INTERNATIONAL ORGANISATION FOR STANDARDISATION

ORGANISATION INTERNATIONALE DE NORMALISATION

ISO/IEC JTC1/SC29/WG11

CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11

MPEG2008/N10032

July 2008, Hannover, Germany

Source: **MPEG audio subgroup**
Title: **Report on the Verification Test of MPEG-4 Enhanced Low Delay AAC**
Status: **Approved**

Executive summary

This document reports on the verification tests of the ISO/IEC 14496-3/AMD9 Enhanced Low Delay AAC standard, a new low delay audio coding scheme for high quality conversational applications. This standard, referred to as AAC ELD, was completed by MPEG in October 2007 to update the earlier MPEG-4 Low Delay AAC (AAC LD) codec with more recently developed MPEG-4 Audio tools that improve coding performance and enable reduction of algorithmic delay.

The verification tests of the AAC ELD codec comprised in total six listening test experiments, each of which were conducted at least by two independent laboratories using the MUSHRA test methodology specified in ITU-R Recommendation BS.1534. Five companies contributed to this exercise as listening test laboratories with a grand total of 152 subjects. The experiments cover the bit rates 24, 32, 48, and 64 kbps in single channel operation.

The verification tests were divided into two sets of experiments. The technology-driven test was designed to assess the performance of selected AAC ELD codec configurations with respect to the current state of the art MPEG-4 audio codecs, MPEG-4 High Efficiency AAC (HE AAC) and AAC LD. The ITU-T Recommendation G.722.1 Annex C, briefly referred to as G.722.1-C, was used as an additional reference. Secondly, the application-driven set of tests was designed to evaluate AAC ELD codec configurations that are foreseen attractive for high quality communication. These tests included clean speech, reverberant clean speech, and speech with office background noise test conditions complemented with music and mixed content conditions. The current state of the art high quality communication codecs AAC LD and G.722.1-C were used as references.

In all experiments, the AAC ELD codec showed good performance with the following overarching conclusions:

- In the experiment conducted at 24 kbps, the performance of the AAC ELD codec at this bit rate was, at the 95% level of significance, comparable to that of the G.722.1-C codec at a bit rate of 32 kbps at both test sites.

- The mean scores for the AAC ELD codec at a bit rate of 32 kbps were in the Good or Excellent range on the MUSHRA scale, and consistently, at the 95% level of significance, higher than those of the AAC LD codec and the G.722.1-C codec at the same bit rate in all experiments. The AAC ELD codec attained this performance gain with over 6 ms lower algorithmic delay compared to these two reference codecs that have an algorithmic delay of 40 ms.

- For a bit rate of 48 kbps, the AAC ELD codec was the only system under test that consistently had the mean score in the Excellent range of the MUSHRA scale in all experiments across all listening test laboratories. At this bit rate, both the AAC ELD and AAC LD codecs consistently outperformed, at the 95% level of significance, the G.722.1-C codec in all experiments.

- At a bit rate of 64 kbps, the AAC ELD codec had, at the 95% level of significance, the mean score equal to that of AAC LD, but has 5 ms (or 25%) lower algorithmic delay than this reference codec, which has an algorithmic delay of 20 ms.

As a further observation, the AAC LD technology showed also good performance throughout the experiments. The mean score of the AAC LD codec was, at the 95% level of significance, equal or better than that of the G.722.1-C codec at the same bit rate in all six experiments across test laboratories with the exception of one experiment conducted in one laboratory.

© ISO/IEC 2008 – All rights reserved

Table of Contents

1 Introduction ... 4

2 Listening Test Experiments ... 4
2.1 Overview ... 4
2.2 Codecs Under Test ... 5
2.2.1 MPEG-4 Enhanced Low Delay AAC ... 5
2.2.2 Algorithmic Delay ... 5
2.2.3 Bitreservoir ... 6
2.3 Technology-driven Test ... 7
2.4 Application-driven Test: High Quality Communication ... 8

3 Test Sites ... 10

4 Test Results ... 10
4.1 Introduction ... 10
4.2 Technology-driven Test ... 11
4.2.1 Experiment T1, Low Bit Rates ... 12
4.2.2 Experiment T2, High Bit Rates ... 13
4.3 Application-driven Test ... 14
4.3.1 Experiment A1 Speech ... 16
4.3.2 Experiment A1 Music ... 16
4.3.3 Experiment A2 Speech ... 17
4.3.4 Experiment A2 Music ... 18

5 Conclusions ... 18

6 Acknowledgements ... 19

7 References ... 19

A Test Methodology ... 21
A.1 MUSHRA test ... 21
A.2 Training of Subjects ... 21
A.3 Processing Steps for 3.5 kHz and 7 kHz Anchors ... 22
A.4 Statistical Analysis ... 22
A.5 Post-screening ... 23
A.6 Test Items ... 24
A.6.1 List of Test Items ... 24
A.6.2 Processing of Speech Items ... 25
A.6.3 Pre- and Post-processing Steps for G.722.1 Annex C ... 27

B Detailed Results of All Test Sites ... 28
B.1 Technology-driven Test, Experiments T1 and T2 ... 28
B.2 Application-driven Test ... 30
B.2.1 Experiment A1: High Quality, Low Bit Rate Communication ... 30
B.2.2 Experiment A2: High Quality Communication ... 33

C Low Power SBR ... 36
C.1 Experiment Overview ... 36
C.2 Test Results ... 36

© ISO/IEC 2008 – All rights reserved

1 Introduction

MPEG has issued several successful audio standards. The International Standards ISO/IEC 11172-3 (MPEG-1 Audio), ISO/IEC 13818-3 (MPEG-2 Audio) and 13818-7 (MPEG-2 Advanced Audio Coding, AAC) were issued in 1992, 1994, and 1997, respectively. These standards were followed by the International Standards ISO/IEC 14496-3 (MPEG-4 Audio Version 1) and ISO/IEC 14496-3 / AMD1 (MPEG-4 Audio Version 2) in 1999 and 2000, respectively. The latter of these standards specifies a low delay derivate of the MPEG-4 Error Resilience AAC coding scheme (AAC LD) for bi-directional communication applications. In 2003, MPEG amended the MPEG-4 Audio specification ISO/IEC 14496-3:2001 / AMD1 with the Spectral Band Replication (SBR) tool that can be combined with MPEG-4 AAC for high quality audio coding at low bit rates. This specification resulted in the standardization of the ISO/IEC 14496-3 High Efficiency AAC (HE AAC) Profile.

In October 2006, at its 78th meeting in Hangzhou, MPEG initiated the development of the standard ISO/IEC 14496-3 / AMD9 entitled Enhanced Low Delay AAC to update the MPEG-4 low delay audio coding scheme by integrating a low delay version of the SBR tool in the existing technology. The technical work on this standard, briefly referred to as AAC ELD, was finalized in the 82nd MPEG meeting held in Shenzhen in October 2007. Finally, the standard was published in July 2008.

This document reports on the final verification tests of the MPEG-4 Enhanced Low Delay AAC. In addition, this document reports the results of the listening tests conducted during the standardization process to characterize the performance of the low power mode of the SBR tool used in the decoder.

This report is organized as follows. Section 2 gives an overview on the listening test experiments conducted for this verification test. Section 3 lists the allocation of the experiments between the participating listening test sites. Section 4 presents the results of the listening test experiments based on post-screened data that is pooled across listening test sites. Furthermore, the complete documentation on the test methodology regarding statistical analysis, post-screening rules and test items is given in Appendix A. The detailed results for the individual test sites are provided in Appendix B followed by the results of the listening tests associated with the low power mode of the SBR tool for use in the AAC ELD decoder in Appendix C.

2 Listening Test Experiments

2.1 Overview

This document reports on the verification tests of the AAC ELD codec organized by the Audio Subgroup of MPEG. These verification tests comprise two sets of listening tests, technology-driven and application-driven, arranged in a total of six experiments that are conducted using the MUSHRA test methodology specified in ITU-R Recommendation BS.1534 [1]. The experiments cover the bit rates 24, 32, 48, and 64 kbps in single channel operation with a total of 32 items from speech, mixed content, and music categories. Each experiment was conducted at least by two independent listening test laboratories. Five companies contributed to this exercise as listening test laboratories with a grand total of 152 subjects.

The technology-driven test was designed to characterize the performance of selected AAC ELD codec configurations across a broad bit rate range with a well-established set of critical test items using two current state of the art MPEG-4 audio codecs, HE AAC and AAC LD, as references. The ITU-T Recommendation G.722.1 Annex C, referred to as G.722.1-C, was added as an additional reference. The AAC LD and G.722.1-C codecs are broadly-used, state of the art communication codecs that operate at bit rates, an algorithmic delay and an audio bandwidth similar to the actual codec under test. To maintain a manageable size for individual experiments, the technology-driven test is divided into two experiments according to the bit rate.

The application-driven test was designed to evaluate some AAC ELD configurations that are foreseen attractive for high quality communication and professional broadcasting applications. The reference codecs and the test items were chosen keeping the targeted use scenario in mind. The test was

© ISO/IEC 2008 – All rights reserved

divided into four experiments based on the content type and the bit rate range to make the size of each experiment manageable. The speech-focused experiments cover clean speech, reverberant clean speech, and reverberant speech with office background noise conditions. The music-focused experiments cover music and mixed content. The AAC LD codec and the G.722.1-C codec were included as references.

2.2 Codecs Under Test

The following codecs were included in the tests:

- AAC ELD (implementation provided by Fraunhofer IIS)
- AAC LD (implementation provided by Fraunhofer IIS)
- HE AAC (implementation provided by Dolby Sweden)
- G.722.1-C (implementation available in ITU-T Recommendations)

The AAC LD and G.722.1-C codecs are current state of the art high quality communication codecs that are broadly deployed especially in videoconferencing systems and operate with bit rates, algorithmic delay and audio bandwidth similar to the actual codec under test. Therefore, these two codecs are used as references throughout the tests. The HE AAC codec is used as a reference in the technology-driven test for the evaluation of the performance of the low delay SBR tool.

According to the MUSHRA test methodology, low-pass filtered versions of the test items with cut-off frequencies at 3500 Hz (LP35) and 7000 Hz (LP70) were included in all experiments. In addition, a copy of the item under test was included as the hidden reference (HR).

2.2.1 MPEG-4 Enhanced Low Delay AAC

The ISO/IEC 14496-3 / AMD9 Enhanced Low Delay AAC standard addresses the task of coding generic audio content for bi-directional communication applications. The algorithmic delay of the codec varies between 15.0 ms and 36.3 ms depending on the configuration. This standard allows a framing of 10 ms or 20 ms. The codec is signaled by the MPEG-4 Audio Object Type (AOT) 39.

Compared to the preceding MPEG-4 Low Delay AAC standard, the new AAC ELD technology allows a further 25% delay reduction of the core codec to 15 ms from 20 ms by utilizing a low delay window in the filter bank. The new standard can also make use of a novel low delay SBR tool, which is derived from MPEG-4 SBR and is optimized for low algorithmic delay. As a result, the low delay SBR tool enables low bit rate operation of AAC ELD while maintaining high audio quality. Both of these benefits provided by the new standard are explored and characterized in the verification tests.

2.2.2 Algorithmic Delay

The algorithmic delay is defined as the latency caused by the encoding and decoding algorithms in a real time system. Additional latency may result from computational, buffering or network transmission delays.

The algorithmic delay of AAC ELD, denoted by d_{ELD} and expressed in seconds, has the following relation to the frame length in samples and the sampling rate of the codec:

$$d_{\mathrm{ELD}} = \frac{1.5 \cdot \text{frame length} + 64}{\text{sampling rate}}. \qquad (1)$$

It should be noted that the frame length of AAC ELD with spectral band replication is twice that of the core. The AAC ELD standard supports frame lengths of 2×480 samples or 2×512 samples. At a

© ISO/IEC 2008 – All rights reserved

sampling frequency of 48 kHz, these frame lengths translate respectively to an algorithmic delay of 31.3 ms or 33.3 ms.

At high bit rates, the SBR tool can be disabled while maintaining high audio quality in order to reduce the algorithmic delay $d_{\text{ELD-S}}$ of this configuration to

$$d_{\text{ELD-S}} = \frac{1.5 \cdot \text{frame length}}{\text{sampling rate}}, \tag{2}$$

where the frame length is now that of the core, either 480 or 512 samples. In this configuration, the algorithmic delay of AAC ELD is respectively reduced down to 15.0 ms or 16.0 ms assuming a sampling rate of 48 kHz.

The algorithmic delay of the AAC LD and G.722.1-C codecs in seconds, denoted respectively by d_{LD} and $d_{1\text{C}}$, has the following relation to the frame length and sampling rate of the codec:

$$d_{\text{LD}} = d_{1\text{C}} = \frac{2 \cdot \text{frame length}}{\text{sampling rate}}. \tag{3}$$

For the HE AAC codec, the algorithmic delay becomes

$$d_{\text{HE}} = \frac{2 \cdot \text{frame length} + 576 + 576 + 384}{\text{sampling rate}}, \tag{4}$$

where the frame length of the codec is twice that of the core.

2.2.3 Bitreservoir

The AAC framework can make use of a bitreservoir in order to shift bits between subsequent frames. In packet-based transmission of bitstreams, the usage of a bitreservoir does not add delay to the transmission chain.

Table 1 reports the effective size of the bitreservoir used for all AAC LD and AAC ELD coding systems in each experiments. The effective size of the bitreservoir is defined here as the maximum size of an access unit (or the compressed representation of one frame) over all items of the experiments less the average bit rate per frame. The effective size of the bitreservoir also determines the additional delay which would be introduced for transmissions via constant bit rate channels such as ISDN. This additional delay is not included in the algorithmic delay presented elsewhere in this report because of the focus on use scenarios over packet-based networks. For the delay estimation in Table 1, it is assumed that the transmission channel has a capacity equal to the bit rate of the codec.

© ISO/IEC 2008 – All rights reserved

Table 1: Effective size of the bitreservoir and the additional delay caused by the bitresevoir in constant bit rate operation.

Experiment	Codec ID	Max. size of access unit in bytes	Average bit rate in bits / frame	Effective bitreservoir size in bits	Additional delay in ms
T1	ELD-24	97	557.3	219	9.11
	ELD-32	116	682.7	245	7.67
	LD-32	118	682.7	261	8.17
T2	ELD-48	183	1024.0	440	9.17
	ELD-64	230	1365.3	475	7.42
	ELD-64-S	119	682.7	269	4.21
	LD-48	132	768.0	288	6.00
	LD-64	119	682.7	269	4.21
A1 Music	ELD-32	109	640.0	232	7.25
	ELD-48	157	960.0	296	6.17
	LD-32	107	640.0	216	6.75
	LD-48	121	720.0	248	5.17
A1 Speech	ELD-32	105	640.0	200	6.25
	ELD-48	162	960.0	336	7.00
	LD-32	110	640.0	240	7.50
	LD-48	120	720.0	240	5.00
A2 Music	ELD-48	157	960.0	296	6.17
	ELD-64-S	110	640.0	240	3.75
	LD-48	121	720.0	248	5.17
	LD-64	111	640.0	248	3.88
A2 Speech	ELD-48	162	960.0	336	7.00
	ELD-64-S	107	640.0	216	3.38
	LD-48	120	720.0	240	5.00
	LD-64	106	640.0	208	3.25

2.3 Technology-driven Test

This test was designed similarly to the MPEG verification test conducted for the evaluation of the SBR tool in the HE AAC Profile [4]. The MPEG Audio Subgroup uses a balanced set of critical test items that exhibit diverse common by challenging signal characteristics for the evaluation of new technologies. This item set is described in detail in Appendix B.

The test consists of two experiments, Experiment T1 that covers bit rates 24 and 32 kbps and Experiment T2 that covers bit rates 48 and 64 kbps as detailed in Table 2 andTable 3, respectively. At high bit rates, the SBR tool of the AAC ELD codec can be disabled to further reduce the algorithmic delay while maintaining high audio quality. The performance of this high bit rate configuration is compared with the nominal configuration of the AAC ELD codec that has the SBR tool enabled in Experiment T2.

© ISO/IEC 2008 – All rights reserved

Table 2: Codecs in the technology-driven low bit rate experiment (Experiment T1).

Codec ID	Codec	Bit rate	Sampling frequency	Frame length in samples	Delay	Audio bandwidth
1C-24	G.722.1-C	24 kbps	32 kHz	640	40.0 ms	14.000 Hz
1C-32	G.722.1-C	32 kbps	32 kHz	640	40.0 ms	14.000 Hz
LD-32	AAC LD	32 kbps	24 kHz	512	42.7 ms	11.600 Hz
ELD-24	AAC ELD	24 kbps	22/44 kHz (1)	2×512	36.3 ms	12.750 Hz
ELD-32	AAC ELD	32 kbps	24/48 kHz (1)	2×512	33.3 ms	14.250 Hz
HE-24	HE AAC	24 kbps	24/48 kHz (1)	2×1024	129.3 ms	15.375 Hz

(1) The MPEG-4 Enhanced Low Delay AAC (ELD-24 and ELD-32) and the MPEG-4 High Efficiency AAC (HE-24) under test are dual rate systems with the underlying AAC codec operating at half the sampling rate of the SBR tool.

Table 3: Codecs in the technology-driven high bit rate experiment (Experiment T2).

Codec ID	Codec	Bit rate	Sampling frequency	Frame length in samples	Delay	Audio bandwidth
1C-48	G.722.1-C	48 kbps	32 kHz	640	40.0 ms	14.000 Hz
LD-48	AAC LD	48 kbps	32 kHz	512	32.0 ms	14.100 Hz
LD-64	AAC LD	64 kbps	48 kHz	512	21.3 ms	14.500 Hz
ELD-48	AAC ELD	48 kbps	24/48 kHz (1)	2×512	33.3 ms	16.875 Hz
ELD-64	AAC ELD	64 kbps	24/48 kHz (1)	2×512	33.3 ms	20.250 Hz
ELD-64-S	AAC ELD	64 kbps	48 kHz (2)	512	16.0 ms	14.500 Hz
HE-32	HE AAC	32 kbps	24/48 kHz (1)	2×1024	129.3 ms	16.875 Hz

(1) The MPEG-4 Enhanced Low Delay AAC (ELD-48 and ELD-64) and the MPEG-4 High Efficiency AAC (HE-32) under test are dual rate systems with the underlying AAC codec operating at half the sampling rate of the SBR tool.
(2) The SBR tool of this MPEG-4 Enhanced Low Delay AAC system under test (ELD-64-S) is disabled.

2.4 Application-driven Test: High Quality Communication

This test focuses on the performance of the AAC ELD codec in high quality communication scenarios, for which this technology is primarily designed. The configurations of AAC ELD, the reference codecs and the test items are chosen for this application scenario. The test is divided into four experiments based on the content type and the bit rate range to make the size of each experiment manageable. The speech-focused experiments cover clean speech, reverberant clean speech, and reverberant speech with office background noise conditions. The music-focused experiments cover music and mixed content.

The speech-focused experiments consists of

- 3 clean speech items
- 3 reverberant speech items
- 6 items with reverberant speech, office background noise and interfering talkers

The detailed description of the process for generating these items is given in Section A.6.2. The processing of these items follows the procedure used by ITU-T for the characterization of Recommendation G.722.1 Annex C. The complete list of the test items is given in Appendix A.6.1.

A second set of test items was defined in order to evaluate the performance of the new technology for the transmission of multimedia content. This set addressed applications including the transmissions of general audio content. Examples for such applications are presentations via video conferencing systems or real-time reporter systems (broadcast phone) with which journalists report live from location and interact with their studio counterparts. For these scenarios, an additional set of 10 items was defined. It consists of speech over background music, applause, jingles and contemporary music. The complete list of these test items is given in Appendix A.6.1.

© ISO/IEC 2008 – All rights reserved

Both sets were tested at a low bit rate (Experiments A1 Speech and Music) as well as a high bit rate (Experiments A2 Speech and Music) configuration of the codecs. The configurations of the AAC ELD and AAC LD codecs were chosen to minimize the algorithmic delay. Thus, the smallest frame length option was employed. The codec configurations used in Experiments A1 Speech and A1 Music are detailed in Table 4 and those of Experiments A2 Speech and A2 Music in Table 5.

Table 4: Codecs in the high quality, low bit rate communication scenario (Experiments A1 Speech and A1 Music).

Codec ID	Codec	Bit rate	Sampling frequency	Frame length in samples	Delay	Audio bandwidth
1C-32	G.722.1-C	32 kbps	32 kHz	640	40.0 ms	14.000 Hz
1C-48	G.722.1-C	48 kbps	32 kHz	640	40.0 ms	14.000 Hz
LD-32	AAC LD	32 kbps	24 kHz	480	40.0 ms	11.600 Hz
LD-48	AAC LD	48 kbps	32 kHz	480	30.0 ms	14.100 Hz
ELD-32	AAC ELD	32 kbps	24/48 kHz (1)	2×480	31.3 ms	14.250 Hz
ELD-48	AAC ELD	48 kbps	24/48 kHz (1)	2×480	31.3 ms	16.875 Hz

(1) The MPEG-4 Enhanced Low Delay AAC (ELD-32 and ELD-48) under test is a dual rate system with the underlying AAC codec operating at half the sampling rate of the SBR tool.

Table 5: Codecs in the high quality communication scenario (Experiments A2 Speech and A2 Music).

Codec ID	Codec	Bit rate	Sampling frequency	Frame length in samples	Delay	Audio bandwidth
1C-48	G.722.1-C	48 kbps	32 kHz	640	40.0 ms	14.000 Hz
LD-48	AAC LD	48 kbps	32 kHz	480	30.0 ms	14.100 Hz
LD-64	AAC LD	64 kbps	48 kHz	480	20.0 ms	14.500 Hz
ELD-48	AAC ELD	48 kbps	24/48 kHz (1)	2×480	31.3 ms	16.875 Hz
ELD-64-S	AAC ELD	64 kbps	48 kHz (2)	480	15.0 ms	14.500 Hz

(1) The MPEG-4 Enhanced Low Delay AAC (ELD-48) under test is a dual rate system with the underlying AAC codec operating at half the sampling rate of the SBR tool.
(2) The SBR tool of this MPEG-4 Enhanced Low Delay AAC system under test (ELD-64-S) is disabled.

© ISO/IEC 2008 – All rights reserved

3 Test Sites

Five companies, Dolby, ETRI, Fraunhofer IIS, LGE and Thomson, contributed to this exercise as listening test laboratories. ETRI and LGE reported their results as one site hereafter referred to as ETRI/LGE. Dolby reported their results from two separate test sites, Dolby San Francisco and Dolby Sweden/Germany, which are hereafter referred to as Dolby SF and Dolby, respectively.

Table 6 reports the allocation of the experiments between the test sites and the number of subjects at each test site after the post-screening procedure of Appendix A.5 has been applied.

Table 6: The allocation of the experiments between contributing test sites and the number of subjects after post-screening

Experiment	Dolby	Dolby SF	ETRI / LGE	Fraunhofer IIS (FhG)	Thomson (DTO)	**Total**
A1 Speech	–	–	9	8	10	**27**
A1 Music	12	–	9	8	–	**29**
A2 Speech	–	8	–	8	11	**27**
A2 Music	–	7	–	8	9	**24**
T1	10	–	–	10	–	**20**
T2	9	–	–	9	–	**18**(1)

(1) The total number of listeners in Experiment T1 is smaller than the minimum number of 20 listeners required by the work plan, but considered sufficient for the purposes of this report.

4 Test Results

4.1 Introduction

A statistical analysis and post screening (see Appendix A.4, A.5) was done on the listening test data. The following plots display the mean values (bar) and associated 95% confidence intervals (vertical tick) averaged over all items for every system. Detailed information of the performance of the codecs on a per-item basis is provided by the Excel sheet attached to this document.

The verified data from various test sites was pooled to enable the most accurate observations and therefore to get the narrowest possible 95% confidence intervals on the mean score. The results from the individual test sites are given in Appendix B.

It should be noted that in reporting subjective test results, it is generally agreed that comparisons of results are valid only for conditions assessed within the same experiment. It is not valid, for example, to directly or statistically compare subjective test results for one codec across two bit rates when those results have been obtained from different experiments. In general, this principle will be observed in the subjective test results presented in the following clauses.

© ISO/IEC 2008 – All rights reserved

4.2 Technology-driven Test

This section provides the results of the two listening test experiments conducted for the technology-driven test. Figure 1 shows the mean values with the 95% confidence intervals for both experiments. A detailed discussion on the results is given in the following subsections.

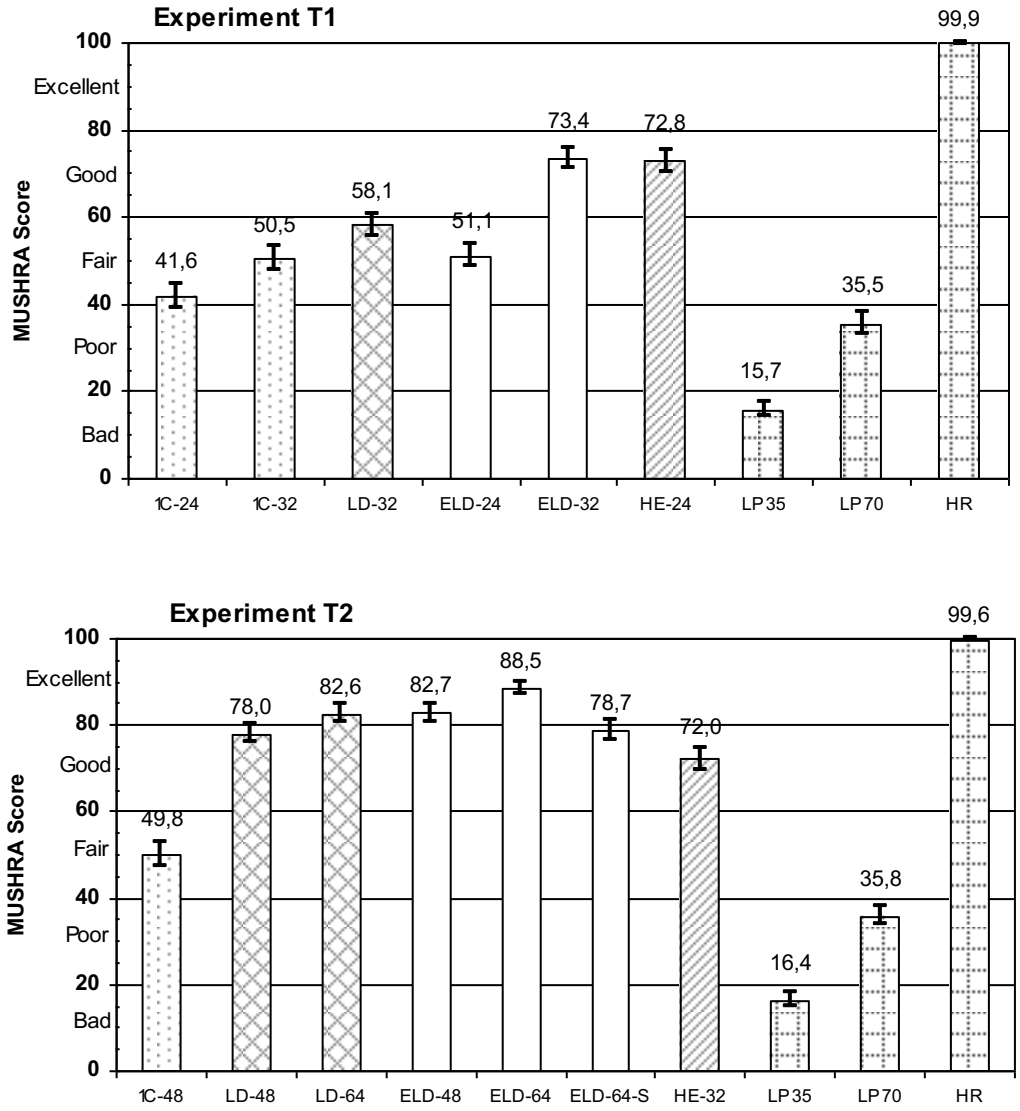


Figure 1: Results of the technology driven test (Experiments T1 and T2).

© ISO/IEC 2008 – All rights reserved

4.2.1 Experiment T1, Low Bit Rates

Table 7: The mean values with the lower and upper 95% confidence interval limits in Experiment T1.

Codec ID	Lower	Mean	Upper
1C-24	38,78	41,57	44,36
1C-32	47,56	50,45	53,34
LD-32	55,40	58,05	60,71
ELD-24	48,48	51,11	53,73
ELD-32	71,08	73,39	75,69
HE-24	70,31	72,75	75,19
LP35	14,03	15,72	17,41
LP70	33,00	35,53	38,06
HR	99,75	99,86	99,97

The following statistical statements can be made concerning the results of this experiment (shown in Table 7) as pooled over all excerpts:

- The mean score for the AAC ELD codec at a bit rate of 32 kbps (ELD-32) is over 15 points higher than the mean score for the G.722.1-C codec (1C-32) or the AAC LD codec (LD-32) at the same bit rate. The AAC ELD codec at a bit rate of 32 kbps is the only low delay coding system in this experiment that scores in the Good range on the MUSHRA scale regardless of having 6 ms (or 22%) lower algorithmic delay than AAC LD and G.722.1-C.
- The mean score for the AAC ELD at a bit rate of 32 kbps (ELD-32) is not different from, at the 95% level of significance, that of the HE AAC codec at a bit rate of 24 kbps (HE-24).
- The mean score for the AAC ELD codec at a bit rate of 24 kbps (ELD-24) is, at the 95% level of significance, higher than that of the G.722.1-C codec at the same bit rate (1C-24). Thus, AAC ELD codec is in this experiment the best low delay coding system for a bit rate of 24 kbps in terms of audio quality.
- The mean score for the AAC LD codec at a bit rate of 32 kbps (LD-32) is, at the 95% level of significance, higher than that of the G.722.1-C codec at the same bit rate (1C-32).

To complement the above observations that are based on the data pooled over all excerpts, we note that the AAC ELD codec at a bit rate of 32 kbps (ELD-32) is the only system in this experiment with the mean score of all items consistently in the Good or Excellent range on the MUSHRA scale.

© ISO/IEC 2008 – All rights reserved

4.2.2 Experiment T2, High Bit Rates

Table 8: The mean values with the lower and upper 95% confidence interval limits in Experiment T2.

Codec ID	Lower	Mean	Upper
1C-48	47,04	49,82	52,61
LD-48	75,70	77,99	80,28
LD-64	80,73	82,62	84,51
ELD-48	80,71	82,70	84,70
ELD-64	87,06	88,53	90,00
ELD-64-S	76,38	78,70	81,02
HE-32	69,39	72,04	74,70
LP35	14,89	16,36	17,83
LP70	33,79	35,81	37,83
HR	99,39	99,58	99,77

The following statistical statements can be made concerning the results of this experiment as pooled over all excerpts and shown in Table 8:

- The mean score for the AAC ELD and AAC LD codecs at a bit rate of 48 kbps (ELD-48, LD-48) is, at the 95% level of significance, higher than the mean score for the G.722.1-C codec at the same bit rate (1C-48).
- The mean score for the AAC ELD codec at a bit rate of 48 kbps (ELD-48) is, at the 95% level of significance, higher than that of the AAC LD codec at the same bit rate (LD-48). Furthermore, AAC ELD is the only system having a mean score in the Excellent range at this bit rate. Thus, AAC ELD is in this experiment the best low delay coding system regarding audio quality.
- The mean score for the AAC ELD codec at a bit rate of 48 kbps (ELD-48) is not different from, at the 95% level of significance, that of the AAC LD codec operating at a bit rate of 64 kbps (LD-64), or at a 33% higher bit rate.
- The mean score for the AAC ELD codec utilizing SBR at a bit rate of 64 kbps (ELD-64) is higher than, at the 95% level of significance, that of the AAC LD codec at the same bit rate (LD-64).
- The mean score for the AAC ELD codec that has SBR disabled at a bit rate of 64 kbps (ELD-64-S) was not different, at the 95% level of significance, from that of the AAC LD codec at the equal bit rate (LD-64), although this AAC ELD configuration operates at 5 ms (or 25%) shorter delay.

Furthermore, the AAC ELD codec utilizing SBR at a bit rate of 64 kbps (ELD-64) is the only system in this experiment with mean scores over all items that are consistently in the Excellent range on the MUSHRA scale.

© ISO/IEC 2008 – All rights reserved

4.3 Application-driven Test

This section provides the results of the four listening test experiments conducted for the application-driven test. Figure 2 shows the mean values with their 95% confidence intervals for the low bit rate experiments while Figure 2 presents the results of the high bit rate experiments. A detailed discussion on the results is given in the following subsections.

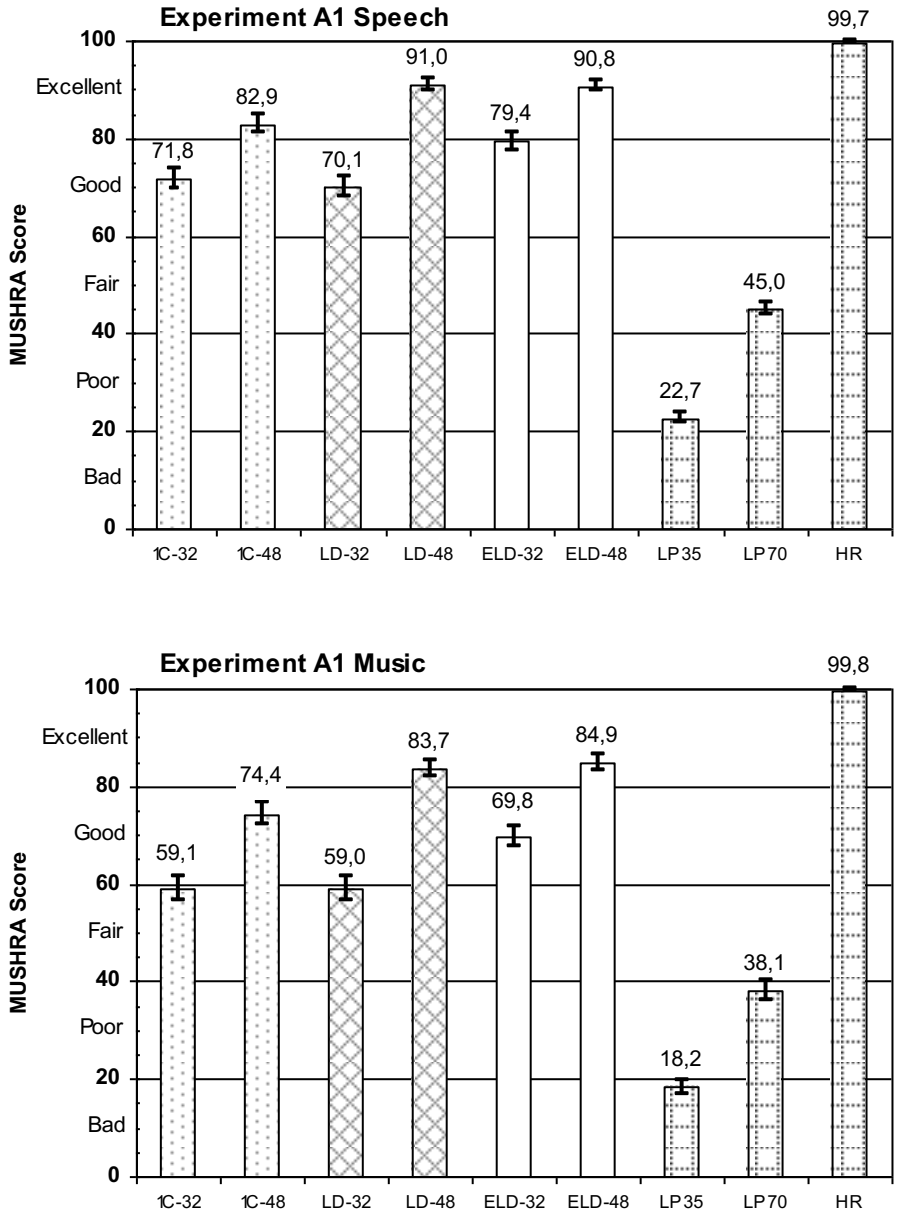
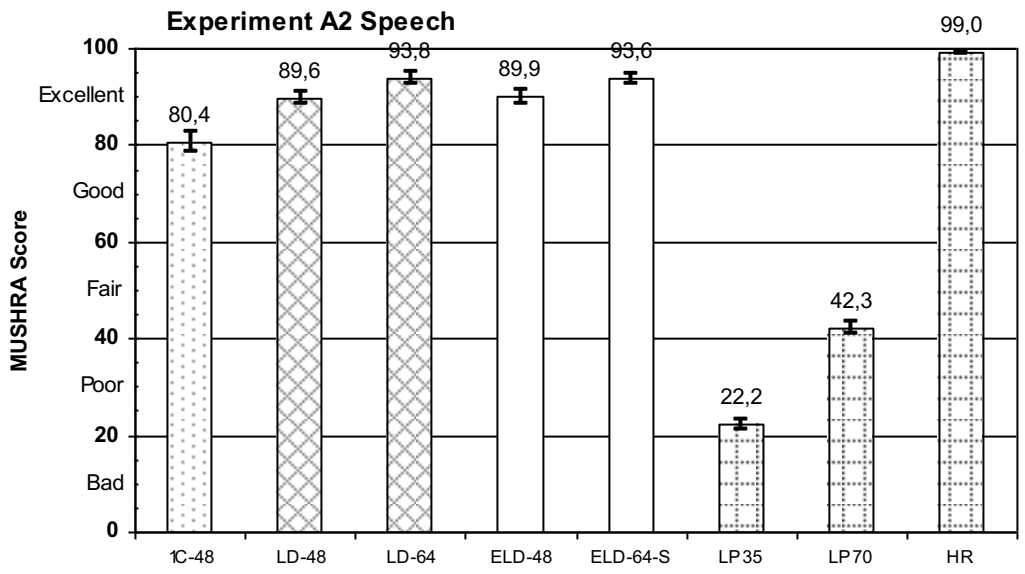


Figure 2: Results of the high quality, low bit rate communication scenario (Experiments A1 Speech and A1 Music).

© ISO/IEC 2008 – All rights reserved



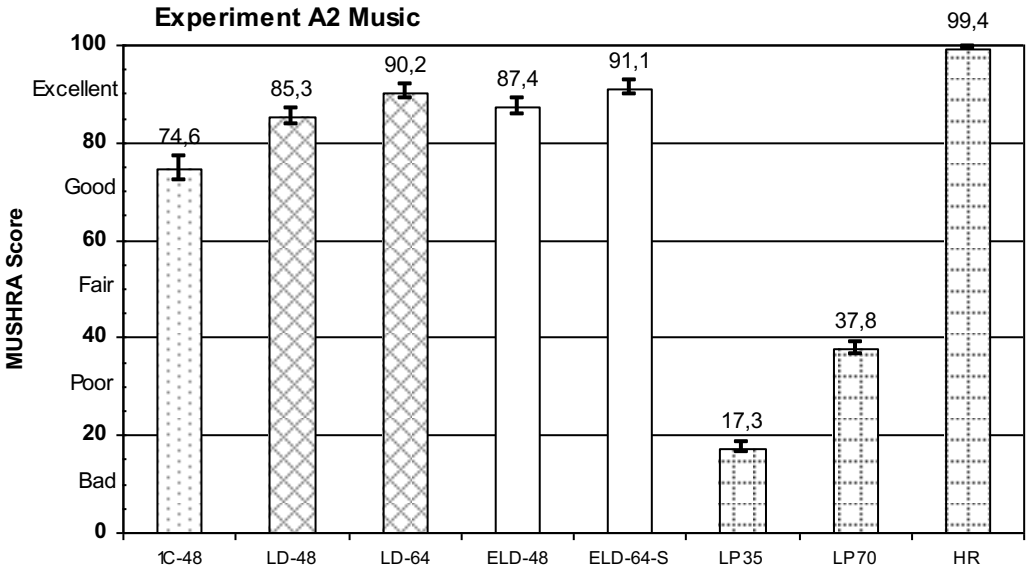


Figure 3: Results of the high quality communication scenario (Experiments A2 Speech and A2 Music).

© ISO/IEC 2008 – All rights reserved

4.3.1 Experiment A1 Speech

Table 9: The mean values with the lower and upper 95% confidence interval limits in Experiment A1 Speech.

Codec ID	Lower	Mean	Upper
1C-32	69,81	71,84	73,86
1C-48	81,06	82,88	84,70
LD-32	67,90	70,12	72,33
LD-48	89,87	90,95	92,04
ELD-32	77,61	79,42	81,23
ELD-48	89,60	90,77	91,93
LP35	21,76	22,73	23,69
LP70	43,67	45,01	46,34
HR	99,58	99,72	99,86

The following statistical statements can be made concerning the results of this experiment as pooled over all excerpts and shown in Table 9:

- The mean score for the AAC ELD codec at a bit rate of 32 kbps (ELD-32) is higher, at the 95% level of significance, than the mean score for the G.722.1-C codec and the AAC LD codec at the same bit rate (1C-32, LD-32). Thus, AAC ELD is the best coding system regarding audio quality in this test for a bit rate of 32 kbps while offering 8 ms (or 22%) lower algorithmic delay compared to these two references.
- The mean score for the AAC ELD codec at a bit rate of 32 kbps (ELD-32) is, at the 95% level of significance, not different from the mean score for the G.722.1-C codec that operates at a bit rate of 48 kbps (1C-48), at 50% higher bit rate.
- The mean scores for the AAC ELD and AAC LD codecs at a bit rate of 48 kbps (ELD-48, LD-48) are, at the 95% level of significance, higher than the mean score for the G.722.1-C codec at the same bit rate (1C-48).

4.3.2 Experiment A1 Music

Table 10: The mean values with the lower and upper 95% confidence interval limits in Experiment A1 Music.

Codec ID	Lower	Mean	Upper
1C-32	56,58	59,09	61,60
1C-48	72,11	74,38	76,65
LD-32	56,70	58,99	61,28
LD-48	82,16	83,71	85,25
ELD-32	67,62	69,76	71,91
ELD-48	83,32	84,94	86,56
LP35	17,00	18,24	19,49
LP70	36,07	38,06	40,04
HR	99,65	99,78	99,91

The following statistical statements can be made concerning the results of this experiment as pooled over all excerpts and shown in Table 10:

- The mean score for the AAC ELD codec at a bit rate of 32 kbps (ELD-32) is, at the 95% level of significance, higher than the mean scores for the G.722.1-C codec (1C-32) and the AAC LD

© ISO/IEC 2008 – All rights reserved

codec (LD-32) at the same bit rate. Furthermore, AAC ELD is the only system having a mean score in the Good range at this bit rate. Thus, AAC ELD is the best coding system regarding audio quality in this experiment for a bit rate of 32 kbps while offering 8 ms (or 22%) lower algorithmic delay compared to these two references.

- The mean scores for the AAC ELD codec (ELD-48) and the AAC LD codec (LD-48) at a bit rate of 48 kbps is higher, at the 95% level of significance, than the mean scores for the G.722.1-C (1C-48) codec at the same bit rate. The AAC ELD and AAC LD codecs at this bit rate are the only systems in this experiment having a mean score in the Excellent range on the MUSHRA scale.

4.3.3 Experiment A2 Speech

Table 11: The mean values with the lower and upper 95% confidence interval limits in Experiment A2 Speech.

Codec ID	Lower	Mean	Upper
1C-48	78,36	80,41	82,47
LD-48	88,25	89,59	90,93
LD-64	92,76	93,85	94,93
ELD-48	88,62	89,89	91,16
ELD-64-S	92,64	93,62	94,60
LP35	21,22	22,19	23,16
LP70	41,06	42,28	43,50
HR	98,78	99,04	99,30

The following statistical statements can be made concerning the results of this test as pooled over all excerpts and shown in Table 11:

- The mean scores for the AAC ELD and AAC LD codecs at a bit rate of 48 kbps (ELD-48, LD-48) were higher, at the 95% level of significance, than the mean score for the G.722.1-C codec (1C-48) at the same bit rate.
- The mean score for the AAC ELD codec that has SBR disabled and operates at a bit rate of 64 kbps (ELD-64-S) was not, at the 95% level of significance, different from the mean score for the to AAC LD codec at the equal bit rate (LD-64), although this AAC ELD configuration offers 5 ms (or 25%) lower algorithmic delay.

As a further observation, the AAC ELD and AAC LD systems included in test are the only systems in this experiment that provide the mean score in the Excellent range for each individual test item.

© ISO/IEC 2008 – All rights reserved

4.3.4 Experiment A2 Music

Table 12: The mean values with the lower and upper 95% confidence interval limits in Experiment A2 Music.

Codec ID	Lower	Mean	Upper
1C-48	72,00	74,55	77,11
LD-48	83,53	85,28	87,03
LD-64	88,73	90,21	91,69
ELD-48	85,68	87,38	89,07
ELD-64-S	89,79	91,14	92,50
LP35	16,29	17,30	18,32
LP70	36,58	37,77	38,95
HR	99,13	99,38	99,63

The following statistical statements can be made concerning the results of this experiment as pooled over all excerpts and shown in Table 12:

- The mean scores for the AAC ELD and AAC LD codecs at a bit rate of 48 kbps (ELD-48, LD-48) were higher, at the 95% level of significance, than the mean score for the G.722.1-C codec at the same bit rate (1C-48). These AAC LD and AAC ELD configurations are the only 48 kbps systems with a mean score in the Excellent range on the MUSHRA scale.
- The mean scores for the AAC ELD and AAC LD codecs at a bit rate of 64 kbps (ELD-64, LD-64) were higher, at the 95% level of significance, than the mean score for the same codecs running at 48 kbps.
- The mean score for the AAC ELD codec with SBR disabled at a bit rate of 64 kbps (ELD-64-S) is not different, at the 95% level of significance, from that of the AAC LD codec at the same bit rate (LD-64) regardless of 25% lower algorithmic delay.
- The mean score of all AAC ELD and AAC LD systems included in this experiment is in the Excellent range on the MUSHRA scale.

As the final observation from the data of individual test items, only the AAC ELD and AAC LD codecs at a bit rate of 64 kbps have in this experiment the mean score of all items consistently in the Excellent range on the MUSHRA scale.

5 Conclusions

The AAC ELD technology showed good performance in the verification tests with consistent and significant improvement in compression efficiency at the lowest tested bit rates, and comparable or better performance at the highest tested bit rates compared the previous low delay codec standard specified by MPEG. This gain in compression efficiency is attained with a substantially lower algorithmic delay that can be expected to contribute to better end-to-end performance in conversational applications compared to current state of the art high quality communication codecs.

More specifically, we can draw the following conclusions from the verification test results:

- In the low bit rate technology driven experiment, the mean score of the AAC ELD codec at a bit rate of 24 kbps (ELD-24) was, at the 95% level of significance, not different to that of the G.722.1-C codec at a bit rate of 32 kbps at both test sites.
- At a bit rate of 32 kbps, the mean score of the AAC ELD codec (ELD-32) is consistently, at the 95% level of significance, higher over all experiments and test sites than that of the AAC LD codec (LD-32) and the G.722.1-C (1C-32) at the same bit rate. At this bit rate, the AAC ELD codec (ELD-32) attains the mean score in the Good or Excellent range on the MUSHRA scale

© ISO/IEC 2008 – All rights reserved

for 33 out of all 34 items included in the experiments across all test sites. The mean score for the one remaining item (applause) was 0.5 points below the Good range.

- For a bit rate of 48 kbps, the AAC-ELD codec (ELD-48) is the only system reaching a mean score consistently in the Excellent range on the MUSHRA scale across all test sites in all five experiments in which this system was tested. At this bit rate, the mean scores of the AAC LD and AAC ELD codecs (LD-48, ELD-48) were consistently, at the 95% level of significance, higher than that of the G.722.1-C codec (1C-48) at the same bit rate in 12 out of 14 test sites. At the two remaining test sites, these systems were equal.
- At a bit rate of 64 kbps, the AAC ELD codec with SBR disabled (ELD-64-S) attains a mean score in the Excellent range in seven out of the eight test sites that participated in the three experiments that assessed this system. Across these eight test sites, this system had, at the 95% level of significance, a mean score not worse than that of the AAC LD codec at the same bit rate (LD-64) regardless of about 25% lower algorithmic delay.

The AAC ELD technology attains the quality level detailed above at a substantially lower algorithmic delay compared to the AAC LD and G.722.1-C codecs included in the verification tests as references. The algorithmic delay of the AAC ELD systems under test at bit rates of 32 kbps and 48 kbps is from 31.3 ms to 33.3 ms depending on the configuration, representing over 20% reduction compared to the G.722.1-C reference codec. At a bit rate of 64 kbps, the AAC ELD technology enables an algorithmic delay as low as 15 ms while maintaining high quality.

6 Acknowledgements

The following persons are acknowledged for their contributions to the verification test:

Name	Company
Schuyler Quackenbush	Audio Research Labs
Markus Schnell	Fraunhofer IIS
Vesa Ruoppila	Dolby Germany
Ralf Geiger	Fraunhofer IIS
Jeongil Seo	ETRI
Kristofer Kjörling	Dolby Sweden
Charles Robinson	Dolby San Francisco
Johannes Boehm	Thomson Hannover
Pierrick Philippe	France Télécom R&D

7 References

[1] ITU-R Recommendation BS.1534, Method for the subjective assessment of intermediate quality levels of coding systems, available at http://www.itu.int

[2] ITU-T Recommendation G.191, Software tools for speech and audio coding standardization, August 2005.

[3] ISO/IEC 14496-3:2005 AMENDMENT 9: Enhanced low delay AAC

[4] Audio Subgroup, "Report on the Verification Tests of MPEG-4 High Efficiency AAC", ISO/IEC JTC1/SC29/WG11/N6009, Brisbane, October 2003

[5] ISO/IEC 14496-3:2005, "Information technology — Coding of audio-visual objects Part 3: Audio"

© ISO/IEC 2008 – All rights reserved

[6] ITU-T Recommendation G.722.1, Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss, 2005.

© ISO/IEC 2008 – All rights reserved

A Test Methodology

A.1 MUSHRA test

The subjective assessment of sound quality (MUSHRA) test methodology specified in ITU-R Recommendation BS.1534 was used with the following stimuli as references throughout the tests:

- full bandwidth reference
- full bandwidth hidden reference
- low-pass filtered hidden reference (3.5 kHz, processed as described in A.3)
- low-pass filtered hidden reference (7.0 kHz, processed as described in A.3)

The listening test design is defined as follows:

- at least 20 listeners in total per experiment across all sites
- at least 8 listeners per site per test
- test sites (see Section 3)
- expert listeners are preferred
- training with the corresponding selected training items
- computer-based MUSHRA presentation is used with enabled instantaneous switching between the items under test
- Reference quality headphones to be used (e.g. Sennheiser HD600 or STAX) for headphone listening test. Test sites are requested to report on the ones they actually used
- one listener at a time due to open headphones and interactive control
- It is recommended to divide the test into about 20 minute sessions in order to prevent listener fatigue
- the test will be preceded by a training

A.2 Training of Subjects

The first step of training is to listen to some coded items in order to become familiar with the characteristics of the artefacts. The subjects can discuss the perceived artefacts, but subjects are not allowed to talk about specific grades in order to avoid bias in individual grading. The randomization of the order of presentation of the training items and the number of repetitions of the items is at the discretion of each listening test site.

The second step of the training is to run a dummy grading of the training items using the grading facility (paper sheet or on-screen display) to become familiar with this tool for the subsequent grading phase.

© ISO/IEC 2008 – All rights reserved

A.3 Processing Steps for 3.5 kHz and 7 kHz Anchors

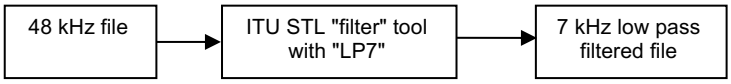


Figure 4: Processing for the 7 kHz anchors.

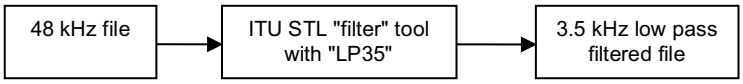


Figure 5: Processing for the 3.5 kHz anchors**.**

A.4 Statistical Analysis

The statistical analysis applied to the listening test data followed the standard MUSHRA procedure. The calculation of the averages of the scores of all listeners remaining after post-screening results in the Mean Subjective Scores (MSS). The first step of the analysis of the results is the calculation of the mean score $\bar{u}_{jk}$ for each of the presentations:

$$\bar{u}_{jk} = \frac{1}{N}\sum_{i=1}^{N} u_{ijk},$$

where u_i is the score of the observer *i* for the given test condition *j* and the sequence *k*, and *N* is the number of observers.

Confidence intervals were also calculated which was derived from the standard deviation and the size of each sample. The 95% confidence interval is given by

$$[\bar{u}_{jk} - \delta_{jk}, \bar{u}_{jk} + \delta_{jk}],$$

where

$$\delta_{jk} = 1.96\frac{S_{jkl}}{\sqrt{N}}$$

and the standard deviation

$$S_{jk} = \sqrt{\sum_{i=1}^{N}\frac{(\bar{u}_{jk} - u_{ijk})^2}{(N-1)}}.$$

With a probability of 95%, the absolute value of the difference between the experimental mean score and the “true” mean score (for a very high number of observers) is smaller than the 95% confidence interval, on condition that the distribution of the individual scores meets certain requirements.

© ISO/IEC 2008 – All rights reserved

A.5 Post-screening

All results in this report are based the data that were post-screened by excluding those subjects who rated the reference stimulus HR below 90 MUSHRA points or rated the reference stimulus and the anchors inconsistently such that the requirement LP3 ≤ LP7 ≤ HR was violated for one or more items in the experiment. This post-screening procedure was applied independently to each experiment.

Table 13 reports the number of subjects for each experiment before and after the post-screening procedure.

Table 13: Number of listeners before and after post-screening.

Experiment	Number of listeners	Number of listeners after post-screening
T1	20	20
T2	19	18
A1 Speech	29	27
A1 Music	29	29
A2 Speech	28	27
A2 Music	27	24

© ISO/IEC 2008 – All rights reserved

A.6 Test Items

A.6.1 List of Test Items

Table 14: Speech items used in Experiments A1 Speech and A2 Speech.

No.	Item name	Category	Source	Duration	Description
1	es04	clean speech	FhG	8 s	Item 2 processed as Figure 6
2	es05	clean speech	FhG	9 s	Item 3 processed as Figure 6
3	nadib09	clean speech	FhG	20 s	Item 5 processed as Figure 6
4	dan_speech_reverb	reverberant speech	FhG	14 s	Item 1 processed as Figure 7
5	jap_male_speech_re…	reverberant speech	FhG	9 s	Item 8 processed as Figure 7
6	nadib09_reverb	reverberant speech	FhG	20 s	Item 5 processed as Figure 7
7	nadib13_nadib08_....	reverberant speech + office noise + interfering talkers	FhG	14 s	Item 6 as main, Item 7 as interferer and office sample 1 processed as Figure 8
8	jap_male_speech_es...	reverberant speech + office noise + interfering talkers	FhG	9 s	Item 8 as main, Item 3 as interferer and noise sample 1 processed as Figure 8
9	nadib01_nadib13_...	reverberant speech + office noise + interfering talkers	FhG	18 s	Item 10 as main, Item 6 as interferer and noise sample 1 processed as Figure 8
10	nadib08_nadib13_...	reverberant speech + office noise + interfering talkers	FhG	12 s	Item 7 as main, Item 6 as interferer, noise sample 1,7,5 processed as Figure 8 with more than one noise items
11	nadib14_es05_48m_...	reverberant speech + office noise + interfering talkers	FhG	12 s	Item 4 as main, Item3 as interferer, noise sample 1,8,3 processed as Figure 8 with more than one noise items
12	nadib10_nadib09_...	reverberant speech + office noise + interfering talkers	FhG	15 s	Item 9 as main, Item 5 as interferer, noise sample 1,2,6 processed as Figure 8 with more than one noise items

Table 15: Music and mixed content items used in Experiments A1 Music and A2 Music.

No.	Item name	Category	Source	Duration	Description
1	nadib28	Speech with background music	Nadib	14 s	French male speech with background jingle
2	V1_fs48_mono	Speech with background music	FhG	10 s	English male speech with background jingle
3	nadib18	Mixed content	Nadib	18 s	English male speech + electronic music
4	nadib27	Mixed content	Nadib	10 s	German male speech + jingle
5	nadib30	Mixed content	Nadib	20 s	French male speech + French female speech + music
6	applause	Mixed content	Nadib	20 s	Music + applause
7	nadib21	Music	Nadib	20 s	English female singing voice with electronic music (dance floor)
8	nadib25	Music	Nadib	10 s	Jingle, intro
9	monowat3	Music	FhG	9 s	Lounge
10	rockyou	Music	FhG	11 s	Pop

© ISO/IEC 2008 – All rights reserved

Table 16: MPEG items used in Experiments T1 and T2.

No.	Item name	Category	Source	Duration	Description
1	es01	Speech	MPEG	11 s	Suzanne Vega, solo
2	es02	Speech	MPEG	9 s	Male German speech
3	es03	Speech	MPEG	8 s	Female English speech
4	sc01	Complex signal	MPEG	10 s	Trumpet
5	sc02	Complex signal	MPEG	13 s	orchestra
6	sc03	Complex signal	MPEG	12 s	Pop music
7	si01	Single instrument	MPEG	8 s	Harpsichord
8	si02	Single instrument	MPEG	8 s	Castanets
9	si03	Single instrument	MPEG	28 s	Pitch pipe
10	sm01	Single instrument	MPEG	11 s	Bag pipe
11	sm02	Single instrument	MPEG	10 s	Plucked strings
12	sm03	Single instrument	MPEG	14 s	Glockenspiel

A.6.2 Processing of Speech Items

The source material for generating speech items consists of a set of clean speech items (Table 17**Error! Reference source not found.**), a set of office background noise items (Table 18**Error! Reference source not found.**) and the software utilities for processing the test items from clean speech and background noise items such that the test items exhibit the desired reverberation and background noise characteristics. For generating clean speech items, the processing according to Figure 6 was applied. The processing for reverberant speech is described in Figure 7 and for reverberant speech with background office noise in Figure 8. For the processing of some items, more than one office noise source has been included.

The tools used for the various processing are part of the ITU-T Software Tool Library, Release 2005 [2]. The level adjustment was applied using the Speech Voltmeter. Depending on the input content, the mode was switched between the speech activity detection mode (P.56) and the RMS mode. The reverberation was performed by using the REVERB tool with the IR48 filter from the package "STL_gorg722.1_qualif.zip".

© ISO/IEC 2008 – All rights reserved

Table 17: Clean speech items used for generating the test items for reverberant speech and reverberant speech with background noise conditions.

No	Item name	Category	Source	Duration	Description
1	dan_speech	Clean speech	FhG	14 s	Danish male speech
2	es04	Clean speech	MPEG	8 s	English male speech
3	es05	Clean speech	MPEG	9 s	German female speech
4	nadib14	Clean speech	Nadib	12 s	Chinese female speech
5	nadib09	Clean speech	Nadib	20 s	German female speech
6	nadib13	Clean speech	Nadib	20 s	French female speech
7	nadib08	Clean speech	Nadib	20 s	Tajik male speech
8	jap_male_speech	Clean speech	FhG	8 s	Japanese male speech
9	nadib10	Clean speech	Nadib	20 s	English female speech
10	nadib01	Clean speech	Nadib	20 s	German male speech

Table 18: Office background noise items used for generating the test items.

No	Item name	Category	Source	Duration	Description
1	office	Office noise	FhG	91 s	sound of keyboards, air conditioner, telephone ringing
2	paper	Office noise	FhG	35 s	paper rustling
3	coffee	Office noise	FhG	29 s	sounds related to a coffee cup
4	computer_fan	Office noise	FhG	15 s	noise of a computer fan
5	telephone	Office noise	FhG	21 s	telephone ringing
6	typing	Office noise	FhG	32 s	typing sounds of a keyboard
7	typing2	Office noise	FhG	29 s	typing sounds of a keyboard
8	drawer	Office noise	FhG	27 s	opening and closing drawers
9	office_ensemble	Office noise	FhG	89 s	Ensemble of different office sounds, high activity
10	office_ensemble2	Office noise	FhG	241 s	Ensemble of different office sounds, lower activity

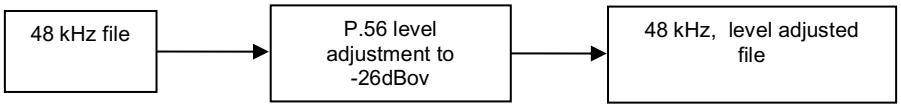


Figure 6: Pre-processing of clean speech.

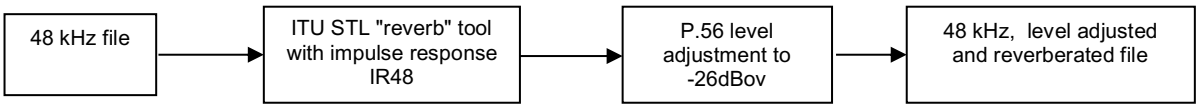


Figure 7: Pre-processing of reverberant speech.

© ISO/IEC 2008 – All rights reserved

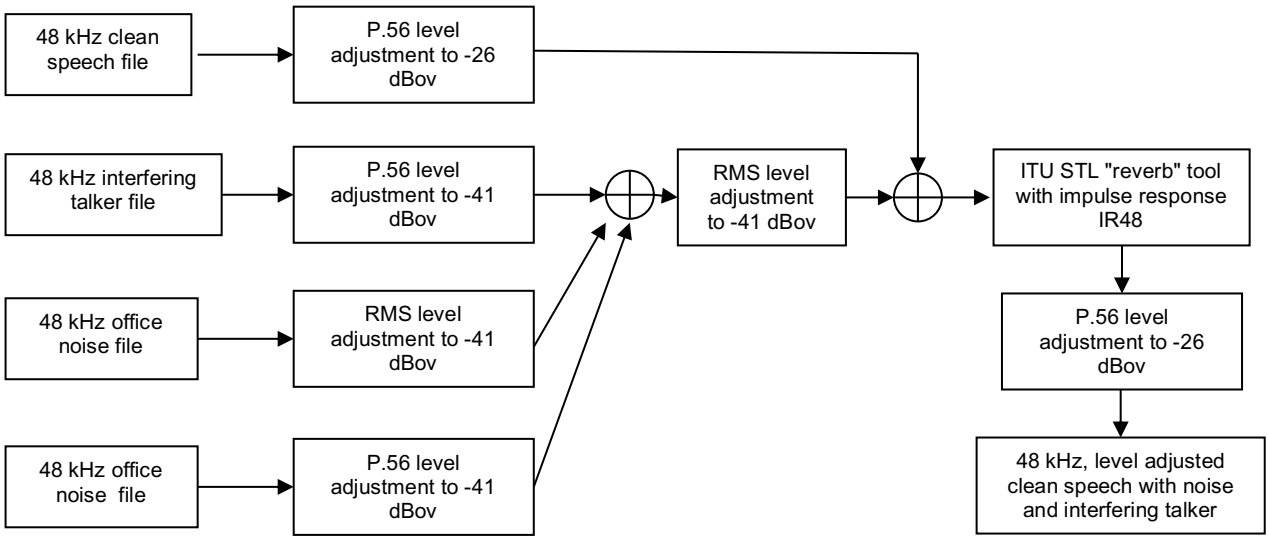


Figure 8: Pre-processing of reverberant speech with office noise and interfering talkers.

A.6.3 Pre- and Post-processing Steps for G.722.1 Annex C

Pre-processing:

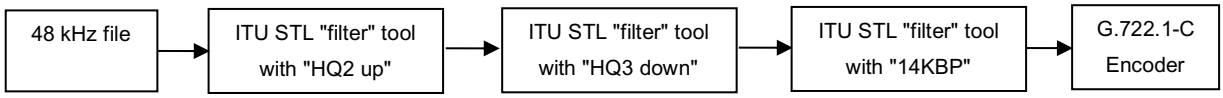


Figure 9: Pre-processing for the G.722.1-C codec conditions.

Post-processing:

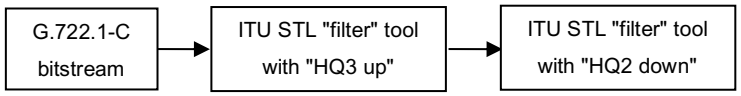


Figure 10: Post-processing for the G.722.1-C codec conditions.

© ISO/IEC 2008 – All rights reserved

B Detailed Results of All Test Sites

This section provides the detailed results of the experiments for each test site. The following figures and tables present the mean and 95% confidence intervals of the results. Note that the reference stimuli HR are not shown in the figures.

B.1 Technology-driven Test, Experiments T1 and T2

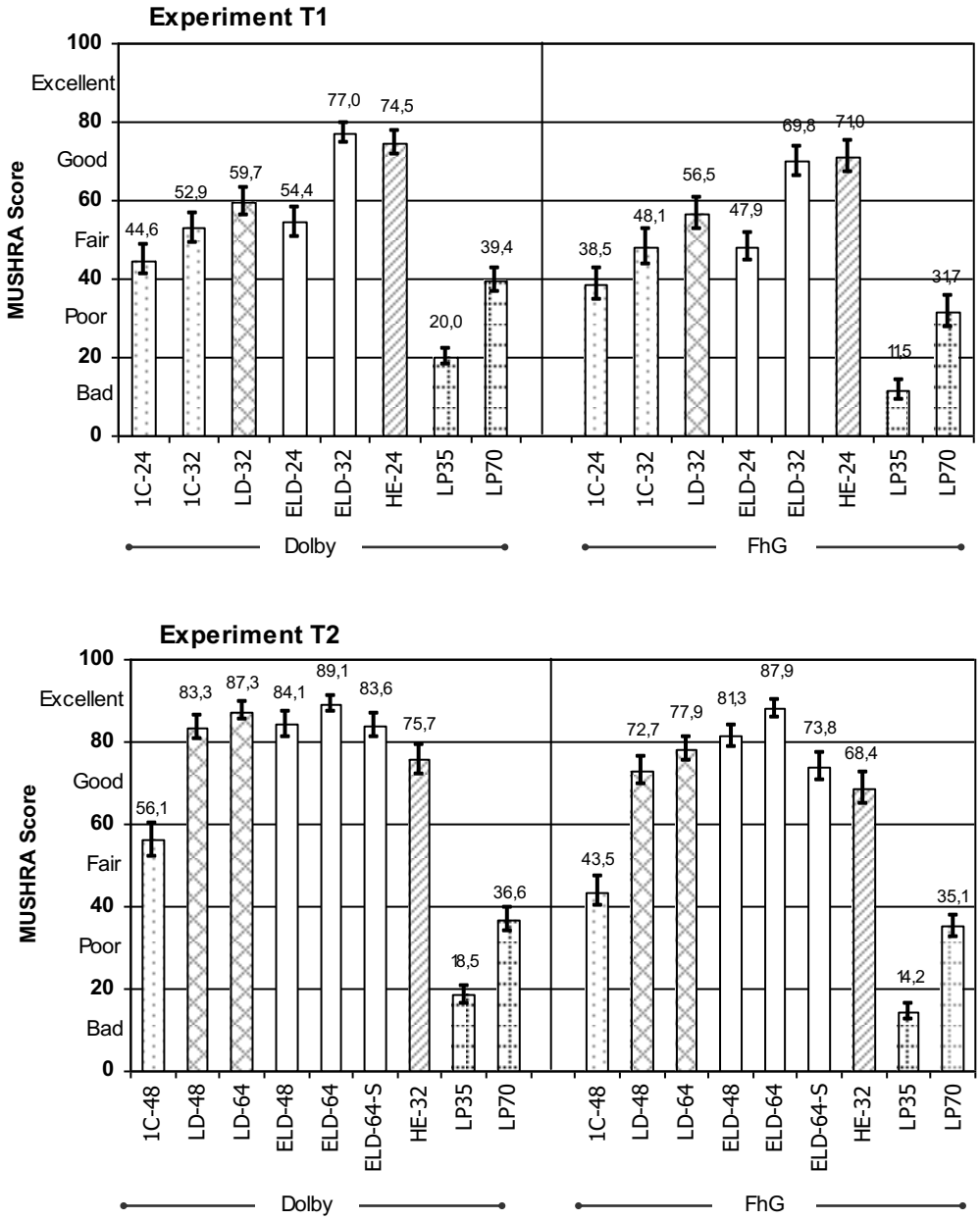


Figure 11: Results of Experiments T1 and T2 per codec and test site.

© ISO/IEC 2008 – All rights reserved

Table 19: The mean values with the lower and upper 95% confidence interval limits in Experiment T1 for both test sites.

Site	Codec ID	Lower	Mean	Upper
Dolby	1C-24	40,84	44,61	48,38
	1C-32	49,01	52,85	56,69
	LD-32	56,16	59,66	63,16
	ELD-24	50,60	54,37	58,13
	ELD-32	74,48	76,96	79,44
	HE-24	71,63	74,53	77,42
	HR	99,70	99,87	100,04
	LP35	17,93	19,99	22,05
	LP70	36,41	39,35	42,29
FhG	1C-24	34,46	38,53	42,60
	1C-32	43,74	48,05	52,36
	LD-32	52,45	56,45	60,45
	ELD-24	44,27	47,85	51,43
	ELD-32	66,02	69,82	73,61
	HE-24	67,07	70,98	74,90
	HR	99,72	99,85	99,98
	LP35	8,99	11,45	13,91
	LP70	27,70	31,72	35,73

Table 20: The mean values with the lower and upper 95% confidence interval limits in Experiment T2 for both test sites.

Site	Codec ID	Lower	Mean	Upper
Dolby	1C-48	52,09	56,13	60,17
	LD-48	80,36	83,25	86,14
	LD-64	85,13	87,31	89,48
	ELD-48	81,09	84,09	87,09
	ELD-64	87,11	89,15	91,19
	ELD-64-S	80,73	83,64	86,55
	HE-32	72,06	75,66	79,25
	HR	99,16	99,46	99,76
	LP35	16,38	18,50	20,62
	LP70	33,61	36,56	39,52
FhG	1C-48	40,07	43,52	46,97
	LD-48	69,44	72,72	76,00
	LD-64	75,09	77,93	80,76
	ELD-48	78,68	81,31	83,94
	ELD-64	85,78	87,91	90,03
	ELD-64-S	70,38	73,76	77,14
	HE-32	64,62	68,43	72,23
	HR	99,47	99,70	99,94
	LP35	12,26	14,22	16,19
	LP70	32,31	35,06	37,80

© ISO/IEC 2008 – All rights reserved

B.2 Application-driven Test

B.2.1 Experiment A1: High Quality, Low Bit Rate Communication

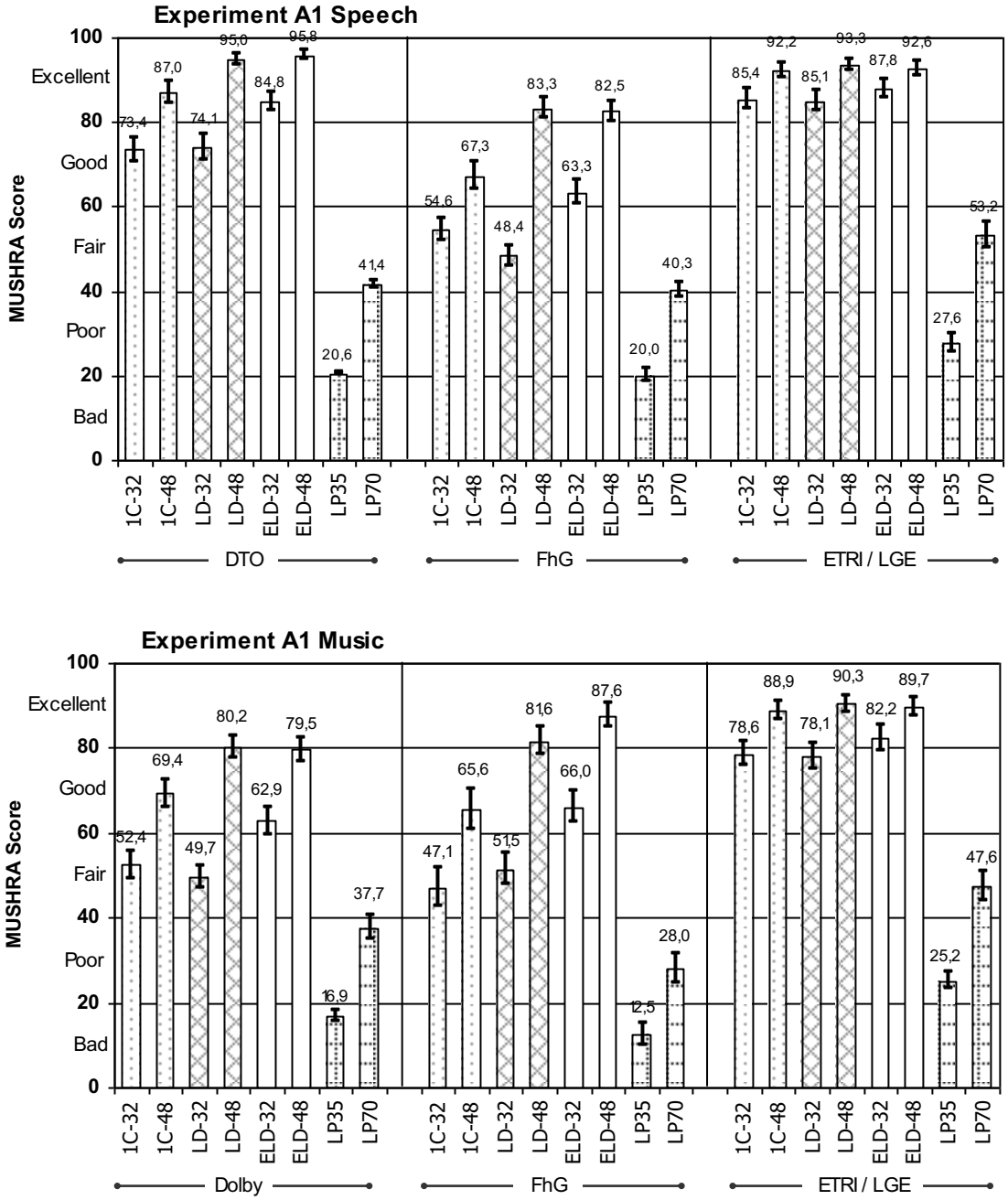


Figure 12: Results of Experiment A1 Speech and A1 Music per codec and test site.

© ISO/IEC 2008 – All rights reserved

Table 21: The mean values with the lower and upper 95% confidence interval limits in Experiment A1 Speech for all three test sites.

Site	Codec ID	Lower	Mean	Upper
DTO	1C-32	70,67	73,38	76,10
	1C-48	84,49	86,95	89,41
	LD-32	71,14	74,08	77,01
	LD-48	93,65	94,95	96,25
	ELD-32	82,50	84,78	87,05
	ELD-48	94,68	95,77	96,86
	HR	99,29	99,57	99,85
	LP35	20,15	20,55	20,95
	LP70	40,53	41,42	42,30
FhG	1C-32	52,00	54,60	57,21
	1C-48	64,15	67,31	70,47
	LD-32	46,06	48,35	50,64
	LD-48	80,96	83,27	85,59
	ELD-32	60,48	63,31	66,14
	ELD-48	79,97	82,48	84,99
	HR	99,50	99,75	100,00
	LP35	18,26	20,02	21,78
	LP70	38,50	40,27	42,05
ETRI/LGE	1C-32	82,90	85,44	87,98
	1C-48	90,45	92,20	93,96
	LD-32	82,64	85,06	87,49
	LD-48	92,05	93,34	94,63
	ELD-32	85,53	87,78	90,03
	ELD-48	91,00	92,58	94,17
	HR	99,72	99,86	100,00
	LP35	25,43	27,56	29,68
	LP70	50,19	53,20	56,22

© ISO/IEC 2008 – All rights reserved

Table 22: The mean values with the lower and upper 95% confidence interval limits in Experiment A1 Music for all three test sites.

Site	Codec ID	Lower	Mean	Upper
Dolby	1C-32	49,30	52,41	55,52
	1C-48	66,11	69,37	72,62
	LD-32	46,95	49,65	52,35
	LD-48	77,74	80,18	82,61
	ELD-32	59,68	62,92	66,15
	ELD-48	76,74	79,54	82,34
	HR	99,69	99,86	100,03
	LP35	15,63	16,89	18,15
	LP70	34,96	37,66	40,36
FhG	1C-32	42,66	47,13	51,59
	1C-48	60,93	65,60	70,27
	LD-32	47,93	51,48	55,02
	LD-48	78,37	81,58	84,78
	ELD-32	62,33	66,03	69,72
	ELD-48	84,80	87,63	90,45
	HR	99,13	99,50	99,87
	LP35	9,82	12,48	15,13
	LP70	24,62	27,95	31,28
ETRI/LGE	1C-32	75,73	78,62	81,51
	1C-48	86,77	88,87	90,97
	LD-32	75,11	78,12	81,13
	LD-48	88,42	90,31	92,20
	ELD-32	79,25	82,21	85,17
	ELD-48	87,69	89,74	91,80
	HR	99,81	99,91	100,02
	LP35	23,08	25,18	27,27
	LP70	44,16	47,57	50,98

© ISO/IEC 2008 – All rights reserved

B.2.2 Experiment A2: High Quality Communication

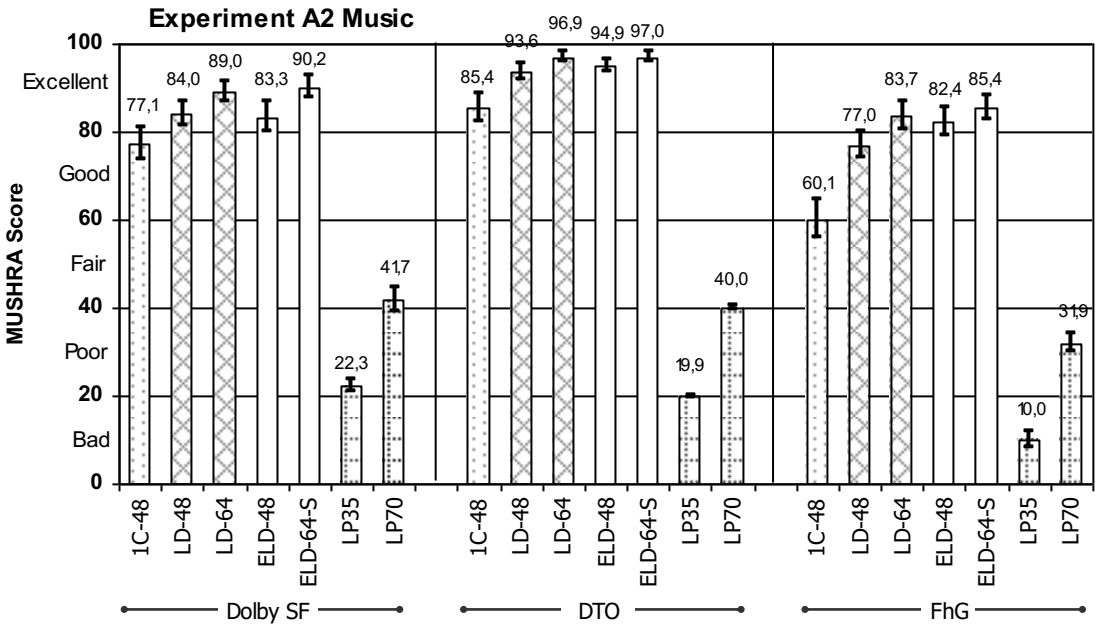


Figure 13: Results of Experiment A2 Speech per codec and test site.

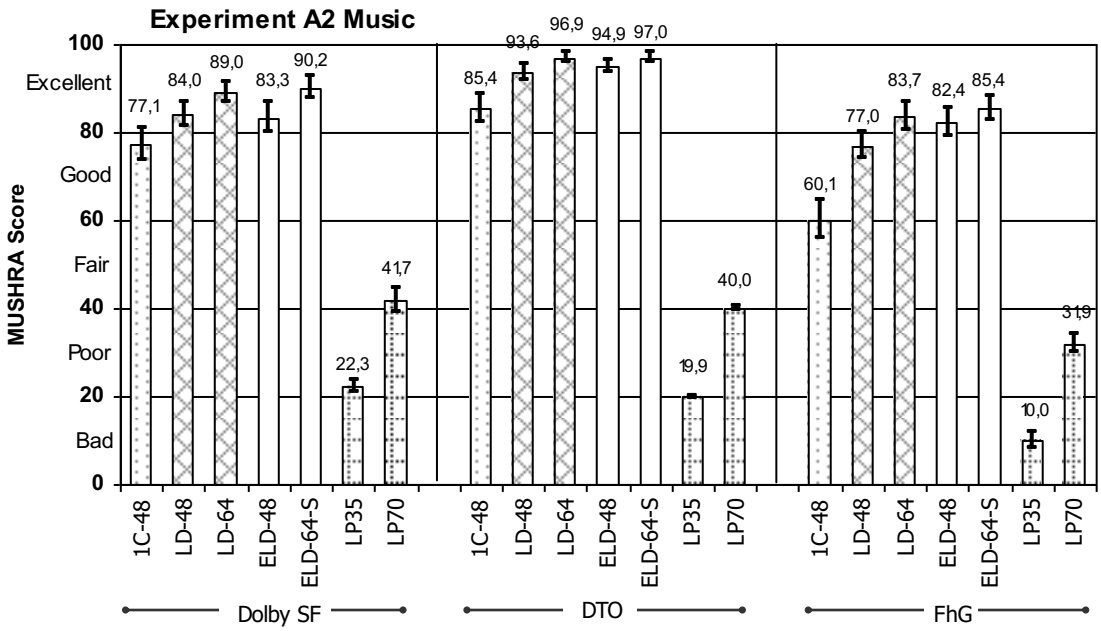


Figure 14: Results of Experiment A2 Speech and A2 Music per codec and test site.

© ISO/IEC 2008 – All rights reserved

Table 23: The mean values with the lower and upper 95% confidence interval limits in Experiment A2 Speech for all three test sites.

Site	Codec ID	Lower	Mean	Upper
Dolby SF	1C-48	79,37	82,39	85,40
	LD-48	86,51	88,43	90,35
	LD-64	91,73	93,17	94,60
	ELD-48	86,65	88,55	90,45
	ELD-64-S	90,51	92,41	94,30
	HR	97,21	97,91	98,60
	LP35	25,39	26,64	27,88
	LP70	43,15	45,96	48,77
DTO	1C-48	84,34	87,51	90,68
	LD-48	94,40	95,78	97,16
	LD-64	96,13	97,73	99,34
	ELD-48	95,99	96,99	98,00
	ELD-64-S	97,91	98,45	99,00
	HR	99,33	99,55	99,76
	LP35	19,87	20,07	20,26
	LP70	40,16	40,58	41,01
FhG	1C-48	65,23	68,69	72,15
	LD-48	79,25	82,25	85,25
	LD-64	86,99	89,19	91,38
	ELD-48	78,76	81,46	84,16
	ELD-64-S	86,12	88,19	90,25
	HR	99,13	99,48	99,83
	LP35	17,83	20,67	23,51
	LP70	38,10	40,94	43,77

© ISO/IEC 2008 – All rights reserved

Table 24: The mean values with the lower and upper 95% confidence interval limits in Experiment A2 Music for all three test sites.

Site	Codec ID	Lower	Mean	Upper
Dolby SF	1C-48	73,47	77,13	80,78
	LD-48	81,23	84,03	86,83
	LD-64	86,69	89,04	91,39
	ELD-48	80,03	83,31	86,60
	ELD-64-S	87,69	90,17	92,66
	HR	98,72	99,21	99,71
	LP35	20,80	22,33	23,85
	LP70	38,92	41,67	44,42
DTO	1C-48	82,13	85,38	88,62
	LD-48	91,64	93,59	95,54
	LD-64	95,80	96,93	98,07
	ELD-48	93,46	94,93	96,41
	ELD-64-S	96,06	97,02	97,99
	HR	99,57	99,80	100,03
	LP35	19,76	19,89	20,01
	LP70	39,73	39,99	40,24
FhG	1C-48	55,79	60,13	64,46
	LD-48	73,90	77,03	80,15
	LD-64	80,61	83,68	86,74
	ELD-48	79,23	82,43	85,62
	ELD-64-S	82,68	85,38	88,07
	HR	98,51	99,05	99,59
	LP35	8,14	10,00	11,86
	LP70	29,78	31,85	33,92

© ISO/IEC 2008 – All rights reserved

C Low Power SBR

C.1 Experiment Overview

In the course of the standardization of AAC ELD, verify the performance of the new complex modulated low delay filter bank utilized in the low delay SBR tool was verified with a listening test. It was of special interest to test the low power (LP) mode of the low delay SBR tool in comparison to its high quality (HQ) mode. In the HQ-mode, the frequency representation and all operations in the SBR module are based on complex values, whereas in the LP-mode the SBR module simply requires a real-valued representation. The low power mode leads to a reduced computational complexity for the SBR decoder part of about 30% [5].

The setup of the codec under test is outlined in Table 25.

Table 25: Codec configuration under test in the low power SBR experiment.

Codec ID	Codec	Bit rate	Sampling frequency	Frame length in samples	Delay	Audio bandwidth
ELD-40	AAC ELD	40 kbps	24/48 kHz (1)	512	33.3 ms	14.250 Hz

(1) The MPEG-4 Enhanced Low Delay AAC (ELD-40) is a dual rate system with the underlying AAC codec operating at half the sampling rate of the SBR tool.

The HQ and LP modes were compared using the MUSHRA methodology and the typical MPEG test item set described in Section A.6.1. This experiment was conducted by two test sites, FhG and Coding Technologies (currently Dolby) with a total number of 17 listeners.

C.2 Test Results

The following Figure presents the results of the low power SBR listening experiment. The plot shows the mean values with the 95% confidence intervals as pooled over all excerpts.

© ISO/IEC 2008 – All rights reserved

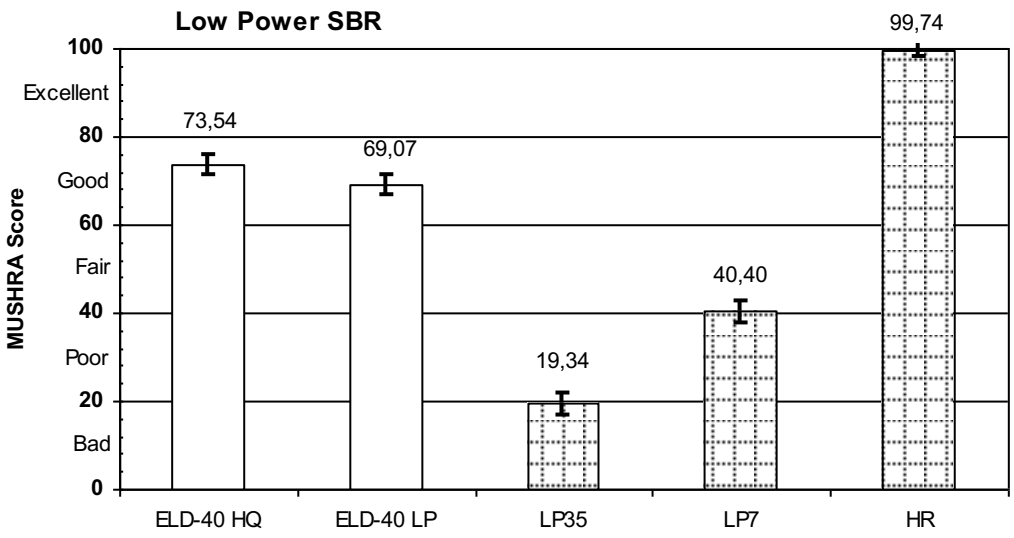


Figure 15: Results of the low power SBR experiment.

Table 26: The mean values with the lower and upper 95% confidence interval limits in the low power SBR experiment.

Codec ID	Lower	Mean	Upper
ELD-40 HQ	71,24	73,54	75,84
ELD-40 LP	66,60	69,07	71,55
LP35	17,73	19,34	20,94
LP7	38,26	40,40	42,53
HR	99,51	99,74	99,97

Based on the results of this experiment as pooled over all excerpts and shown in Table 26, the mean score for the AAC ELD codec utilizing the low power mode of the decoder was not different, at the 95% level of significance, from the mean score for the AAC ELD codec utilizing the high quality version of the decoder. The same statistical statement holds true on a per item based observation.

© ISO/IEC 2008 – All rights reserved