalizes level differences. The speech signals are amplified for those IP phones providing a higher SLR (phone A and E) compared to the phones with higher sensitivities (B, C and D). This also increases the noise level and contributes to a slightly lower MOS-LQO value.

Table 2 shows the MOS-LQO values measured in receiving direction for IP phone A, B, C and D applying different pressure forces of 2 N, 8 N and 13 N between the handset and the artificial ear.

	A	B	C	D	ISDN	GSM
MOS-LQO 2 N	**2.7**	**2.2**	**3.4**	**2.7**	**2.4**	**3.2**
MOS-LQO 8 N	**3.9**	**3.1**	**4.1**	**3.4**	**3.4**	**3.5**
MOS-LQO 13 N	**4.1**	**3.8**	**4.1**	**3.8**	**3.9**	**3.5**

Table 2: MOS-LQO in receiving direction

Due to the different leakage sensitivity of the handsets the MOS-LQO differences are higher for the lower pressure force of 2 N (MOS-LQO between 2.2 and 3.4) compared to the higher pressure forces of 8 N and 13 N (MOS-LQO between 3.8 and 4.1). The clearest differences can be seen between the two IP phones B and C. The low MOS-LQO values for the low pressure forces are surprisingly taking into account that the same analysis for the GSM mobile leads to MOS-LQO value between 3.2 (2 N) and 3.5 (8 N and 13 N). This MOS-LQO value of 3.2 for the 2 N pressure force is higher compared to the IP phones A, B and C.

Further analysis point out that the dominant quality parameter in receiving direction is the frequency response. Figure 2, 3 and 4 represent the measured curves for the IP phones A, B and C. The different colors indicate the pressure forces between handset and artificial ear (green: 2N, red: 8N and light blue 13 N).

The handset of IP phone A shows a distinct high pass characteristic for the 2 N pressure force (fig. 2). The MOS-LQO value is determined to 2.7 (see table 1). A significant improvement can be measured for 8 N whereas higher pressure forces (13 N) have a minor influence.

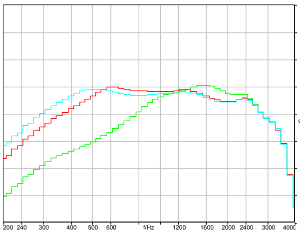

Fig. 2: IP phone A

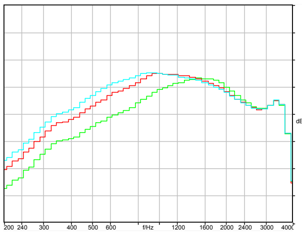

Fig. 3: IP phone B

A distinct high pass characteristic is also measured for IP phone B (fig. 4). The higher pressure forces improve the sensitivity in the lower frequency range but the frequencies below 800 Hz are not sufficiently coupled. The transmitted speech sounds "thin", accordingly the MOS-LQO values are low compared to the other IP phones (see table 2). The measured curves for the IP phone C are better balanced around a mid-frequency of 1 kHz. This leads to higher MOS-LQO values.

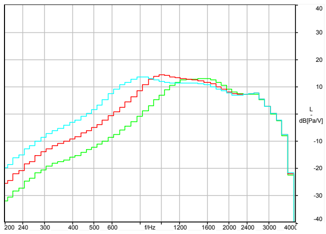

Fig. 4: IP phone C

The terminal coupling loss values (TCL_w) are given in table 3. Note that TIA 810 tests require a TCL_w value of at least 52 dB for IP phones [1]. This can typically be met only by implementing additional signal processing (echo suppression). Consequently this influences double talk performance which is not covered by the TIA 810A tests.

	A	B	C	D	E	ISDN	GSM
TCL_w [dB]	**42.0**	**49.7**	**42.5**	**48.4**	**47.1**	**48.7**	**41.7**

Table 3: Echo measurement results (TCL_w value)

None of the devices under test fulfill these 52 dB. Moreover these results indicate two different kind of implementations. The two IP phones A and C provide a significantly lower echo attenuation (around 42 dB) compared to the implementations B, D und E. The reason can be found in the implemented echo suppression: Both IP phones A and C do not have any additional signal processing implemented. The TCL_w value is lower compared to the other solutions B, D, and E providing echo suppression to attenuate the sending direction (reduce the coupled echo). The influence is obvious, the TCL_w values are significantly higher.

On the other hand the lower sensitivity in sending direction also influences the double talk performance. This can be shown by the double talk measurement results in figure 4, 5 and 6 comparing the implementation in IP phone C, B and D. This analysis characterizes the double talk performance of hands-free terminals (ITU-T P.340 [4]). A level variation of approximately 9 dB can be measured for the IP phone B (dotted line in fig. 5). This leads to a "type 2c" characterization [4]. No level variation occur for the IP phones C and D.

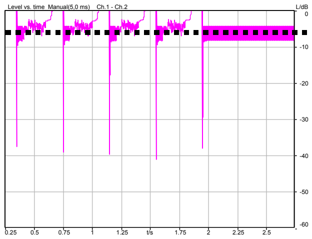

Fig 4: IP phone C

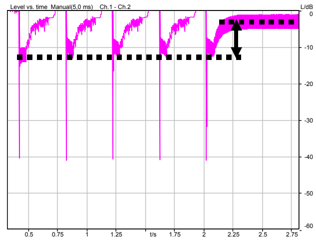

Fig 5: IP phone B

An interesting implementation is represented by the result given in fig. 6 for the IP phone D. The double talk capability is not impaired ("type 1") although the echo attenuation is high (48.4 dB TCL_w, table 3).

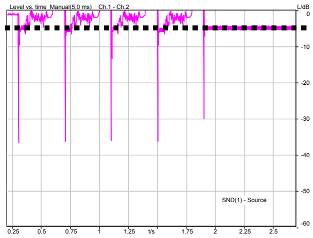

Fig. 6: IP phone D

Summary

Fig. 7 summarizes the results of the different quality aspects. The MOS-LQO values scale the axes from 1 ("bad") to 5 ("excellent"). The scaling for the TCLw axis is chosen between 35 dB and 55 dB. The "DT type" axis indicates the characterization given in [4]. The overview indicates the highest acoustical quality in terms of MOS-LQO values in sending and receiving direction for IP phone C (magenta). IP phone D provides a low echo level without impairing double talk performance (blue).

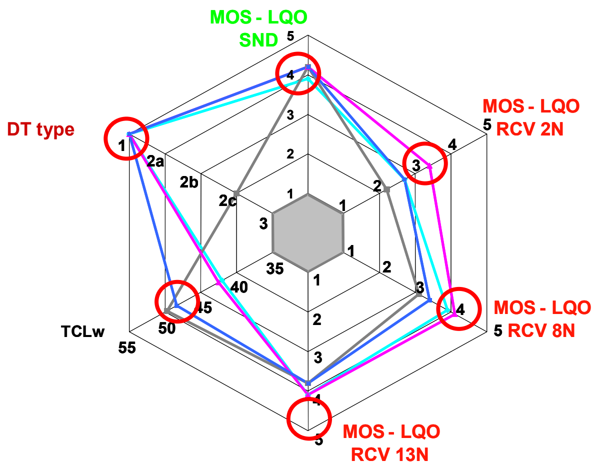

Fig. 7: "Quick Check" Results (A: light blue, B: grey, C: magenta, D: blue)

References

[1] ANSI / TIA / EIA-810-A, Transmission Requirements for Narrowband Voice over IP and Voice over PCM Digital Wireline Telephones
[2] 1st ETSI SQTE, Anonymized Test Report, Oct. 2000
[3] 2nd ETSI SQTE, Anonymized Test Report, April 2002
[4] ITU-T Recommendations P.Series
[5] ITU-T Contr. 12-20-E, Results of Objective Speech Quality Assessment Including Receiving Terminals Using the Advanced TOSQA2001