EBU subjective listening tests on low-bitrate audio codecs

June 2003

EBU·UER

fuboTV Media Inc.
Exhibit 1055
fuboTV v. Dish
IPR2024-00917, -00918, -00919
Ex. 1055-001

Summary

This document contains the results of recent evaluations carried out on several commercial audio codecs by EBU Project Group B/AIM (Audio In Multimedia), chaired by Gerhard Stoll (IRT). For the first time, a new methodology – specifically developed to evaluate "intermediate-quality" [1] codecs subjectively – was successfully used. This methodology is known as MUSHRA and is standardized within the ITU as ITU-R Recommendation BS.1534.

B/AIM members have considerable experience in evaluating audio codecs. In October 2000, the EBU published a document BPN 029 entitled "*EBU Report on the Subjective Listening Tests of Some Commercial Internet Audio Codecs*". This publication was well received by EBU Members and other organizations involved in multimedia and internet broadcasting. It was valued as an impartial evaluation of the intermediate-quality audio codecs available on the internet market at the time.

Following the publication of that document, many EBU Members requested that evaluations be carried out on the newer generation of codecs, such as Windows Media 8, CTaacPlus, mp3PRO, RealAudio 8 and others. The present document shows that, at low bitrates (i.e. 16 to 32 kbit/s), these newer codecs do not offer significant improvements over the previous-generation codecs. However, at 48 kbit/s to 64 kbit/s, a new technology – Spectral Band Replication (SBR) – has contributed to some significant improvements. This technology has been built into several existing codecs and has helped AAC and mp3 to become virtually transparent at 64 kbit/s. These are extremely encouraging developments.

The present study confirms what has been known before, namely that the quality of coded signals depends strongly on the signal content. Some codecs are very good for music while others are more suitable for speech.

It is believed that the present tests provide useful independent guidance to EBU Members about which codecs are technically suitable for a given purpose. It should be mentioned, however, that the technical quality alone is often not a decisive criterion when selecting a codec for internet operations. Other criteria – such as cost, digital rights management capabilities, encoding facilities, processing power required, streaming capability and availability of hardware implementation – are often more important.

As new codecs are becoming commercially available almost every day, the EBU plans to continue carrying out its independent, impartial, tests. Such work should not be done too often though, as it is felt that the quality difference between two successive generations is not sufficient to warrant such efforts. Nevertheless, it is felt that a new series of tests should be conducted every two or three years.

In conclusion, we believe that the best service the EBU can provide for its Members is to conduct high-quality, professional and reliable evaluations on market products such as internet audio codecs.
B/AIM is committed to continue doing precisely that.

Franc Kozamernik, Secretary of B/AIM

1. The term "intermediate quality" has been chosen to distinguish the quality usually available from internet streaming, from the true "broadcast quality" that is normally obtainable in digital broadcasting (e.g. DAB or DVB).

Ex. 1055-002

Ex. 1055-003

Contents

1. Introduction 5

2. Audio codecs under test 6

2.1. Microsoft® Windows Media™ 8 6
2.2. MPEG-2, MPEG-4 AAC (Advanced Audio Coding) 6
2.3. AAC+SBR 6
2.4. mp3PRO 7
2.5. RealNetworks RealAudio 8 7
2.6. AMR Wide Band 7
2.7. mp3 (MPEG-1, MPEG-2 Layer III derivative) 7
2.8. RealNetworks G2 7

3. Encoding process 7

4. Test material 8

5. Experimental design 9

5.1. Test method 9
5.2. Training phase 9
5.3. User interface 9
5.4. Listening panels 10
5.5. Test duration 10
5.6. Listening conditions 10

6. Main results 10

6.1. Results for 16 kbit/s mono 11
6.2. Results for 20 kbit/s stereo 12
6.3. Results for 32 kbit/s stereo 13
6.4. Results for 48 kbit/s stereo 14
6.5. Results for 64 kbit/s stereo 15
6.6. Codec performance as a function of bitrate 17

7. References 17

Appendix A: The detailed results 19

Ex. 1055-004

Ex. 1055-005

1. Introduction

In 1999, the EBU carried out subjective listening tests on intermediate-quality audio codecs and, in October 2000, published the results in document BPN 029 [1]. This document was well received by EBU Members and other organizations involved in multimedia and internet broadcasting. It was particularly valued as an impartial evaluation of which codecs would be technically suitable for a given purpose – as many codec proponents had been carrying out their own evaluations, often with some significant commercial bias. These EBU tests were unique in that all the major audio codecs available on the market were evaluated concurrently in a single series of tests, carried out by several independent EBU laboratories.

Since then, the market for internet audio codecs has evolved rapidly, with new codecs (and new versions of existing codecs) being launched at frequent intervals. These newer codecs have a range of different characteristics, are available in different configurations and can use different bitrates to suit different objectives.

For this reason, the EBU decided to launch a new series of evaluations in late 2001, on the commercial audio codecs then currently available. All the codecs tested were available in software and allowed for real-time encoding and decoding.

This work was entrusted again to EBU Project Group B/AIM (Audio in Multimedia), chaired by Gerhard Stoll (IRT).

The listening tests were performed by the following EBU Members and their partner organizations:

- **British Broadcasting Corporation** (BBC);
- **Institut für Rundfunktechnik** (IRT);
- **France Télécom R&D**;
- **Norsk Rikskringkasting** (NRK);
- **Fraunhofer AEMT** (FhG);
- **Telewizja Polska** (TVP);
- **T-Systems**.

The objective of these tests remained basically the same as before; i.e. to help EBU Members in choosing the codecs that are most suited for given internet applications using bitrates in the range from 16 to 64 kbit/s.

The tests were conducted using the new EBU MUSHRA test method [2], which is a recommended evaluation method adopted by the ITU [3]. This method has been specifically designed to assess the subjective audio quality of intermediate-quality audio codecs; that is, codecs that exhibit relatively large impairments. The test method given in ITU-R Recommendation BS.1116-1 [4] is not appropriate for the kind of quality considered in these tests.

In order to be able to relate the present test results to the previous ones, it was decided to reuse the same reference and hidden anchors, as well as two codecs (mp3 and RealNetworks version G2). In addition, several audio test sequences (but not all) were retained from the previous tests. In this way, the consistency and validity of the test results could be confirmed.

Ex. 1055-006

2. Audio codecs under test

The following audio codecs were tested:

1) **Microsoft Media 8** (WM8Eutil.exe);
2) **MPEG-2 AAC** (Advanced Audio Coding) – Implementation by FhG-IIS;
3) **AAC+SBR** (or CTaacPlus) – Implementation by Coding Technologies;
4) **mp3PRO** (Thomson Multimedia) – Implementation by Coding Technologies;
5) **RealNetworks RealAudio 8**;
6) **AMR Wide Band** – Implementation by FhG-IIS.

In order to test for consistency with the previous EBU evaluations (1999), the following two codecs used in those tests were also included in the new series of tests:

7) **mp3** (close to MPEG-1 and MPEG-2 Layer III) – Implementation by Opticom);
8) **RealNetworks G2**.

A short description of each of the above codecs is given in the following sections:

2.1. Microsoft® Windows Media™ 8

The Microsoft® Windows Media™ Audio codec is designed to handle all types of audio content, from speech-only audio – recorded with a sampling rate of 8 kHz – to 48 kHz high-quality stereo music. The Windows Media Audio codec scales its bitrate to suit the target network bandwidth. It produces bitstreams with bitrates ranging between 5 kbit/s and 192 kbit/s. A Digital Rights Management (DRM) system is available for Windows Media. A DRM-enabled file is encrypted and cannot be played without the decryption key provided in the user licence.

2.2. MPEG-2, MPEG-4 AAC (Advanced Audio Coding)

AAC forms part of the MPEG-2 and MPEG-4 standards. It uses waveform coding, based on the traditional MPEG audio time/frequency transform. The AAC (Advanced Audio Coding) coder used in this test was the MPEG-2 AAC Main profile encoder, according to ISO/IEC 13818-7 and implemented by FhG-IIS (Fraunhofer Gesellschaft, Erlangen, Germany). AAC was used with four sampling rates, ranging from 8 to 32 kHz (depending on the bitrate adopted). MPEG-2 AAC provides up to 48 channels of audio in one stream, at sampling rates of up to 96 kHz.

In MPEG-4, the number of audio objects is handled by MPEG-4 systems and is practically unlimited. Developed and standardized within ISO/IEC MPEG-4, AAC is supported by a growing number of hardware and software manufacturers. For example, Envivio, RealNetworks, Liquid Audio and Packet Video are using, or intend to use, AAC for internet streaming. The ISMA (Internet Streaming Media Alliance) has included AAC in their current specification, and it has been adopted as the audio standard for digital broadcasting in Japan.

2.3. AAC+SBR

SBR (Spectral Band Replication) is a new audio coding enhancement tool. It offers the possibility of improving the performance of low-bitrate audio and speech codecs – by either increasing the audio bandwidth at a given bitrate or by improving the coding efficiency at a given quality level. In these tests, the application of SBR to AAC was provided by Coding Technologies and, since May 2002, has been called “CTaacPlus” [5]. Only one bitrate (48 kbit/s stereo) was available at the time of the tests. CTaacPlus is used in the Digital Radio Mondiale (DRM) system [6].

Ex. 1055-007

2.4. mp3PRO

mp3PRO is the combination of mp3 and SBR in a backward-compatible way. Conventional mp3 players can still decode the compatible part of the bitstream, which represents the lower part of the frequency range of the audio signal. mp3PRO players are capable of playing legacy mp3 streams. The implementation in the test was provided by Thomson Multimedia. Two bitrates (48 and 64 kbit/s stereo) were available at the time of the tests. However, bitrates in the range from 24 to 48 kbit/s mono and from 32 to 96 kbit/s stereo are now available.

2.5. RealNetworks RealAudio 8

RealAudio 8 was commercially introduced in October 2000 as a new audio standard for internet audio. Developed in part through a strategic alliance with Sony Corporation, RealAudio 8 includes ATRAC3, Sony's quality audio compression technology. The system was designed to deliver quality audio at the widest range of bitrates. It delivers comparable or better audio quality when compared to RealAudio G2, the prior RealNetworks standard for internet audio. RealAudio 8 files can be created using RealProducer, RealJukebox® or a wide range of third party products.

2.6. AMR Wide Band

The AMR (Adaptive Multi-Rate) Wide Band system is a codec for wideband speech and was specified by 3GPP for GSM/3G and UMTS systems. It works at a sampling rate of 16 kHz and can be used at bitrates from 6.6 kbit/s to 23.85 kbit/s.

2.7. mp3 (MPEG-1, MPEG-2 Layer III derivative)

mp3 is a file format which is mainly used for the streaming or downloading of audio files. It is a proprietary system, patented by the Fraunhofer Institute, and is based on the ISO/IEC MPEG Layer III standard. There are several versions of mp3 on the market. For these tests, Opticom's Layer III was used in two different modes, depending on the bitrate. At bitrates of 48 kbit/s and 64 kbit/s, mp3 was used in compliance with the MPEG standard, whereas at the lower bitrates, it was used with the proprietary ultra-low sampling rate extension called "MPEG 2.5" at a sampling frequency of 11 kHz. The latter is not included in the MPEG Audio standard.

2.8. RealNetworks G2

This audio system is used exclusively for live streaming of audio or streaming of audio files.

It was also included in the earlier EBU tests *(see* [1]*)*.

3. Encoding process

For the encoding of the audio items, only software encoders were used. Where complete decoders were available, WAV files were produced directly by the decoders. In the case where only a player was available, "Total Recorder" – a software which redirects the output of the player to a WAV file (a "virtual soundcard") – was used. Total Recorder records streaming audio and stores it on the PC's hard-drive in uncompressed WAV format. For sampling rates other than 44.1 kHz, a software sampling-rate converter, which is part of the SEK'D 2496 package, was introduced.

Table 1 gives the encoding parameters that were set for each of the different codecs during the tests.

Ex. 1055-008

Table 1
Encoding parameters of the different codecs

Codec	Parameter	64 kbit/s stereo	48 kbit/s stereo	32 kbit/s stereo	20 kbit/s stereo	16 kbit/s mono
AAC	Bandwidth (kHz)	11.5*	11.5*	10*	5*	6*
AAC+	F_s (kHz)		44.1			
AMR	Bandwidth (kHz)					6*
mp3	F_s (kHz)	22.05	22.05			
mp3PRO	F_s (kHz)	44.1	44.1			
REAL 8	Bandwidth (kHz) Setting name	16 HiRA8	16 HiRA8	10 HiRA8	5.5 HiRA8	8 Music
REAL G2	Bandwidth (kHz)	15*	12.5*	10*	5.5*	8*
WMA 8	F_s (kHz)	44.1	44.1	32	22.1	22.1

Note: The bandwidth values with an asterisk (*) were measured by spectral analysis of the decoded output signal. Other values are encoder settings.

4. Test material

Three sites collaborated in the selection of the test material. A large selection of test items (sound clips) were coded and decoded. Small groups of expert listeners at the IRT, BBC and T-Systems independently listened to all these items, in carefully-controlled listening environments. Each group produced a short-list of about ten test clips which fulfilled the requirements described in the test method specification. A consolidation of these three short-lists produced a final list of eight sound clips – the minimum requirement for these tests. They are listed in *Table 2*.

Table 2
The eight sound clips used for the subjective tests

Test sequence name	Description	Status
De	German male speech	New
Eu	Dutch female speech with background music	No. 2 in BPN 029
Ff	Fanfare, Orff : Carmina Burana	New
Gr	Pop music	New
Pp	Pop music	No. 8 in BPN 029
Rq	Mozart's Requiem	No. 1 in BPN 029
Sy	Symphony	New
Ve	Female vocal (Susan Vega)	No. 9 in BPN 029

The selection contained a variety of audio types: speech, music, and speech combined with music. Four of the test items had been used in the previous tests [1]. The duration of each sound clip was typically 12 seconds.

The overall test programme was divided into five sessions – each one dedicated to one of the five bitrates being tested *(see Table 1)*. The tests at 16 kbit/s used mono test signals derived from the original stereo signals by downmixing (L+R) and multiplying by 0.707.

5. Experimental design

5.1. Test method

The present tests used the MUSHRA method, described in [2]. This was developed in 1999 by EBU Project Group B/AIM, in collaboration with ITU-R Working Party 6Q. An important feature of this method is the inclusion of the hidden reference and two bandwidth-limited anchor signals (7 kHz and 3.5 kHz).

5.2. Training phase

The purpose of the training phase was to allow each expert listener to achieve two objectives as follows:

- **PART A:** to become familiar with all the sound excerpts under test and their quality-level ranges;
- **PART B:** to learn how to use the test equipment and the grading scale.

The process that was followed, in order to train the listeners prior to the tests, is described in the MUSHRA specification.

5.3. User interface

Implementations of MUSHRA user-interfaces from CRC and the IRT were used in these tests. The BBC and IRT used the CRC SEAQ test software in its MUSHRA operating mode. NRK used software from the IRT.

A screenshot of one implementation of a MUSHRA user-interface is shown in *Fig. 1*.

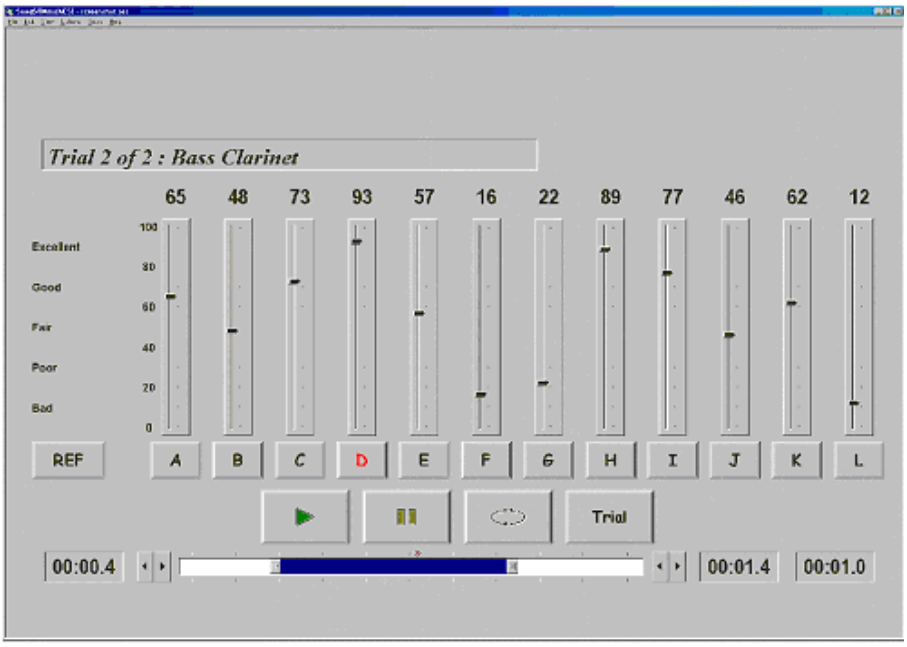


Figure 1
Screenshot of a MUSHRA user-interface implementation

The buttons along the bottom represent the reference source (REF), which is specially displayed top-left, and all the codecs under test (labelled A to L), including the hidden reference and the two bandwidth-limited anchor signals. Above each button, with the exception of the REF button, a slider is used to grade the quality of the test item, according to the continuous MUSHRA quality scale.

For each of the test items, the signals under test were randomly assigned. In addition, the test items were randomised for each subject within a session. To avoid sequential effects, each subject took the five sessions in randomised order.

5.4. Listening panels

The tests were conducted at the sites of three EBU Members:

- **BBC Research & Development** (Kingswood Warren, south of London);
- **NRK** (Oslo);
- **IRT** (Munich).

The benefits of carrying out the tests at multiple sites were significant. On the one hand, the effort could be divided among several partners and, on the other, it was possible to check the reproducibility of the results.

The ten subjects from the BBC, the eleven from the IRT and the nine from NRK (making a total of 30 subjects) were mostly listeners who had some previous experience with listening tests of this type. The results of each individual subject were subsequently analysed, to check for consistency, as described in the MUSHRA specification. The scores of two subjects then had to be discarded because of clear unreliability (for example, assigning a much lower grade to the hidden reference than to a considerably impaired signal). Thus, the final analysis of results is based on the individual results of the 28 remaining subjects.

5.5. Test duration

The training session for each listener took approximately 30 minutes, including an explanation about the tests and equipment, and a practice grading session.

The grading phase consisted of five test sessions, each one containing eight trials with eight to ten test signals, depending on the bitrate. Each session for a particular bitrate took a subject, on average, about 30 minutes. It was generally found that the higher-bitrate tests took longer to complete than the lower-bitrate tests. Subjects were allowed a rest period between each session, but not during a session.

5.6. Listening conditions

The tests at the BBC and NRK used Stax Signature SR-404 headphones. The tests at the IRT used Stax Lambda Pro headphones. The tests at all sites were conducted in rooms with low background noise levels.

Test items were replayed from the following equipment:

NRK: Windows NT PC, IRT software and Digigram PCX11 digital audio interface card;

BBC: Windows NT PC, CRC SEAQ software, Turtle Beach Montego II digital audio interface card, Prism Sound Dream DA-1 DAC.;

IRT: Windows NT PC, CRC SEAQ software, RME Digi96/8 Pro digital audio interface card with external Nvision NV1000 Terminal Equipment DAC.

At NRK, each subject was free to adjust the listening level during their training phase and then it stayed fixed for all their grading sessions. At the BBC and IRT, the subjects were requested to adjust the level at the start of their test but not during a session; at the BBC, no subject found it necessary to change the level.

6. Main results

In the following sections, the main results of each session are described. The statistical analysis method described in the MUSHRA specification was used to process the test data. The results are presented as **mean grades** and **95% confidence intervals**.

Experience has shown that the scores obtained for different test sequences are dependent on the criticality of the test material used. Therefore the mean grades and the 95% confidence intervals have been included, in order to provide a more complete understanding of codec performance – by presenting the results for different test sequences separately, rather than only as aggregated averages across all the test sequences used in the assessment.

For the legend of the different abbreviations used in the figures, please refer to *page 19*.

Ex. 1055-011

6.1. Results for 16 kbit/s mono

The test results for a bitrate of 16 kbit/s per mono signal are given in *Fig. 2*. They show that the sound quality provided by all the codecs, when tested at a bitrate of 16 kbit/s mono, was significantly lower than the subjective quality of the 7 kHz low-pass anchor. Moreover, at this bitrate, only two codecs – RealNetworks G2 and RealNetworks RealAudio 8 – were slightly better than the 3.5 kHz low-pass anchor. The difference between the various codecs was relatively small, with a grade of just under 35 (out of 100) for the best and 15 for the worst. This means that the results for all the codecs, when running at 16 kbit/s mono, were in the regions "bad" and "poor" *(see Fig. 1)*.

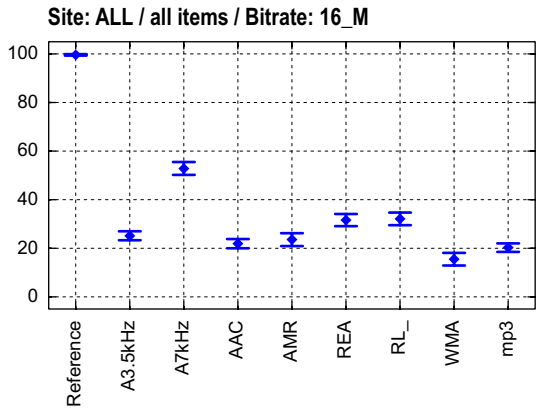


Figure 2
Mean and 95% confidence interval at an audio bitrate of 16 kbit/s mono, hidden reference, bandwidth-limited references, all codecs, 28 subjects – aggregate values averaged over all items

However, when looking in detail at the results – in particular, the individual sound-clip results for each codec – it became obvious that there were some differences among the codecs. For example, the AMR wide-band speech codec gave a "fair" result for speech-like audio signals, even at a bitrate of 16 kbit/s, whereas the same codec produced a "bad" result with the more music-like signals. However, none of the other codecs in this test showed a quality grade better than "poor", for any audio example, which can be seen from *Fig. 3*.

It should be noted that, when compared to the EBU B/AIM tests in 1999/2000 *(see* [1]*)*, the new audio codecs showed no improvement in audio quality at this low bitrate.

If a reasonable audio quality is required, a bitrate higher than 16 kbit/s should be used for generic audio material.

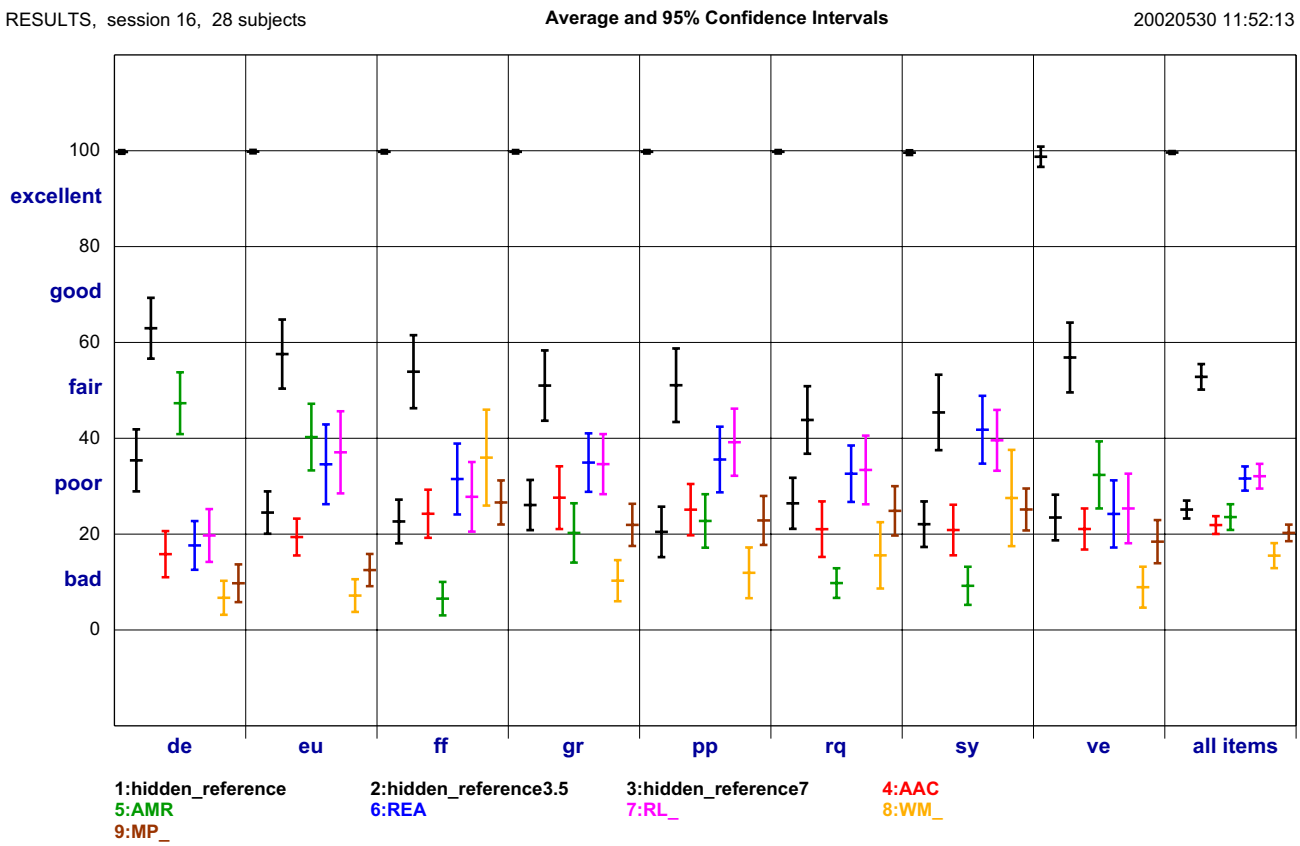


Figure 3
Mean and 95% confidence interval at an audio bitrate of 16 kbit/s mono, hidden reference, bandwidth-limited references, all test items, all codecs, 28 subjects

6.2. Results for 20 kbit/s stereo

The results for a bitrate of 20 kbit/s stereo are given in *Fig. 4*. They show that the quality provided by all the tested codecs was still significantly lower than the subjective quality of the 7 kHz low-pass anchor.

Compared to the results with 16 kbit/s mono, the quality of the 20 kbit/s stereo codecs was even lower. With the exception of one codec (RealNetworks RealAudio 8), the results from all the codecs were worse than that of the 3.5 kHz low-pass anchor at 20 kbit/s stereo. In fact, the subjective difference between the low-pass filtered anchors at 16 kbit/s mono and 20 kbit/s stereo was found to be negligible.

The individual results of all test signals at 20 kbit/s stereo were in the two regions "bad" and "poor", which can be seen from *Fig. 5*. As in the case of the 16 kbit/s mono results, it should be noted that, when compared to the EBU B/AIM tests in 1999/2000 *(see* [1]*)*, there was no improvement in audio quality at 20 kbit/s stereo across the spectrum of available audio codecs.

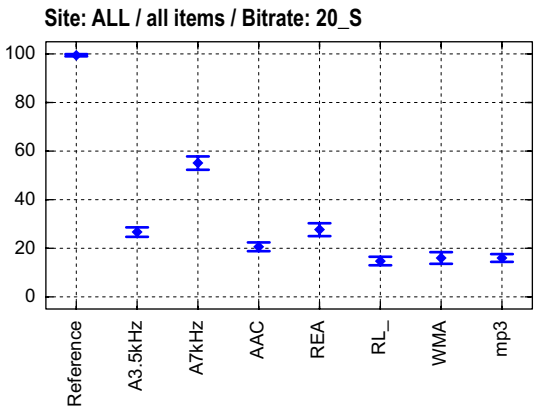


Figure 4
Mean and 95% confidence interval at an audio bitrate of 20 kbit/s stereo, hidden reference, bandwidth-limited references, all codecs, 27 subjects

If a reasonable stereophonic audio quality for generic audio material is required, a bitrate of 20 kbit/s or less should definitely be avoided.

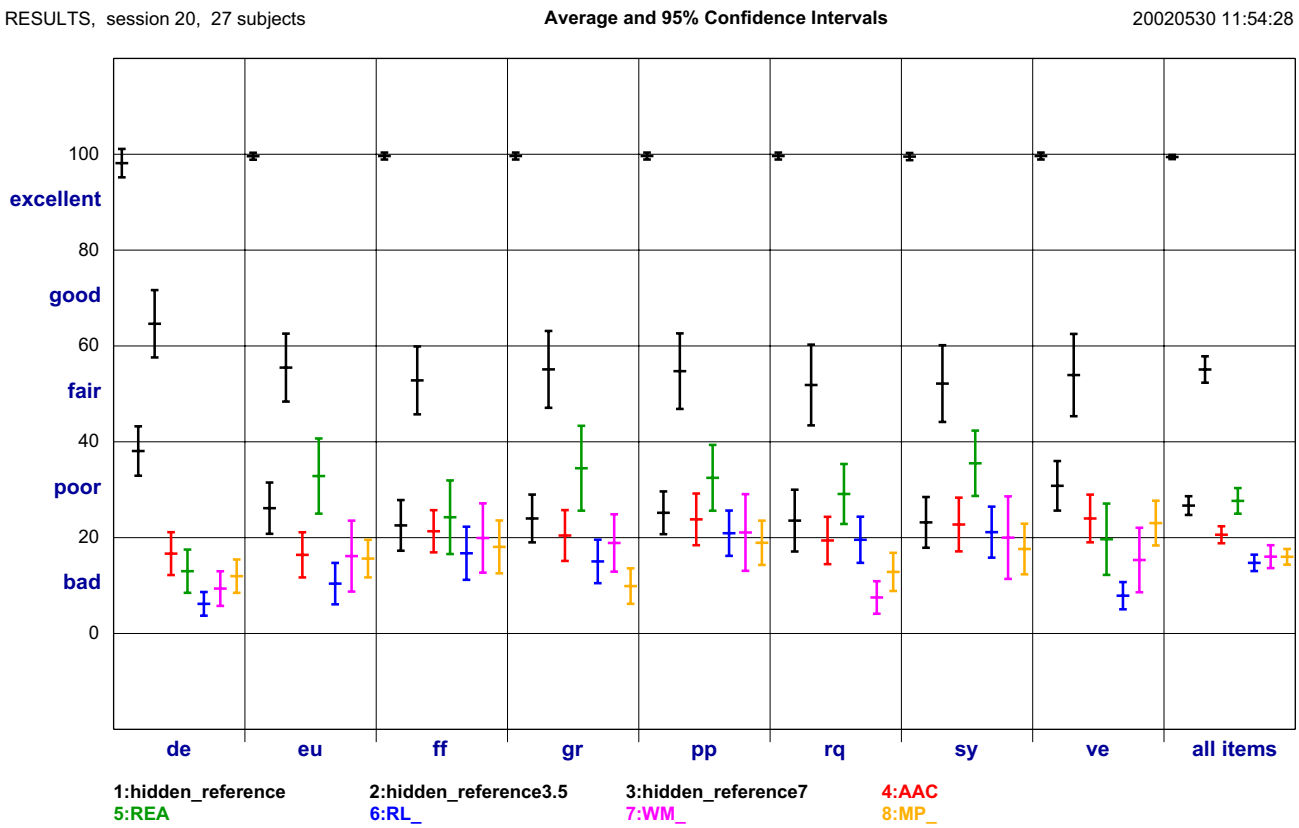


Figure 5
Mean and 95% confidence interval at an audio bitrate of 20 kbit/s stereo, hidden reference, bandwidth-limited references, all test items, all codecs, 27 subjects

6.3. Results for 32 kbit/s stereo

The results for a bitrate of 32 kbit/s stereo are given in *Fig. 6*. The most obvious result here is that the differences between the various codecs has become more pronounced. The difference between the best and the worst codec was about 30 points on a 100-point scale, compared to not much more than 10 points at 16 kbit/s mono. [It is interesting to note that the same effect occurred during the earlier EBU B/AIM tests *(see* [1]*)*, with a spread of about 25 points at 32 kbit/s stereo.]

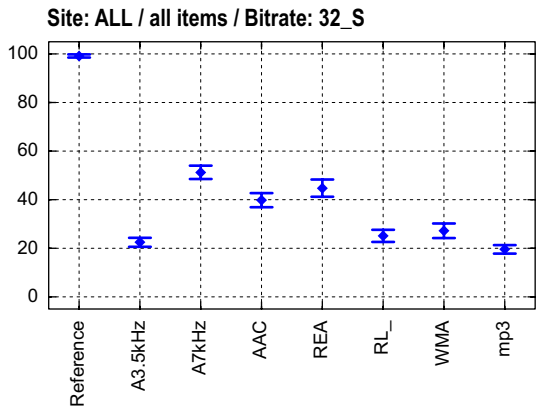


Figure 6
Mean and 95% confidence interval at an audio bitrate of 32 kbit/s stereo, hidden reference, bandwidth-limited references, all codecs, 28 subjects

The two best codecs in these tests, i.e. RealNetworks RealAudio 8 and MPEG-2 AAC, provided a quality level close to that of the 7 kHz low-pass anchor. The other codecs were no better than the 3.5 kHz low-pass anchor. This means that only the two best codecs were able to offer a quality in the region of "poor" and "fair".

RealAudio 8 and AAC showed a rather balanced behaviour between the results of the different test signals *(see Fig. 7)*. This means that practically none of the test items could "kill" these two codecs – with the (marginal) exception of German male speech, in the case of the RealAudio 8 codec. Such a balanced behaviour is exactly what the user, i.e. the content or service provider, expects.

The results for the hidden reference indicate that the subjects were always able to identify it correctly. This is only possible if there is a significant difference between the quality level of the codecs under test and that of the hidden reference.

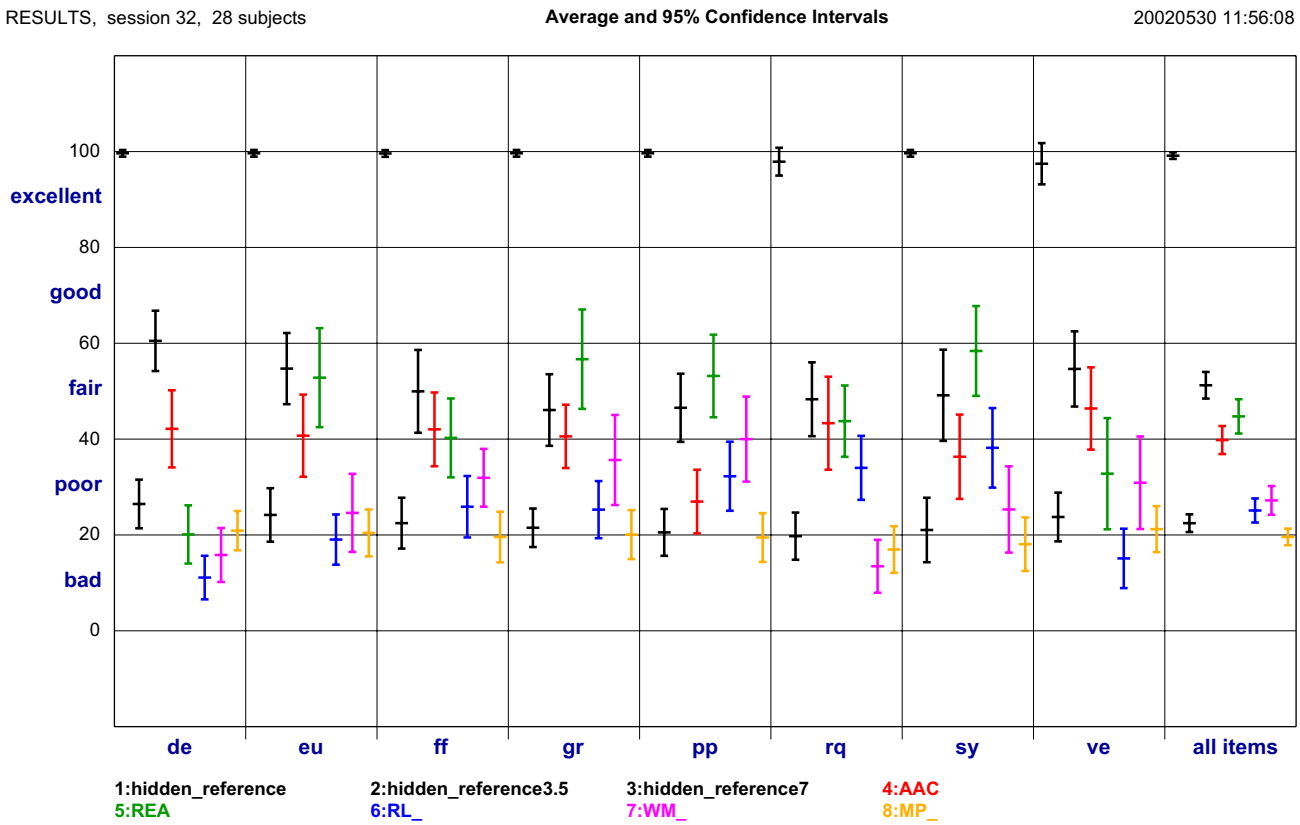


Figure 7
Mean and 95% confidence interval at an audio bitrate of 32 kbit/s stereo, hidden reference, bandwidth-limited references, all test items, all codecs, 28 subjects

Generally speaking, we can conclude that, when compared to the tests reported in [1], there has been no quality improvement in the new generation of audio codecs at this bitrate.

32 kbit/s seems to be the lowest bitrate limit for obtaining a quality level that is better than "bad", for generic stereophonic audio signals.

6.4. Results for 48 kbit/s stereo

The results of the tests at 48 kbit/s stereo are given in *Fig. 8*. It shows a clear winner: MPEG-2/4 AAC plus SBR, nowadays called **CTaacPlus**. This codec showed an average result, from all the test items, of between "good" and "excellent". On the other hand, *Fig. 8* also shows a "poor" rating for the worst codec in these tests – which was even worse than the quality of the 7 kHz low-pass filtered anchor.

The difference between the best and worst codecs at 48 kbit/s was now 50 points on a 100-point scale, which is remarkable. Two of the codecs, i.e. Windows Media 8 and mp3, showed a quality that was comparable to the quality of the 7 kHz low-pass filtered anchor.

The second best codec, i.e. mp3PRO, showed a mean result in the region of "good". MPEG-2/4 AAC came close with a result between "fair" and "good".

Fig. 9 shows that the three best codecs produced a fairly balanced

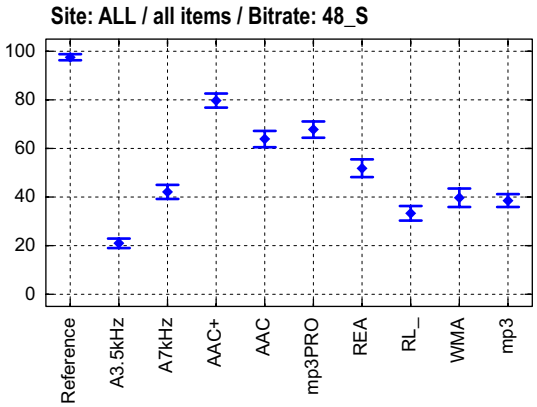


Figure 8
Mean and 95% confidence interval at an audio bitrate of 48 kbit/s stereo, hidden reference, bandwidth-limited references, all codecs, 28 subjects

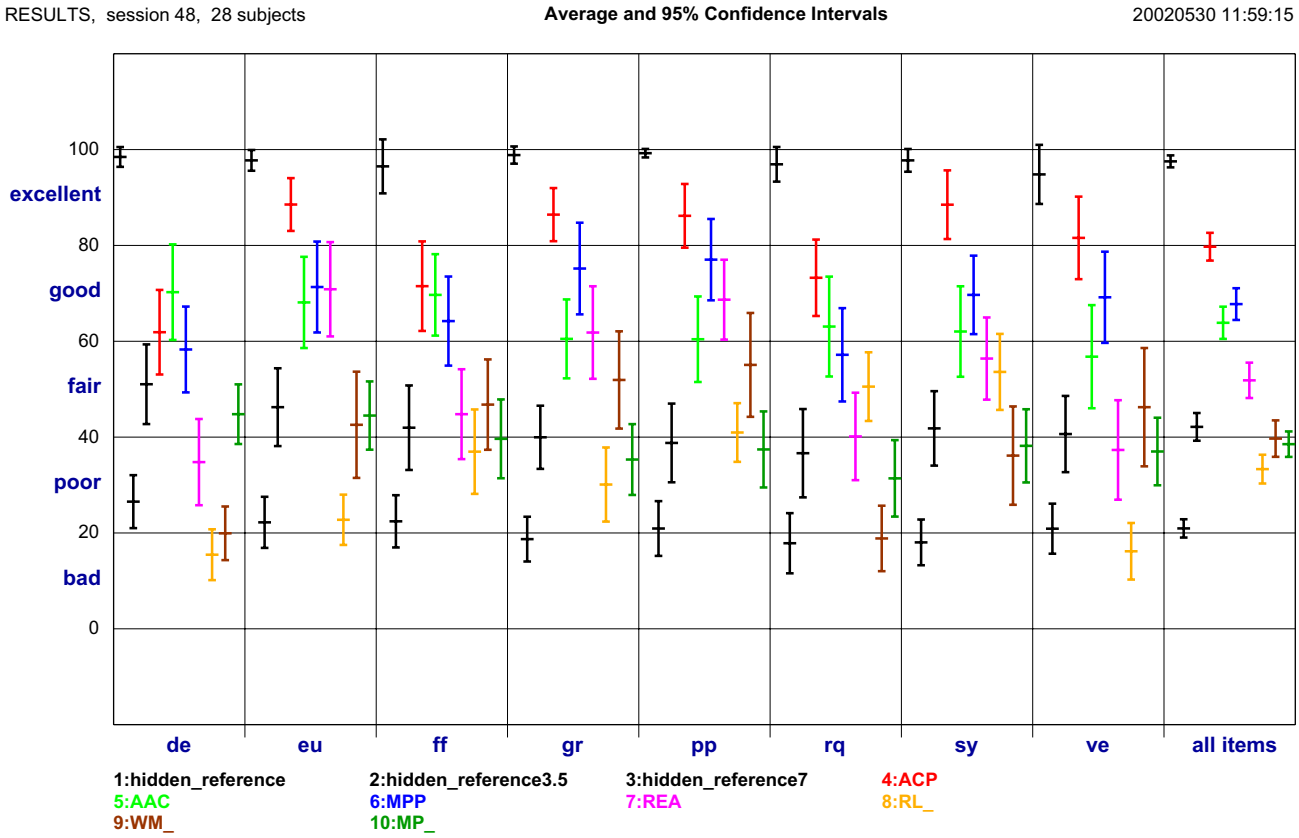


Figure 9
Mean and 95% confidence interval at an audio bitrate of 48 kbit/s stereo, hidden reference, bandwidth-limited references, all test items, all codecs, 28 subjects

result between the different test items, with no particularly critical item.

Compared to the earlier tests reported in [1], it becomes obvious that the new generation of codecs which use the spectral band replication (SBR) technique produce a significantly better quality when compared to the best codecs in the earlier tests (which were MPEG2/4 AAC and mp3).

48 kbit/s offers the programme or service provider the possibility to stream or deliver generic stereophonic audio signals with good and even excellent quality.

6.5. Results for 64 kbit/s stereo

The results for a bitrate of 64 kbit/s stereo are shown in *Fig. 10*.

It should be pointed out that CTaacPlus could not be included in this test. Instead, the second best codec in the 48 kbit/s stereo tests (i.e. mp3PRO) became the best codec in the 64 kbit/s stereo tests. mp3PRO showed a clear lead of about 20 points (which corresponds to one quality category) over the remaining codecs. All the other codecs produced similar scores, with quality assessments between "fair" and "good". The worst codec at this bitrate (i.e. mp3) was only in the "fair" category.

The quality range between the best and worst codec at this bitrate has shrunk to about 30 points, which is significantly less than for the assessments made at 48 kbit/s

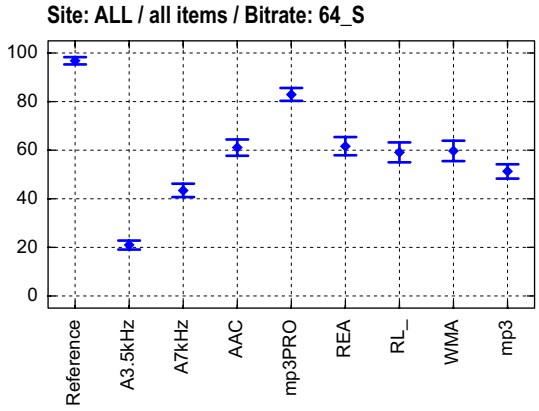


Figure 10
Mean and 95% confidence interval at an audio bitrate of 64 kbit/s stereo, hidden reference, bandwidth-limited references, all codecs, 28 subjects

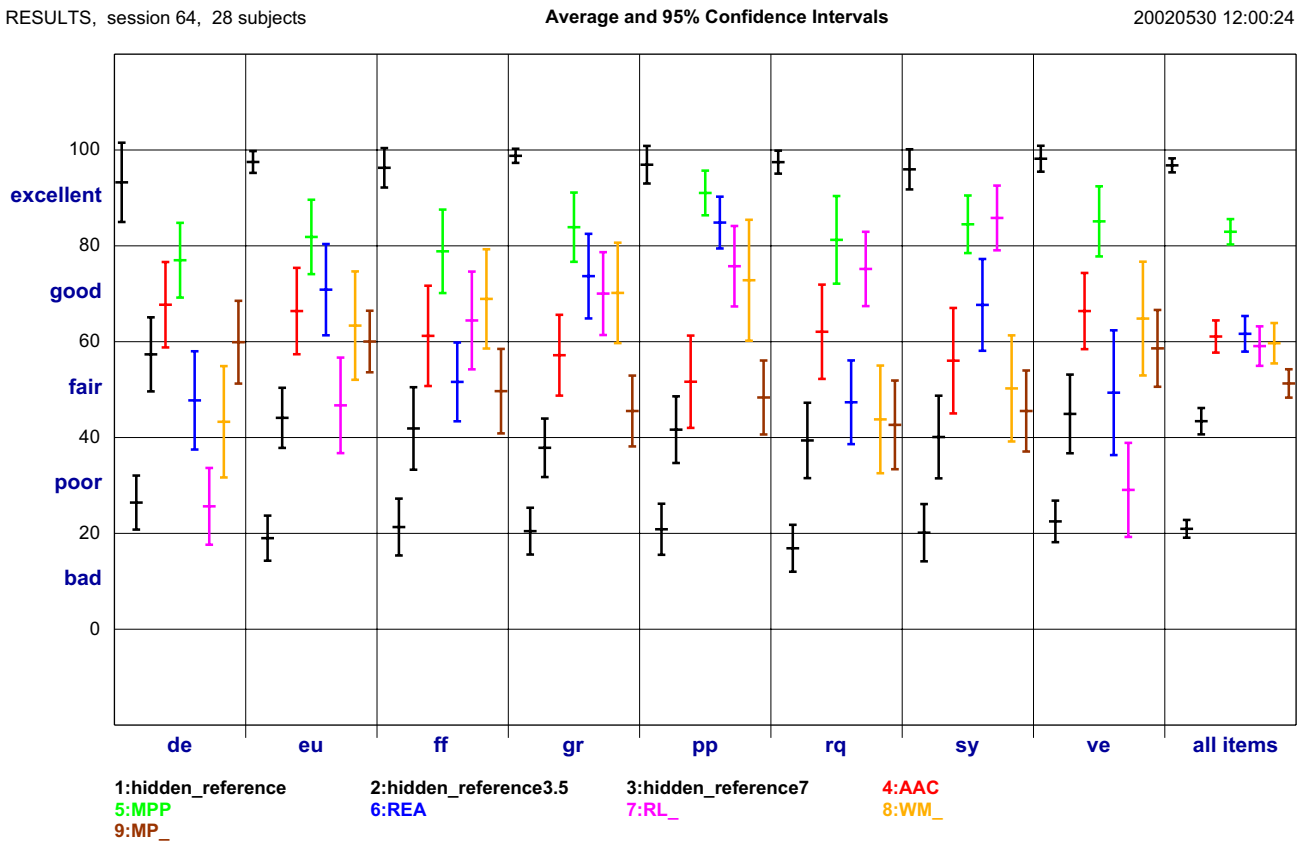


Figure 11
Mean and 95% confidence interval at an audio bitrate of 64 kbit/s stereo, hidden reference, bandwidth-limited references, all test items, all codecs, 28 subjects

stereo (where the quality range spanned 50 points). This means that the difference between the codecs is less pronounced today than it was in the earlier EBU tests *(see* [1]*)* when it was more than 40 points. In the present tests at 64 kbit/s stereo, each codec showed a better performance than the quality of the 7 kHz low-pass filtered anchor: in the 1999 tests, only one codec (MPEG-2/4 AAC) was judged to be better than the 7 kHz anchor.

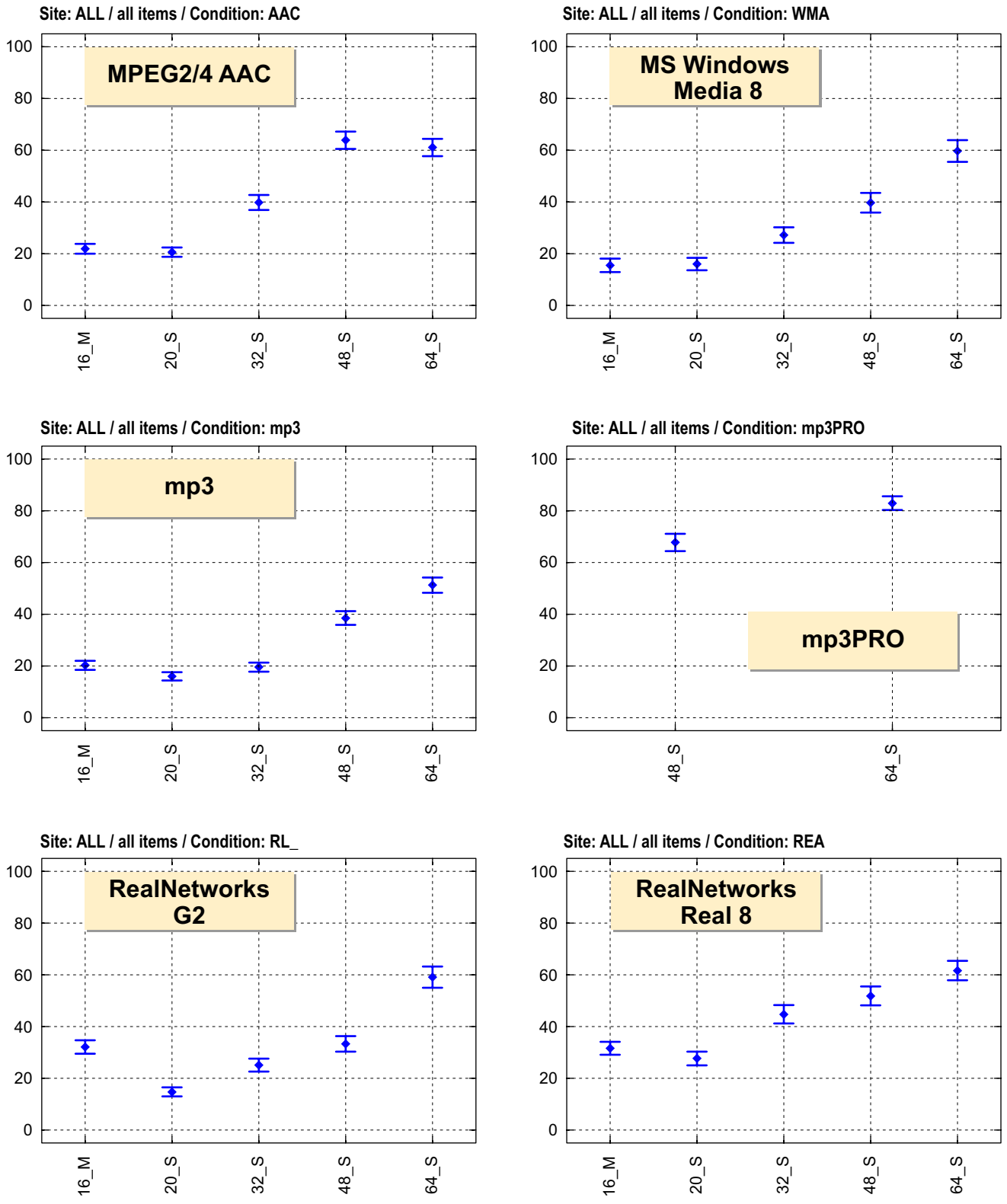


Figure 12
Mean and 95% confidence interval for all the audio test items, excluding the hidden reference and bandwidth-limited references, at different audio bitrates for the six codecs tested, 28 subjects

Thus it can be concluded that, in the last two years or so, most of the new codecs have improved the subjective quality at 64 kbit/s stereo. However, as we have seen before, this is certainly not the case at bitrates equal to or less than 32 kbit/s.

Fig. 11 shows that mp3PRO exhibited a very balanced behaviour across the results for different test items (this behaviour was similar to that at 48 kbit/s). In most cases this meant "excellent". Only two items fell in the category "good". Other codecs, in particular RealNetworks G2, showed a rather unbalanced behaviour across the different sequences, the quality varying by as much as two categories.

As noted above, the 64 kbit/s version of CTaacPlus was not available for these tests. Nevertheless, as CTaacPlus performed so well at 48 kbit/s, it can be expected that its performance might have been even better than that of mp3PRO at 64 kbit/s.

Programme and service providers can deliver an excellent quality for generic stereophonic audio signals – if using CTaacPlus or mp3PRO at a stereo bitrate of 64 kbit/s. This bitrate is convenient, in practice, as it fits into one B channel of an ISDN connection.

6.6. Codec performance as a function of bitrate

Generally it can be expected that, by increasing the bitrate, the performance of codecs will improve and vice versa. Regarding the B/AIM tests in 1999/2000 *(see* [1]*)*, this was mostly true, except in two cases:

- Increasing the bitrate from 16 kbit/s to 20 kbit/s did not result in a better quality. In fact, quite the opposite happened, due to going from monophonic audio signals at a bitrate of 16 kbit/s to stereophonic audio signals at 20 kbit/s.
- One codec in the earlier tests, the Q-Design Music-Codec 2, did not perform better with increasing bitrates. In fact, it showed a remarkable quality at 20 kbit/s (in particular, for music), but did not improve very much at the higher bitrates *(see* [1]*)*. However, the Q-Design Music-Codec 2 had very unbalanced results for different types of audio signals.

Fig. 12 shows the mean value and 95%-confidence interval – averaged over all the audio test items – for the codecs tested in 2002. It can be seen, again, that increasing the bitrate from 16 kbit/s to 20 kbit/s, and going from mono to stereo at the same time, does not lead to an increase in the audio quality. For most of the new-generation codecs, the quality was at least maintained. However, the older generation of codecs, e.g. RealNetworks G2 and mp3, showed a clear decrease in quality when going from 16 kbit/s mono to 20 kbit/s stereo. One reason for such a different behaviour could be the better stereophonic coding techniques used by the newer codec generation.

All the codecs showed an increasing quality with bitrate when progressing from 20 kbit/s to 64 kbit/s. The only exception was the MPEG-2/4 AAC codec, which did not provide a better quality when increasing the bitrate from 48 kbit/s to 64 kbit/s. According to the information received from the codec developers, the version under test had not been optimized for bitrates higher than 48 kbit/s.

7. References

[1] BPN 029: **EBU Report on the Subjective Listening Tests of Some Commercial Internet Audio Codecs** Contribution from EBU Project Group B/AIM, October 2000 (available to EBU members only).

[2] B/AIM 087: Draft EBU contribution to ITU-R 6Q meeting (March 2002) on BS.1534: Draft new Recommendation ITU-R BS.[DOC. 6/106] – **Method for the subjective assessment of intermediate audio quality**
(Question ITU-R 220/10) – BMC 767

[3] ITU-R Recommendation BS.1534 (June 2001): **Method for the subjective assessment of intermediate quality level of coding systems**
Approved in accordance with Resolution ITU-R 45. Status: In force.

Ex. 1055-018

[4] ITU-R Recommendation BS.1116-1: **Methods for the subjective assessment of small impairments in audio systems including multi-channel sound systems**.

[5] Martin Dietz and Stefan Meltzer: **CT-aacPlus — a state-of-the-art audio coding system**
EBU Technical Review No. 291, July 2002.
http://www.ebu.ch/trev_291-dietz.pdf

[6] Jonathan Stott: **DRM — key technical features**
EBU Technical Review No. 286, March 2001.
http://www.ebu.ch/trev_286-stott.pdf

Appendix A: The detailed results

The figures in this Appendix present all the detailed results, averaged over all three test sites, as follows:

Pages	Title	y-axis	x-axis	Parameters
20 - 21	Overview	Continuous quality scale 0 - 100	a) Codecs b) Bitrates	a) Bitrates b) Codecs
22 - 27	Codec / bitrate	Continuous quality scale 0 - 100	Items	Codecs, Bitrates
28 - 38	Codec / item	Continuous quality scale 0 - 100	Bitrates	Codecs, Items
39-43	Item / bitrate	Continuous quality scale 0 - 100	Codecs	Bitrates, Items

Bitrates:

16 kbit/s mono
20 kbit/s stereo
32 kbit/s stereo
48 kbit/s stereo
64 kbit/s stereo

Items:

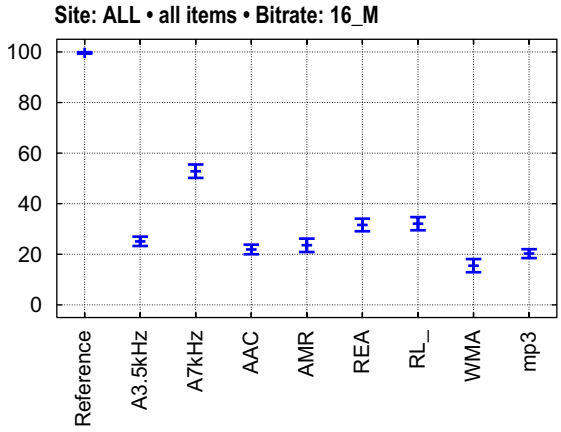
De	German male speech
Eu	Dutch female speech with background music
Ff	Fanfare
Pp	Pop music
Rq	Requiem music
Sy	Symphony music
Ve	Pop music (Susan Vega)

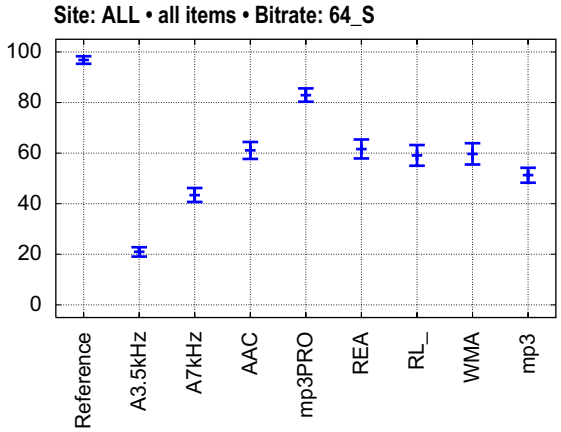
Codecs:

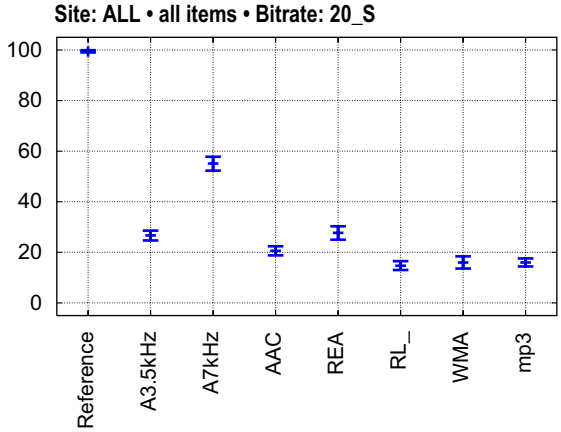
Anchor3.5	3.5 kHz filtered anchor
Anchor7	7.0 kHz filtered anchor
AAC	MPEG-2 AAC
AAC+	CTaacPlus
AMR	AMR Wide Band
mp3	Old mp3
mp3PRO	Thomson Multimedia
REA	RealAudio 8
RL_	RealAudio G2
WMA	Microsoft Media 8

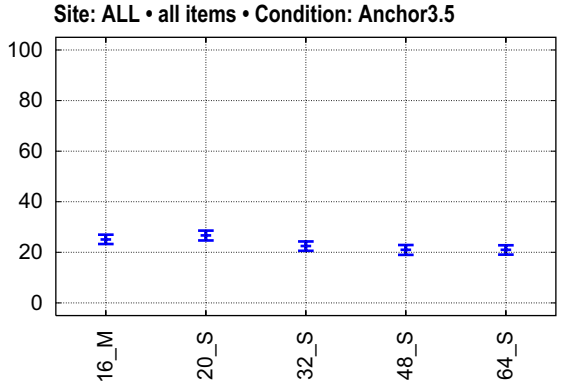
Ex. 1055-020

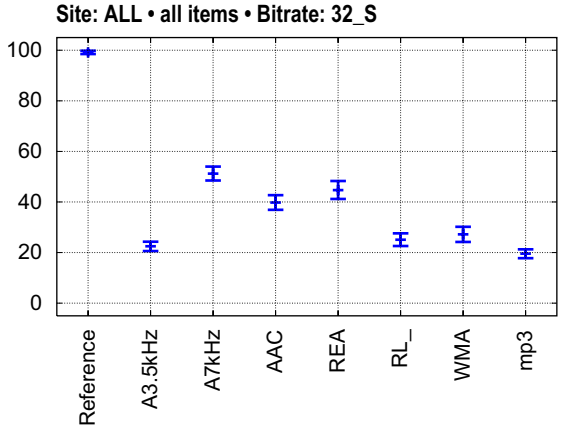
Appendix A — Overview

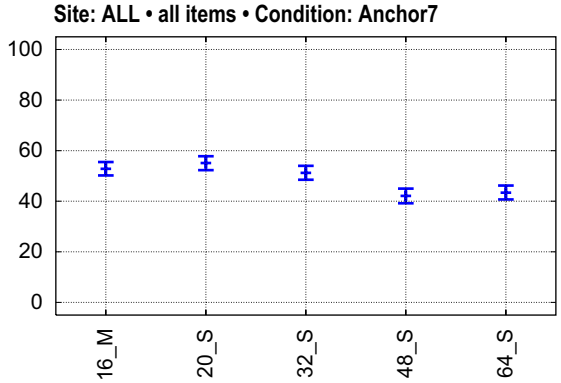


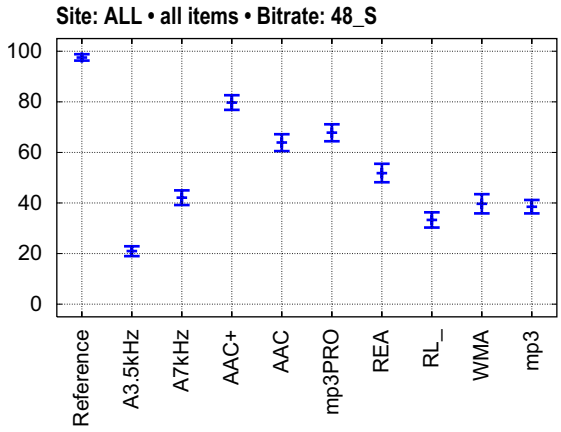


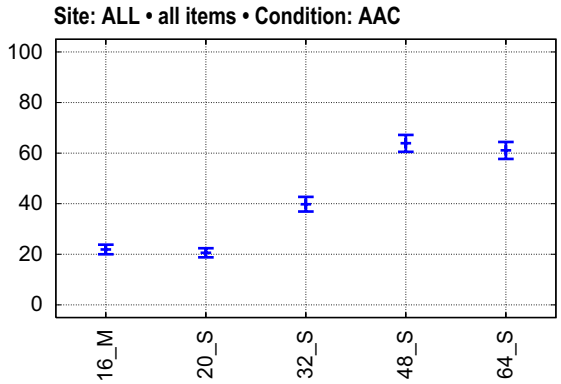




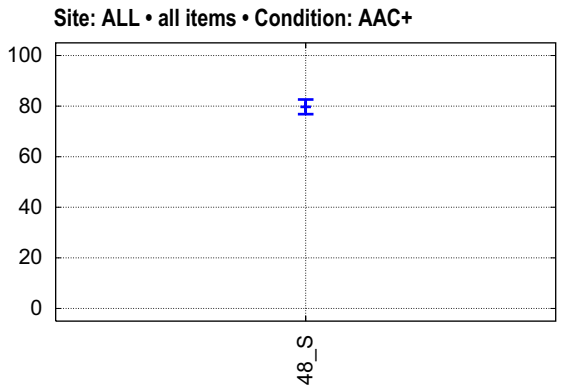


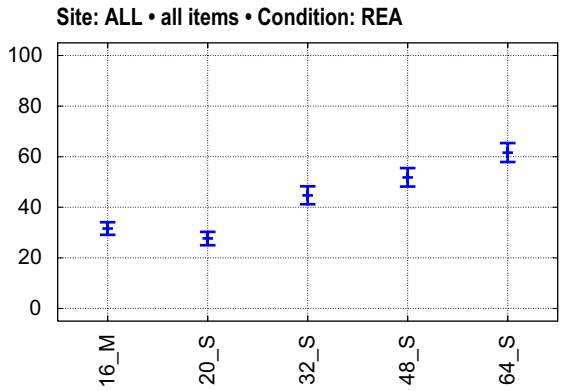


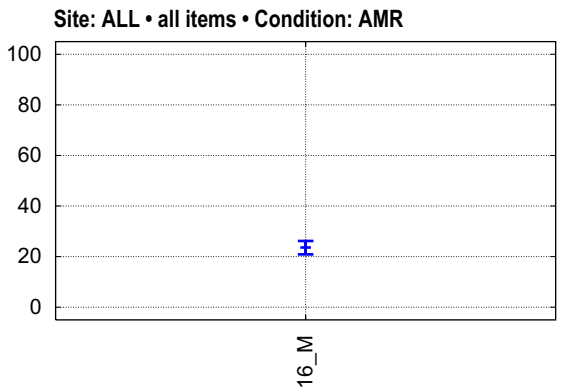


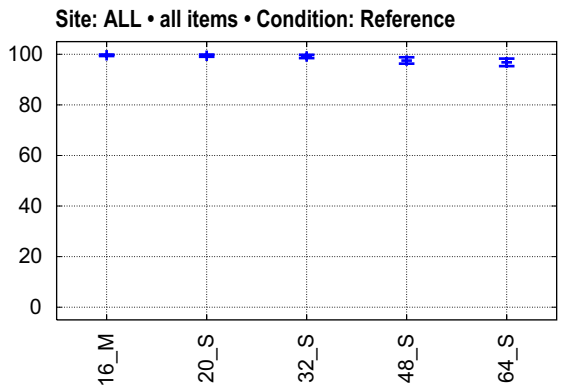


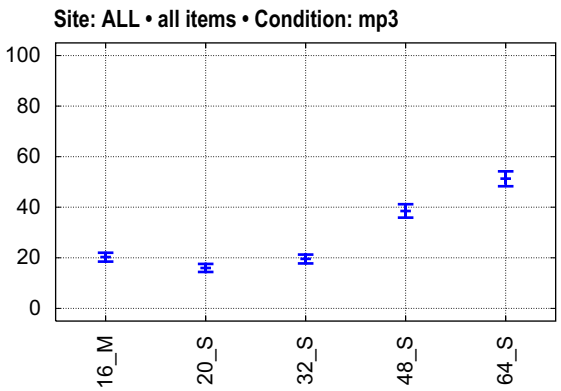
Appendix A — Overview

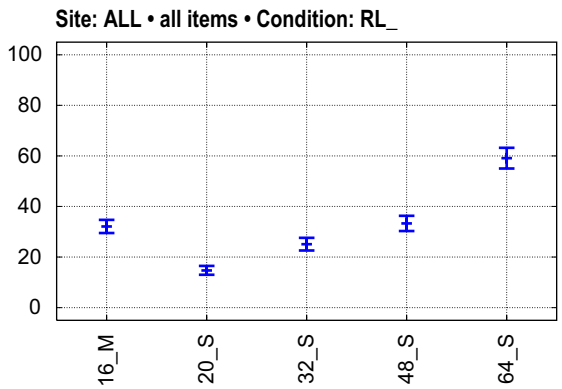


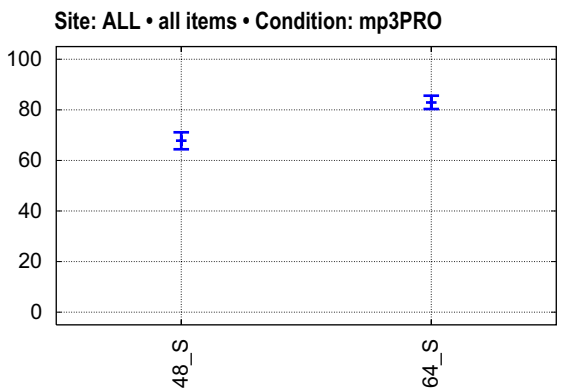


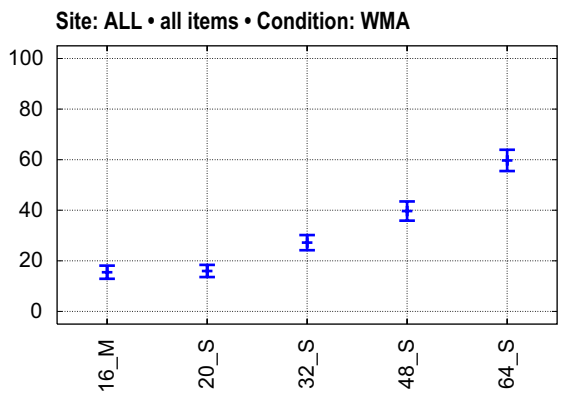






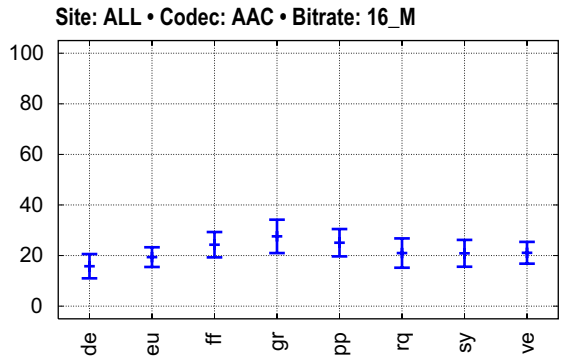


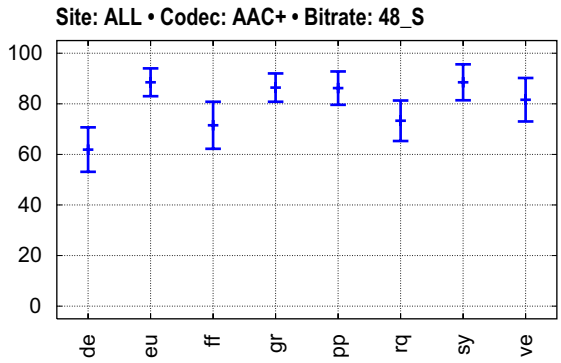


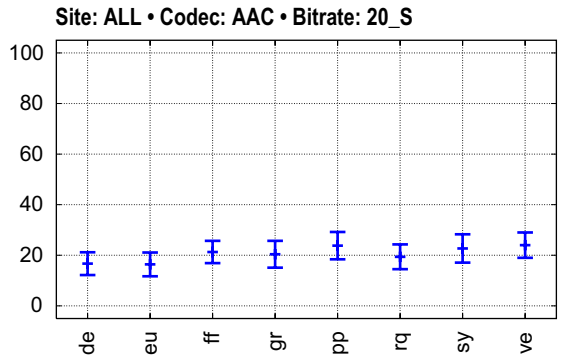


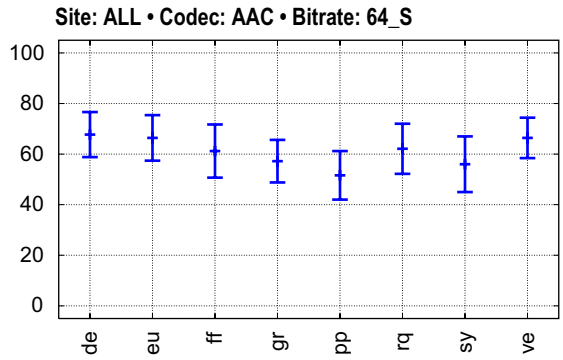
Ex. 1055-022

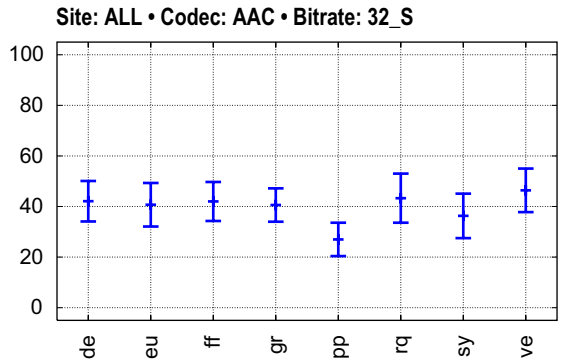
Appendix A — Codec / Bitrate

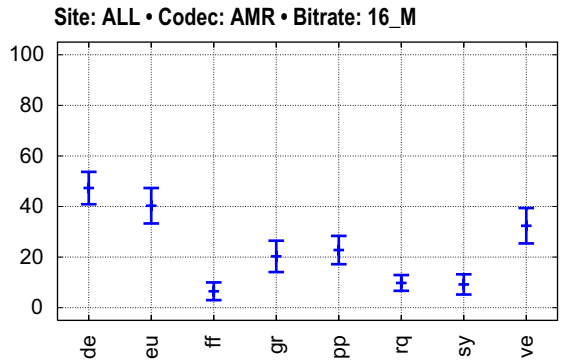


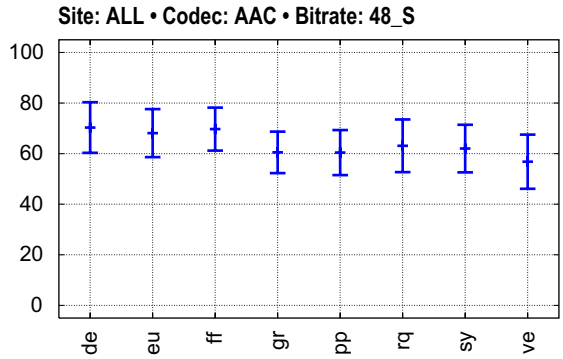


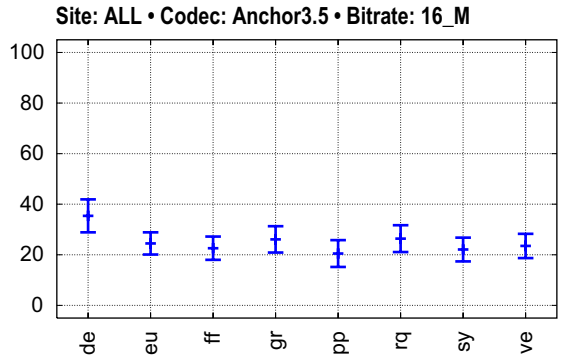




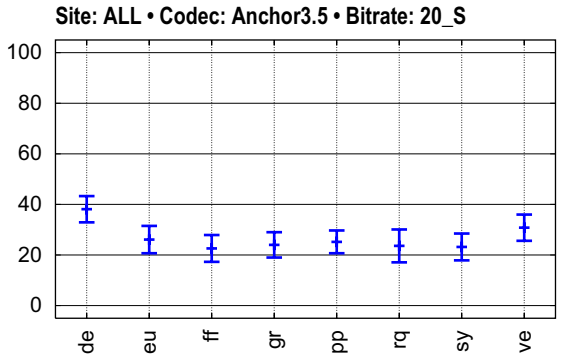


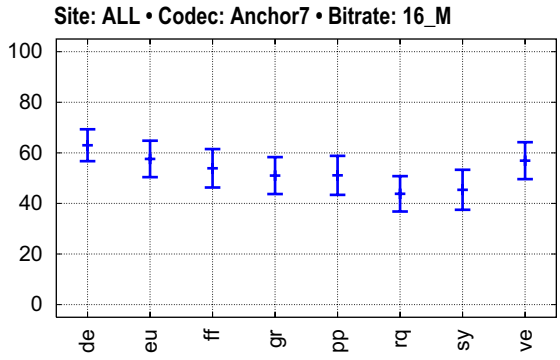


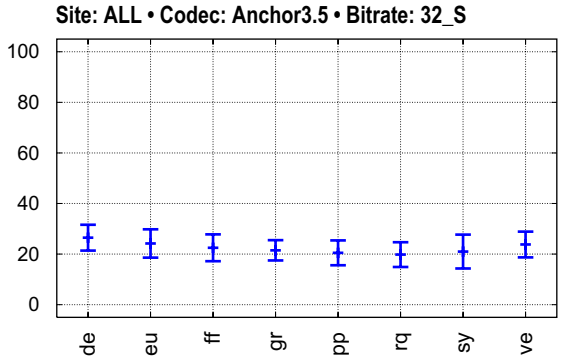


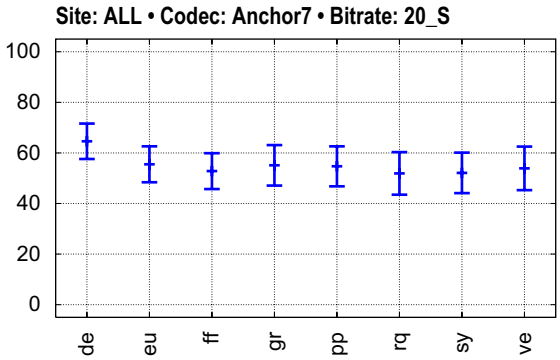


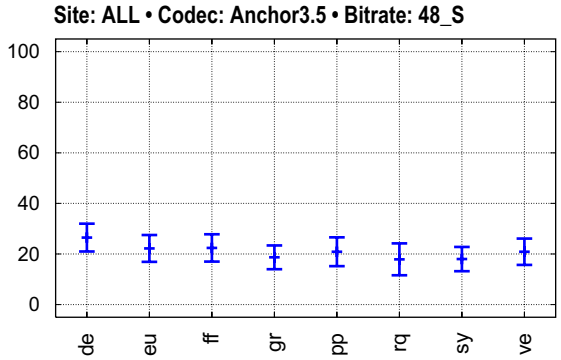
Appendix A — Codec / Bitrate

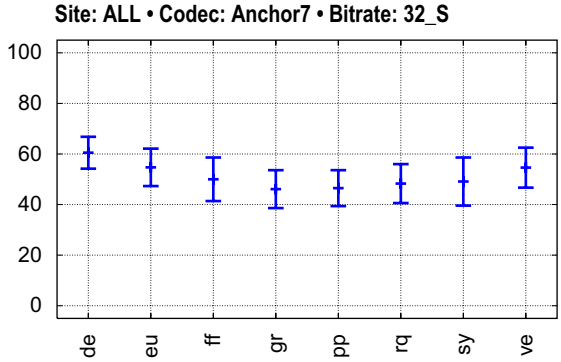


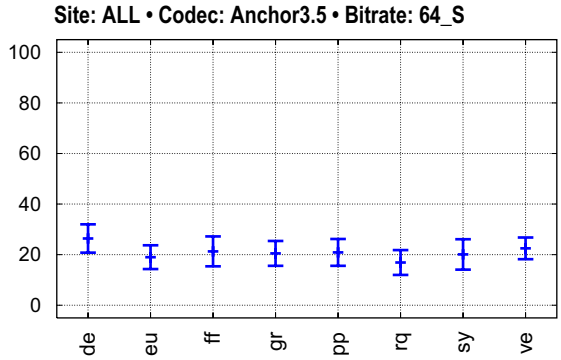


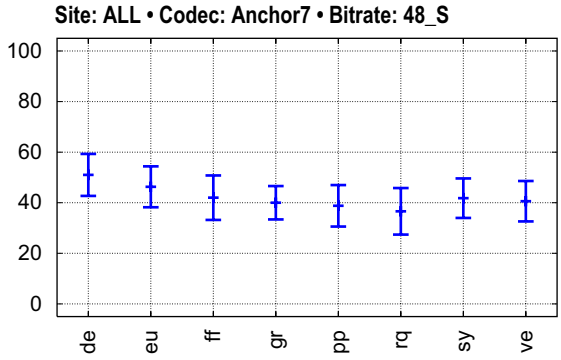








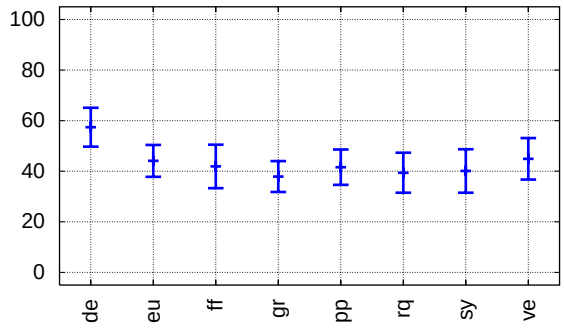




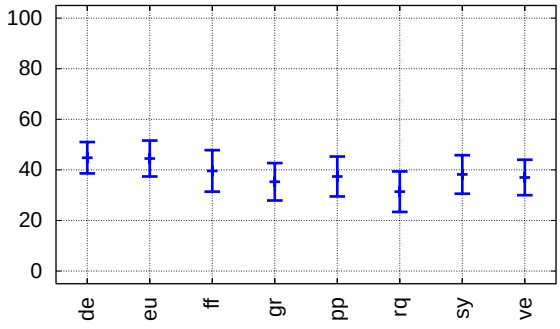
Ex. 1055-024

Appendix A — Codec / Bitrate

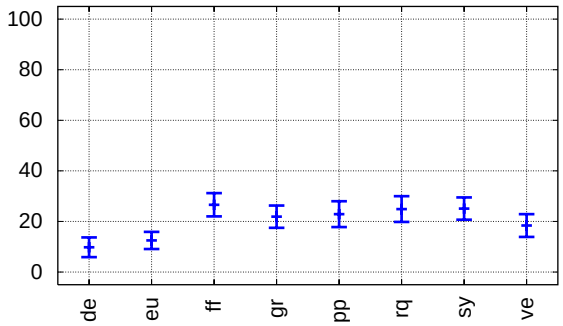
Site: ALL • Codec: Anchor7 • Bitrate: 64_S



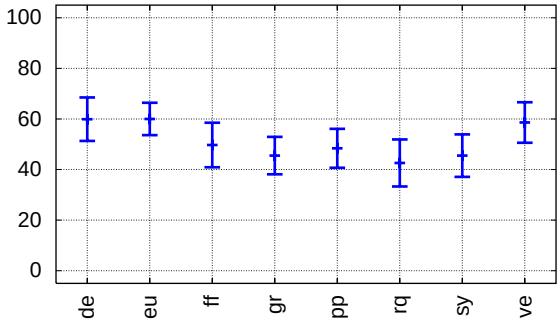
Site: ALL • Codec: mp3 • Bitrate: 48_S



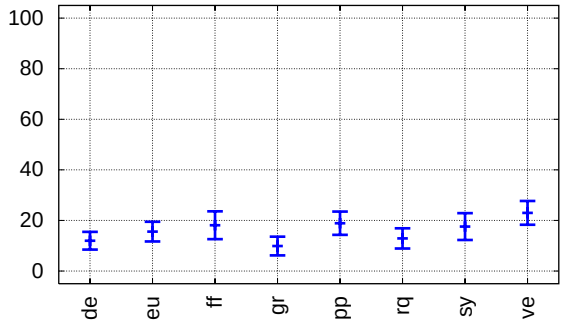
Site: ALL • Codec: mp3 • Bitrate: 16_M



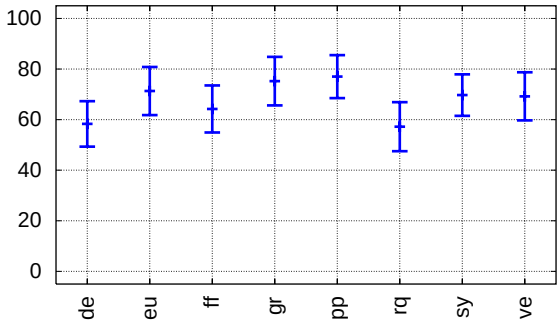
Site: ALL • Codec: mp3 • Bitrate: 64_S



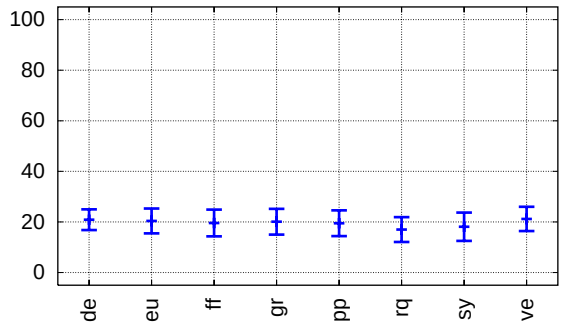
Site: ALL • Codec: mp3 • Bitrate: 20_S



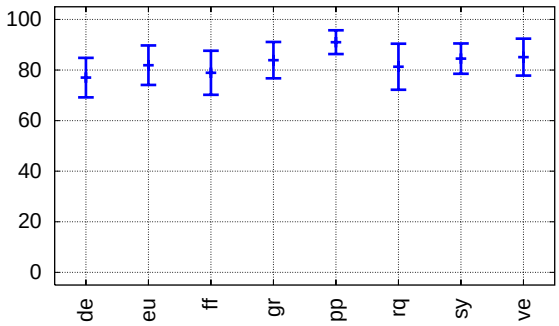
Site: ALL • Codec: mp3PRO • Bitrate: 48_S



Site: ALL • Codec: mp3 • Bitrate: 32_S

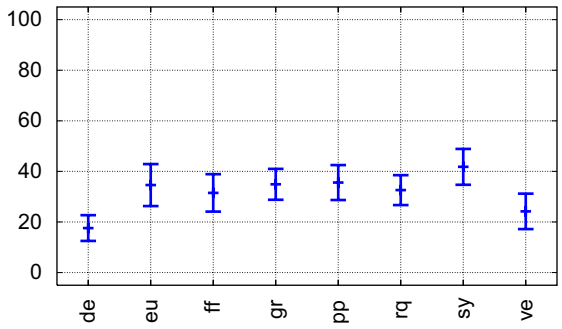


Site: ALL • Codec: mp3PRO • Bitrate: 64_S

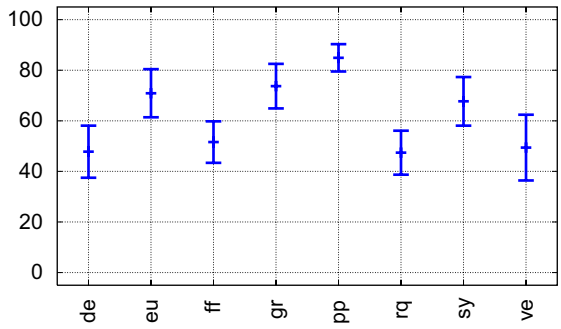


Appendix A — Codec / Bitrate

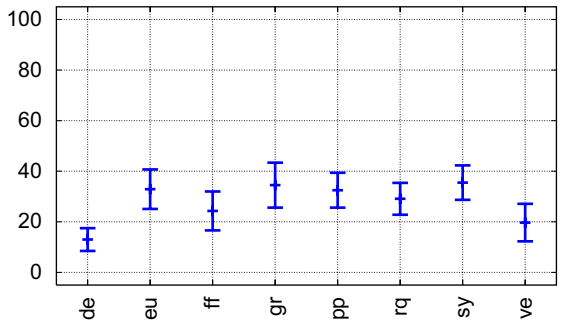
Site: ALL • Codec: REA • Bitrate: 16_M



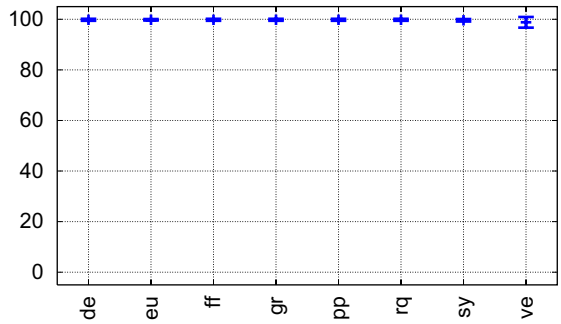
Site: ALL • Codec: REA • Bitrate: 64_S



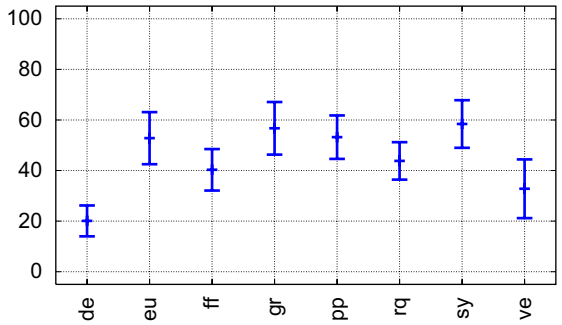
Site: ALL • Codec: REA • Bitrate: 20_S



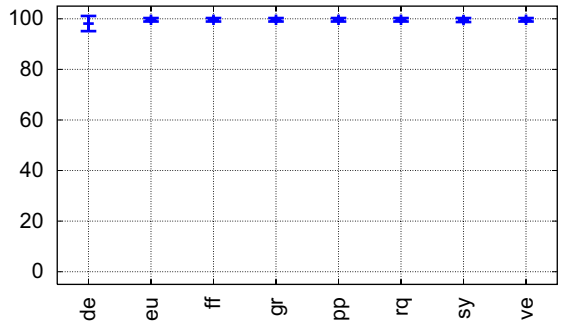
Site: ALL • Codec: Reference • Bitrate: 16_M



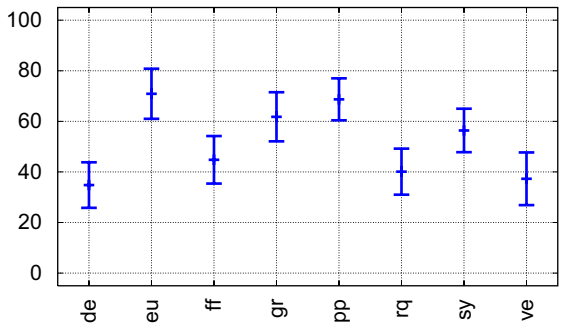
Site: ALL • Codec: REA • Bitrate: 32_S



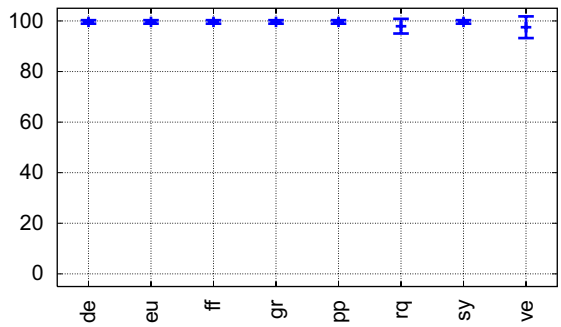
Site: ALL • Codec: Reference • Bitrate: 20_S



Site: ALL • Codec: REA • Bitrate: 48_S

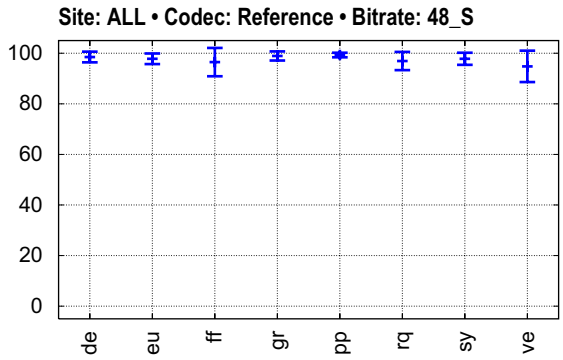


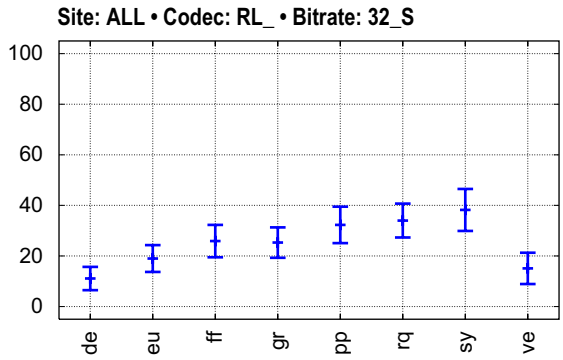
Site: ALL • Codec: Reference • Bitrate: 32_S

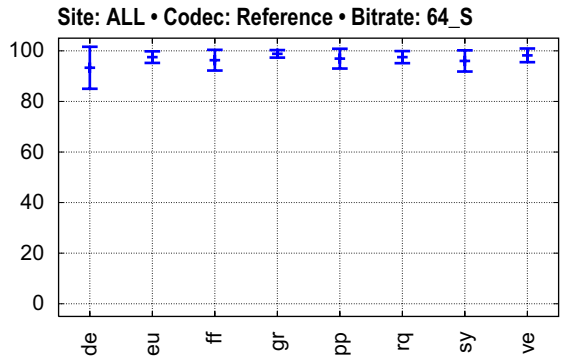


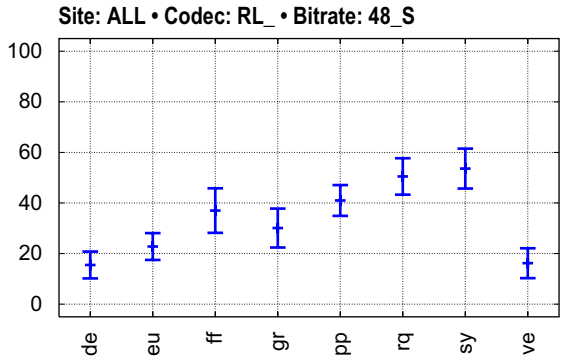
Ex. 1055-026

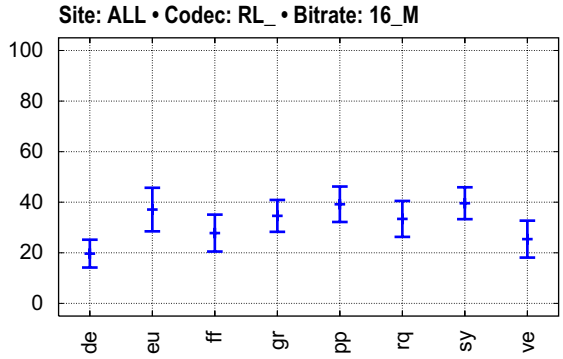
Appendix A — Codec / Bitrate

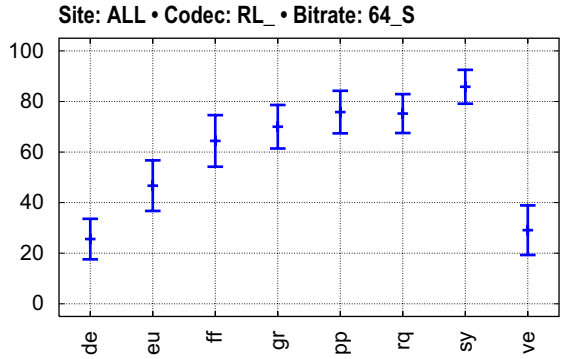


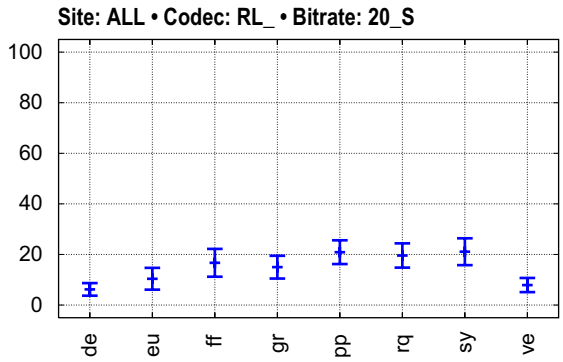


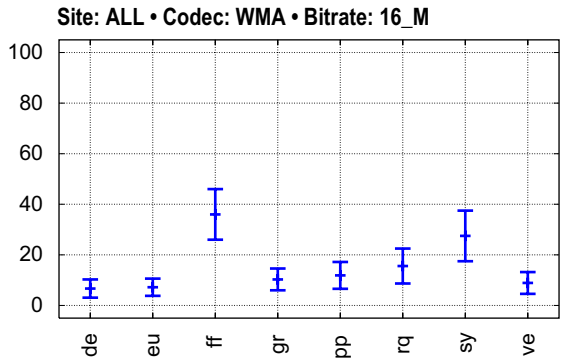




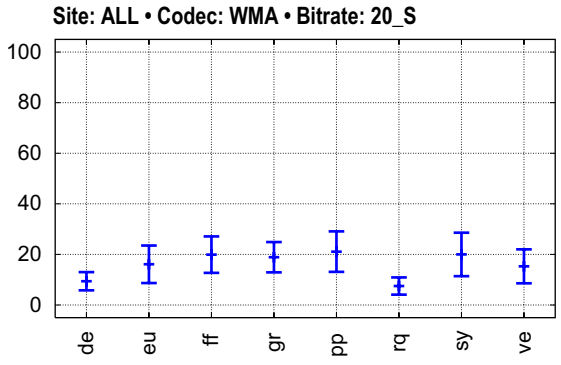


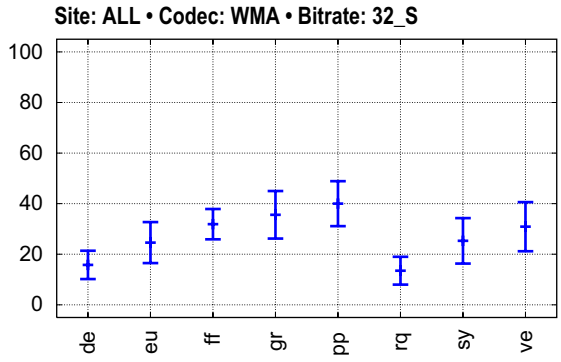


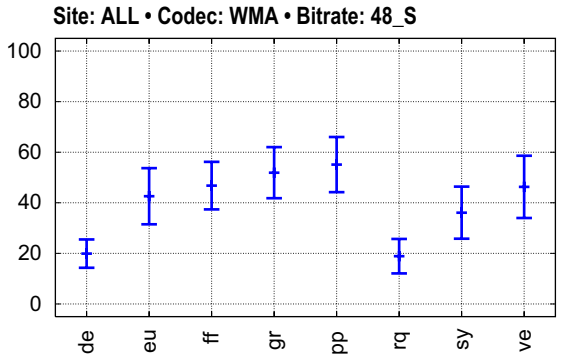


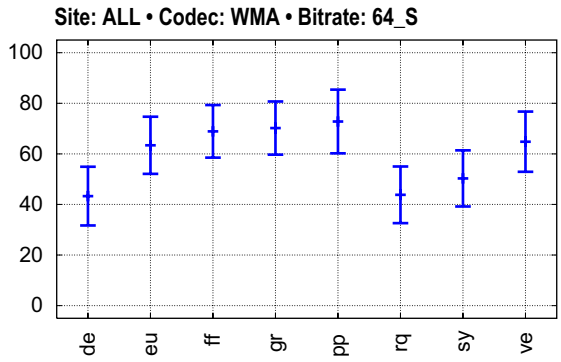


Appendix A — Codec / Bitrate

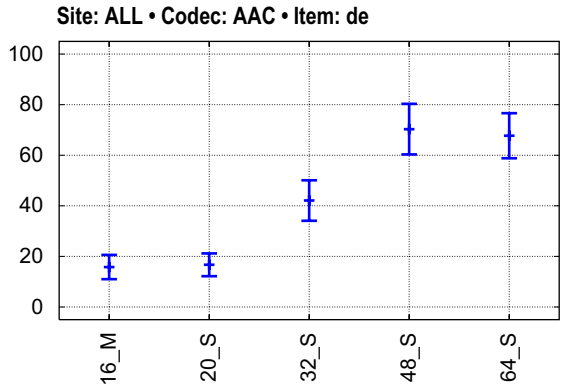


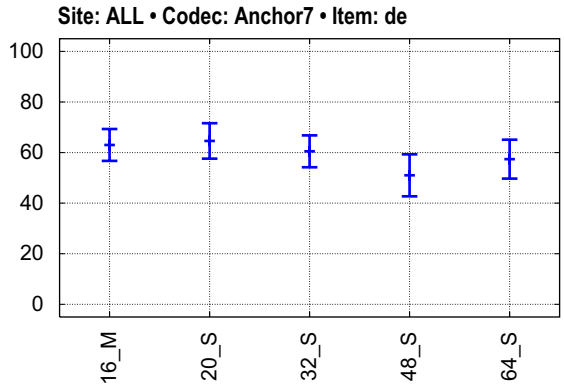


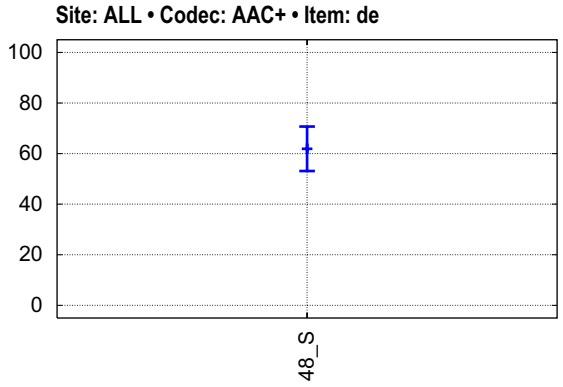


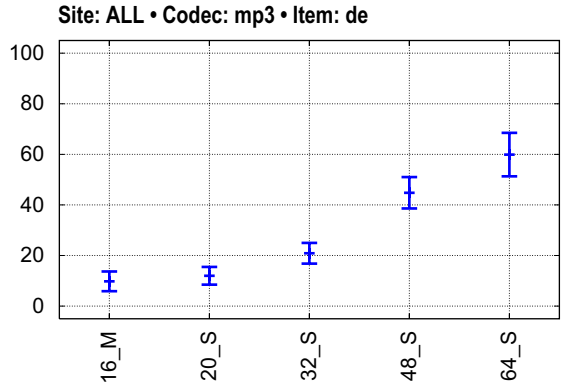


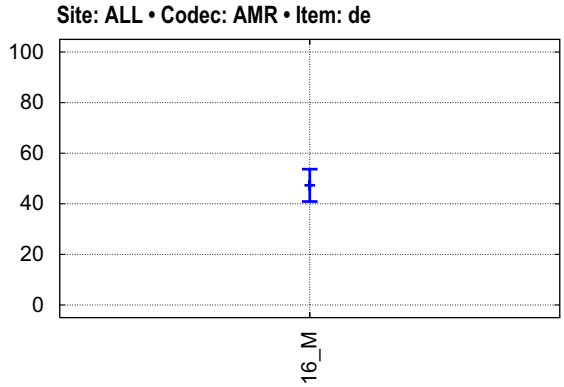
Appendix A — Codec / Item

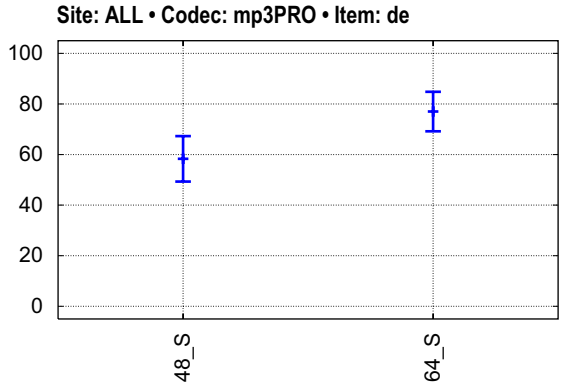


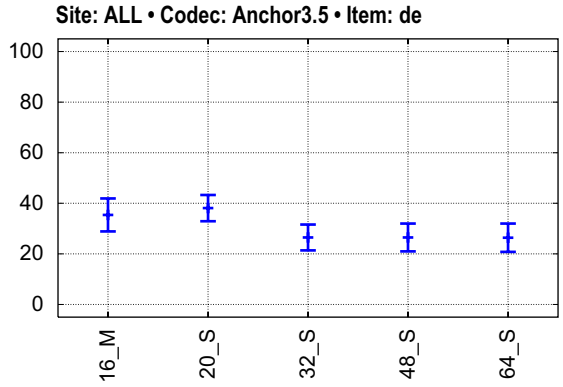


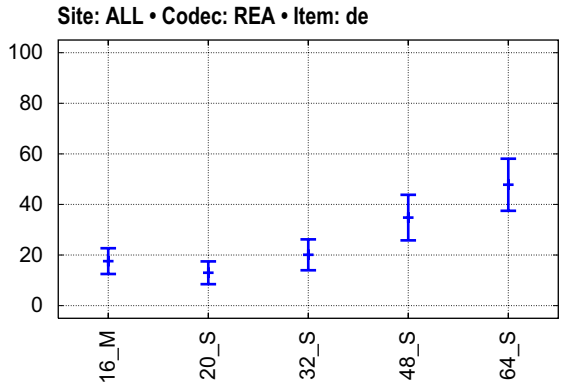






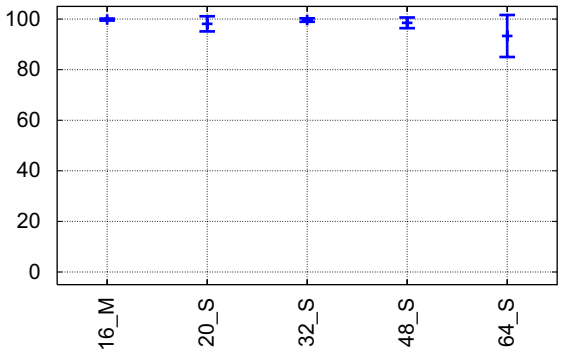




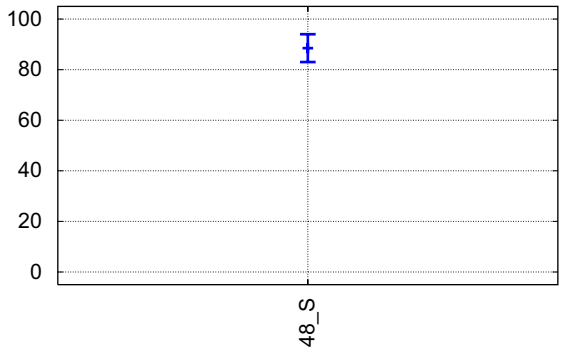


Appendix A — Codec / Item

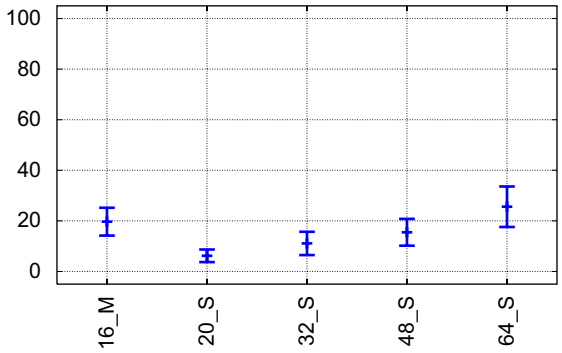
Site: ALL • Codec: Reference • Item: de



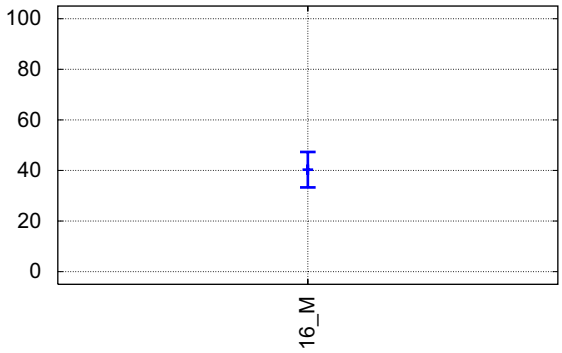
Site: ALL • Codec: AAC+ • Item: eu



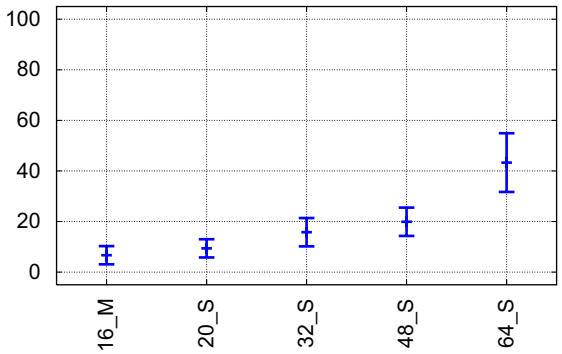
Site: ALL • Codec: RL_ • Item: de



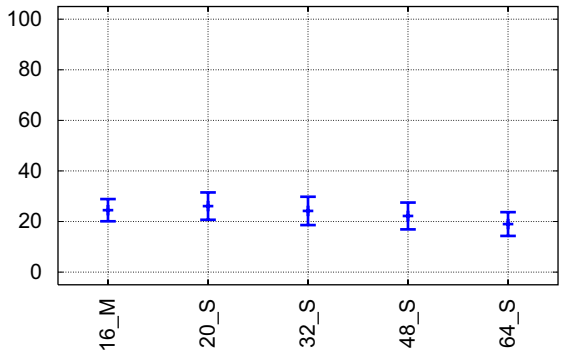
Site: ALL • Codec: AMR • Item: eu



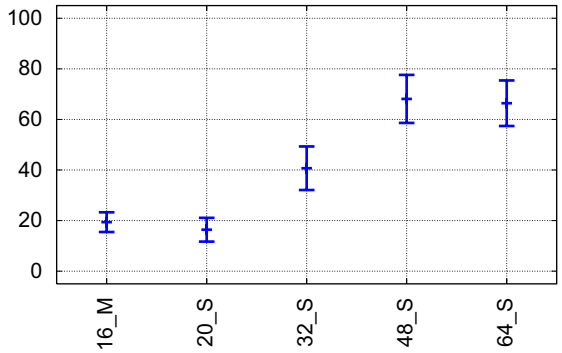
Site: ALL • Codec: WMA • Item: de



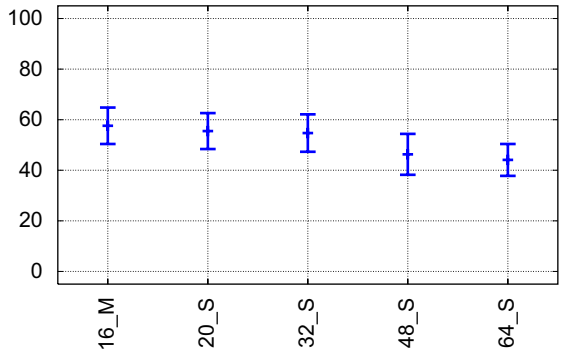
Site: ALL • Codec: Anchor3.5 • Item: eu



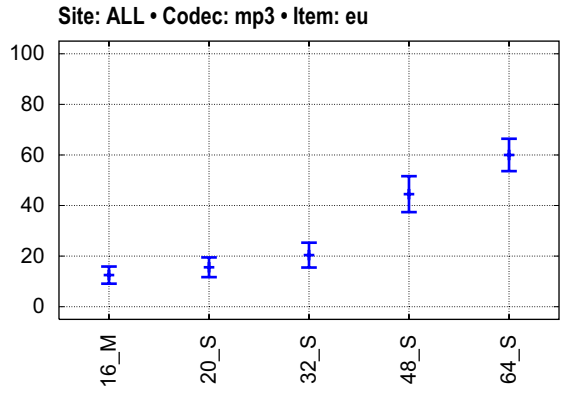
Site: ALL • Codec: AAC • Item: eu

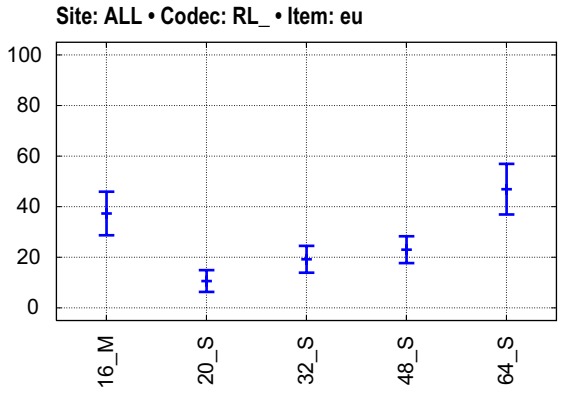


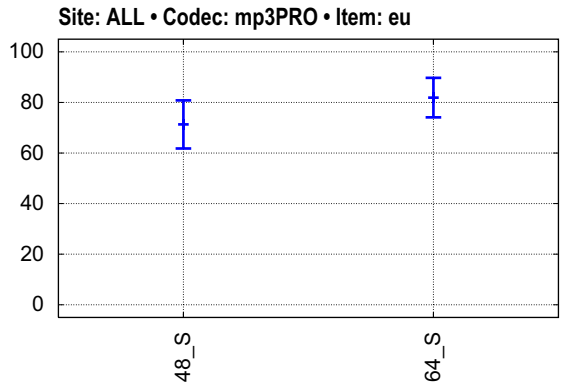
Site: ALL • Codec: Anchor7 • Item: eu

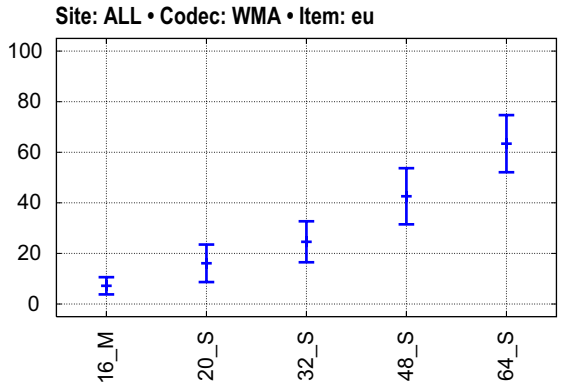


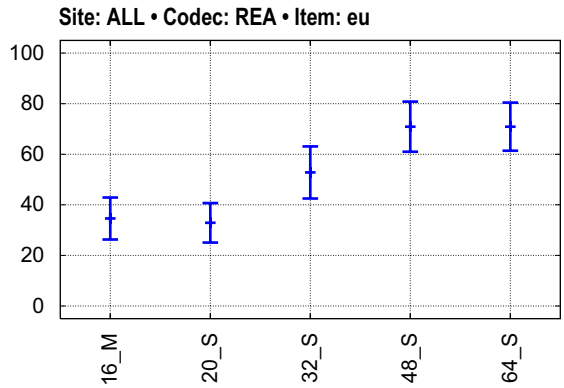
Appendix A — Codec / Item

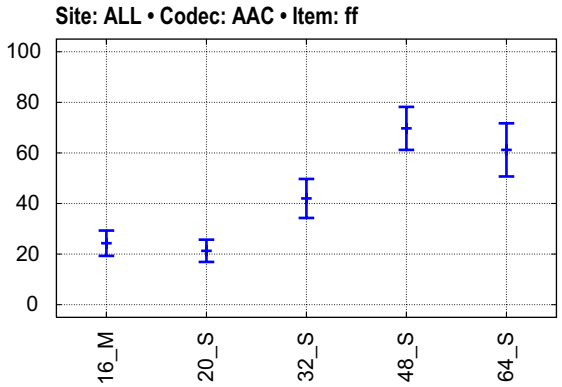


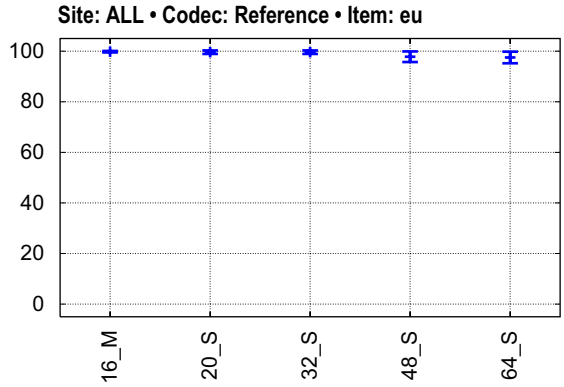


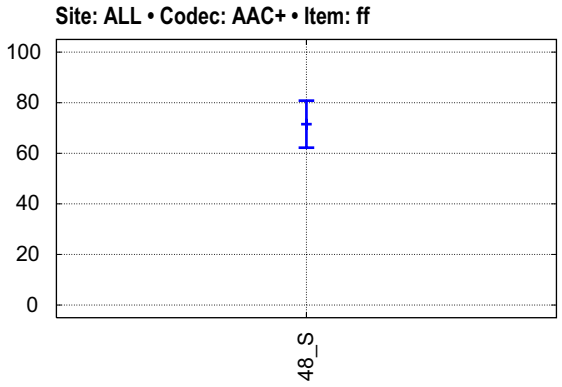




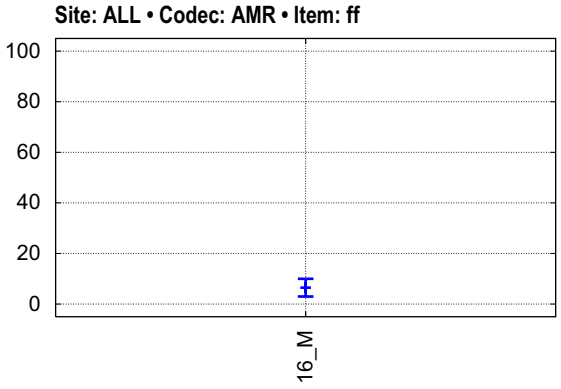


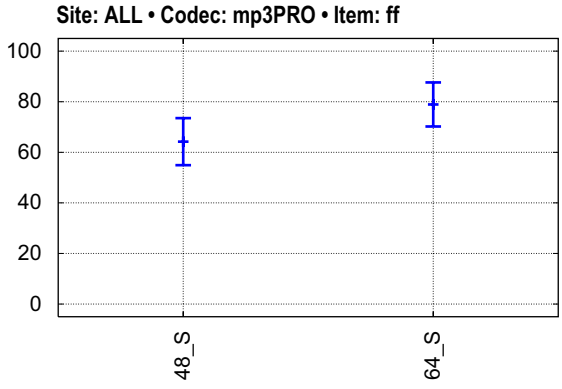


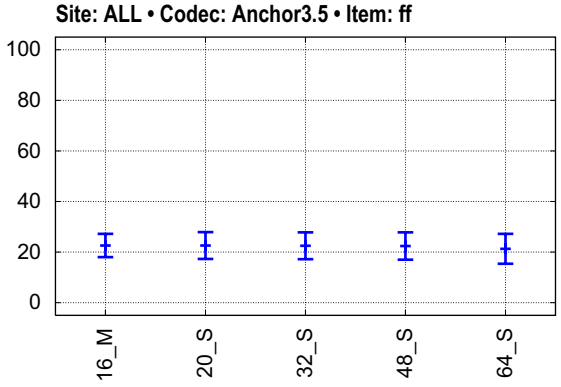


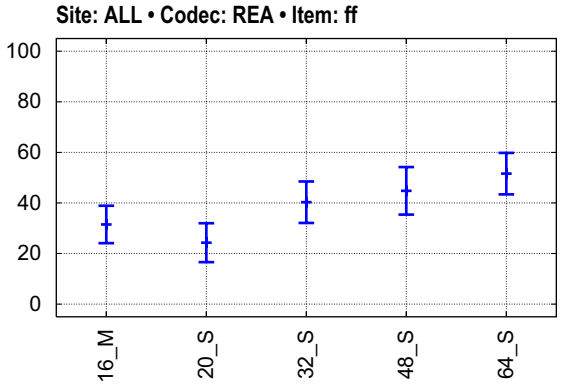


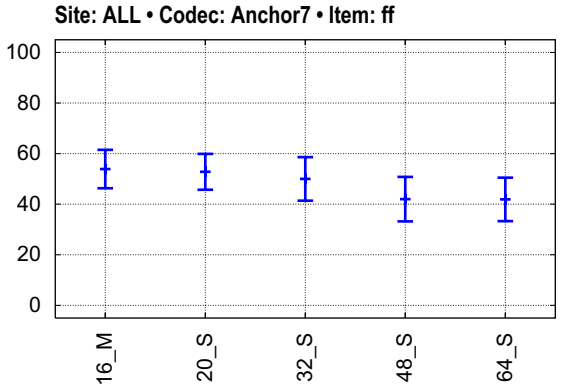
Appendix A — Codec / Item

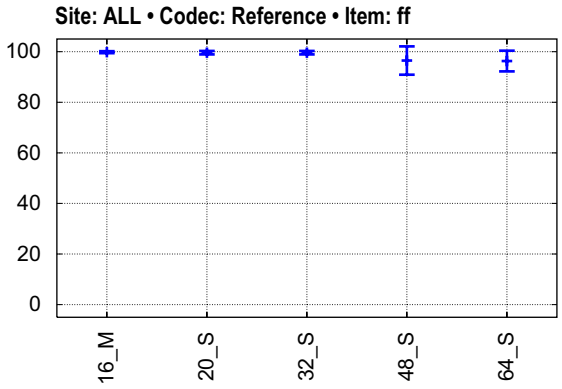


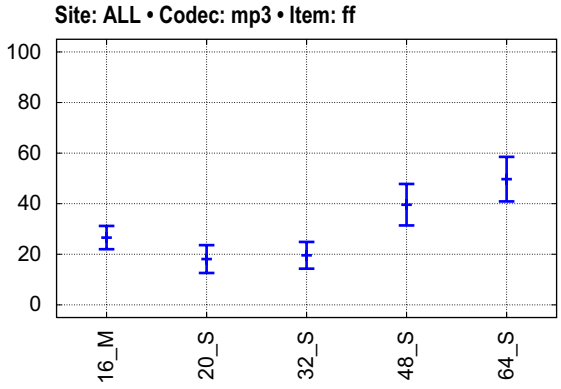


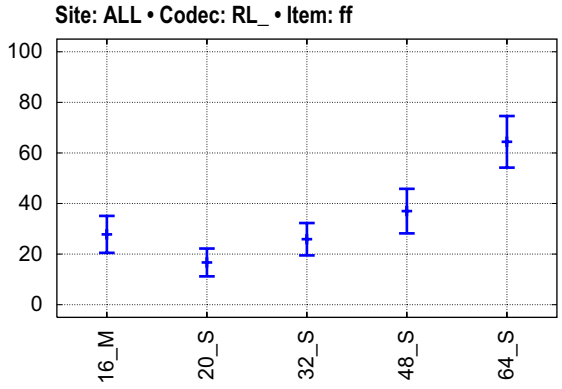






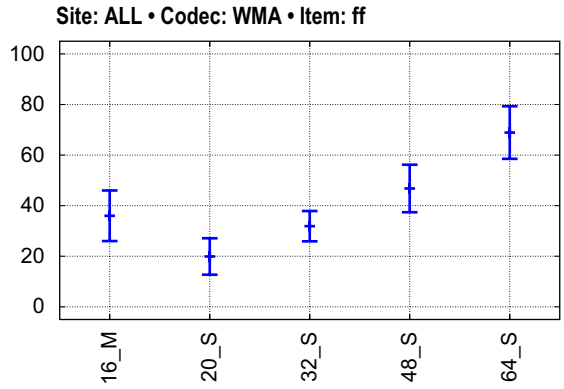


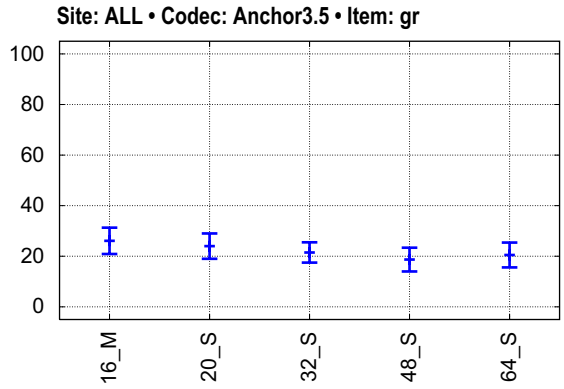


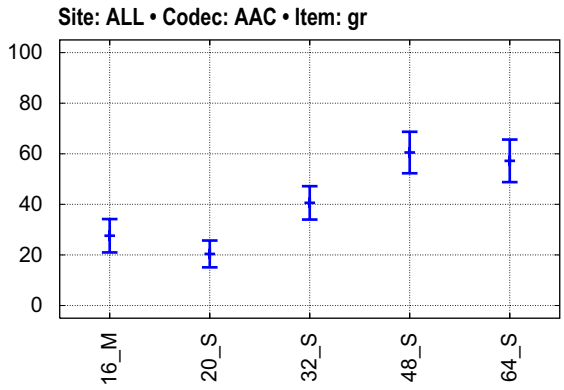


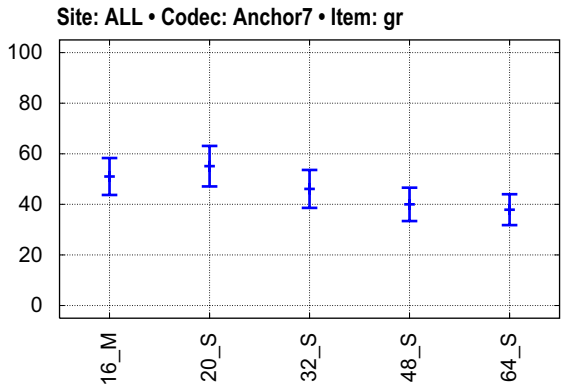
Ex. 1055-032

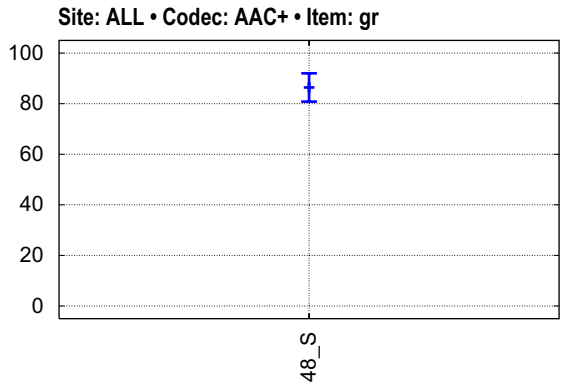
Appendix A — Codec / Item

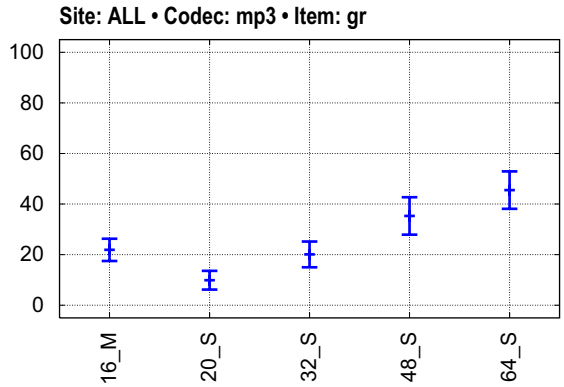


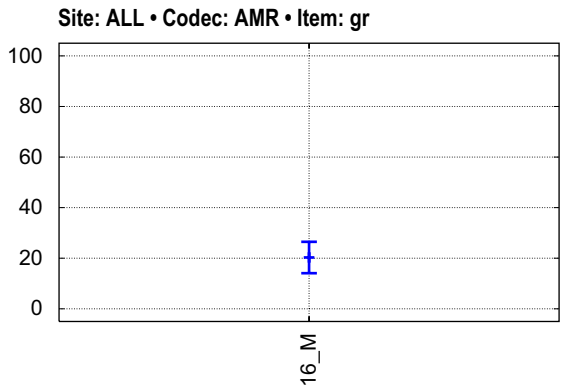


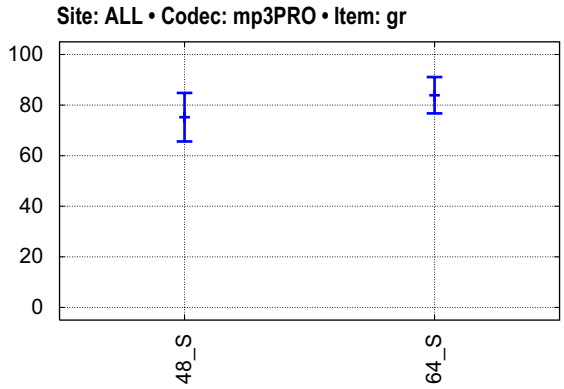




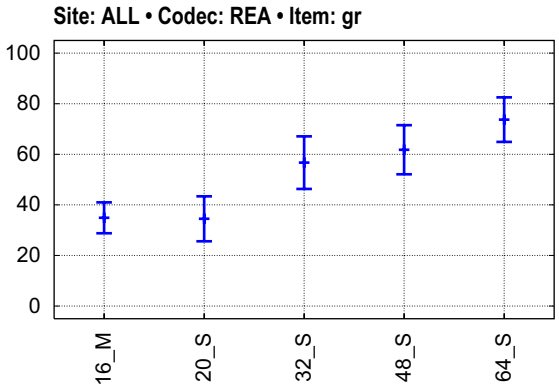


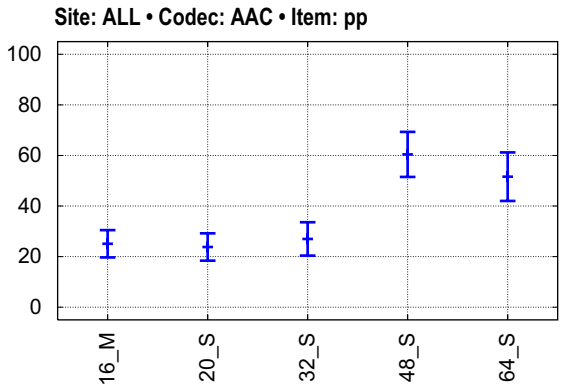


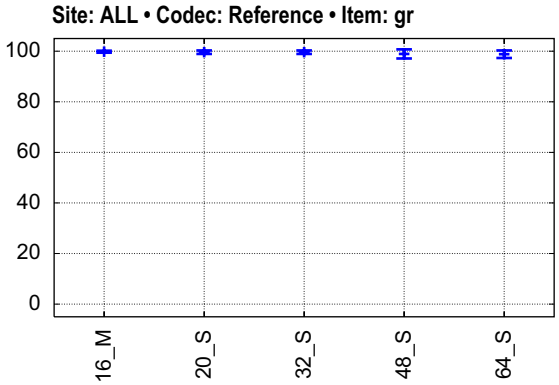


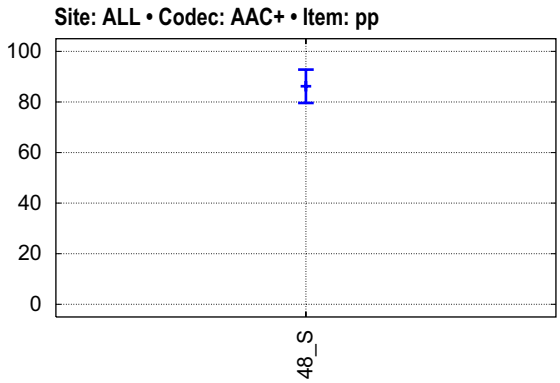


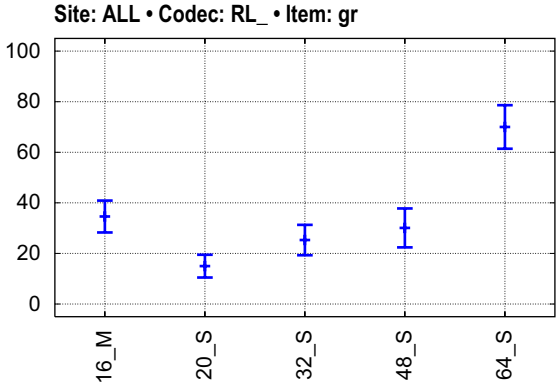
Appendix A — Codec / Item

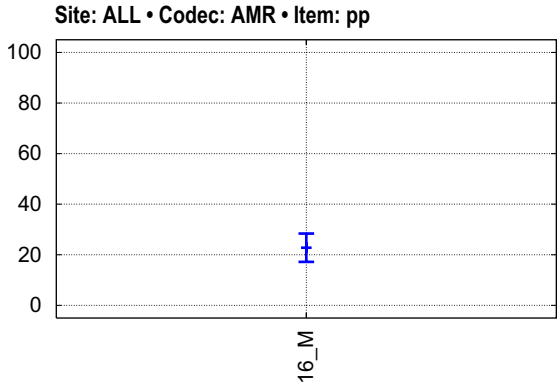


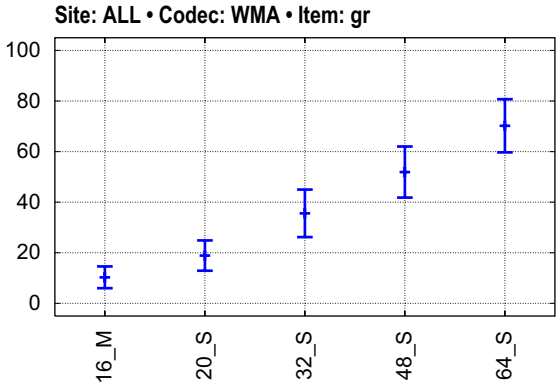


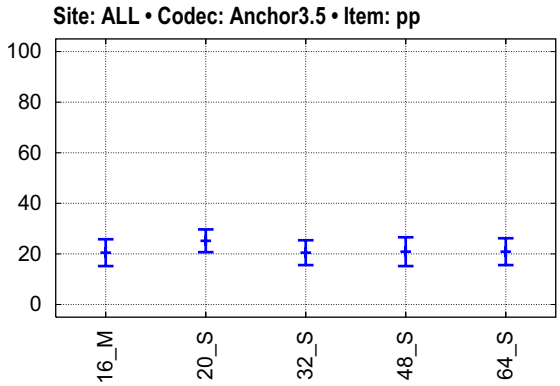








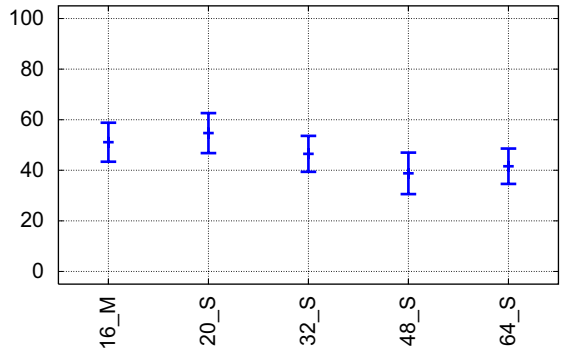




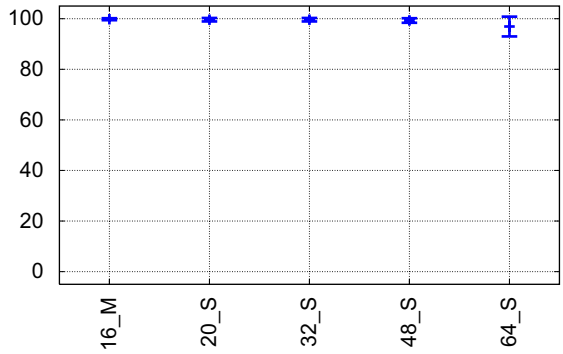
Ex. 1055-034

Appendix A — Codec / Item

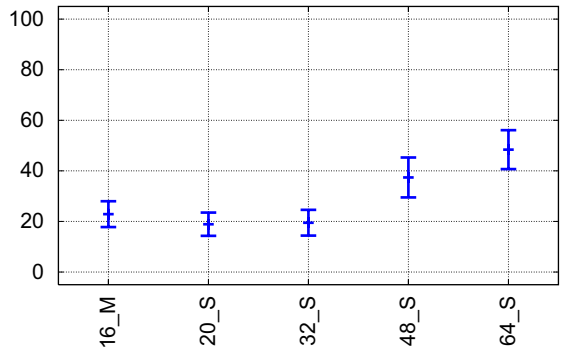
Site: ALL • Codec: Anchor7 • Item: pp



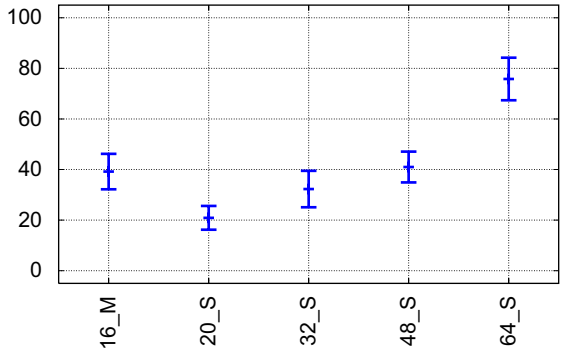
Site: ALL • Codec: Reference • Item: pp



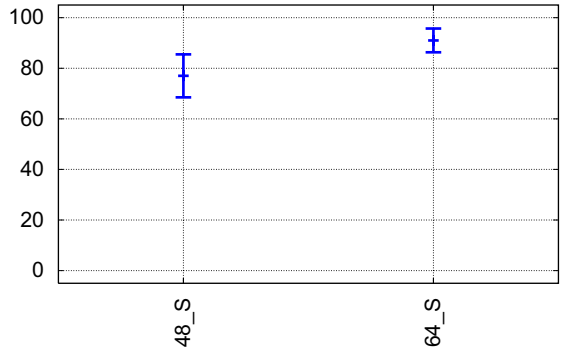
Site: ALL • Codec: mp3 • Item: pp



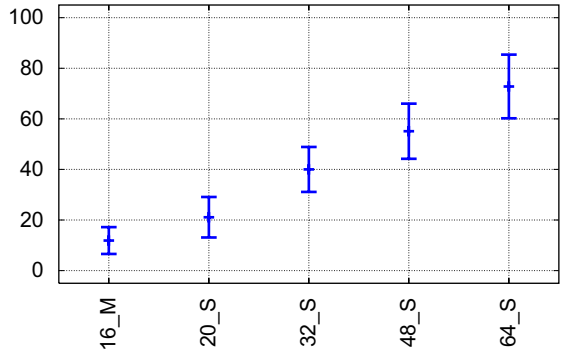
Site: ALL • Codec: RL_ • Item: pp



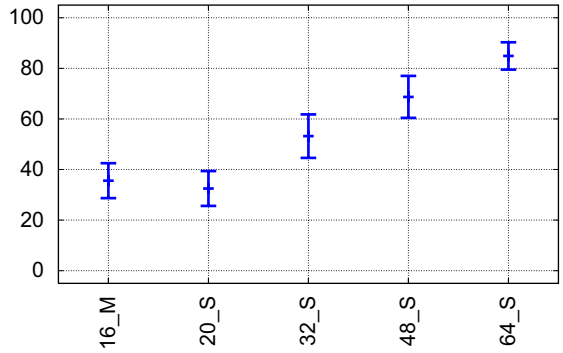
Site: ALL • Codec: mp3PRO • Item: pp



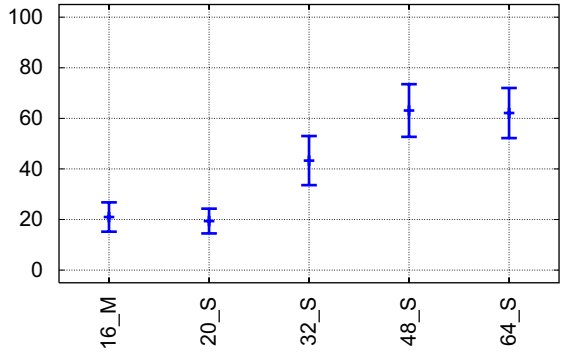
Site: ALL • Codec: WMA • Item: pp



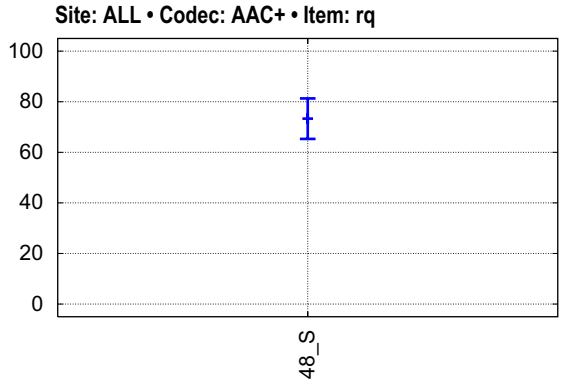
Site: ALL • Codec: REA • Item: pp

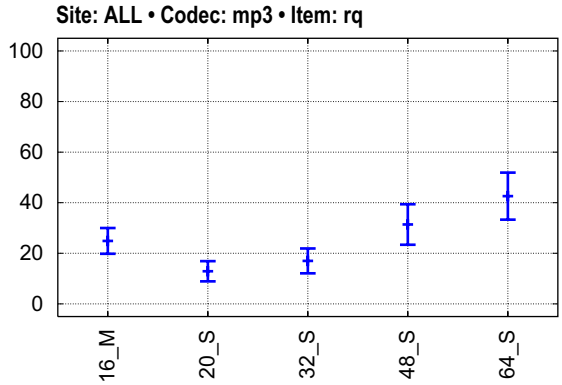


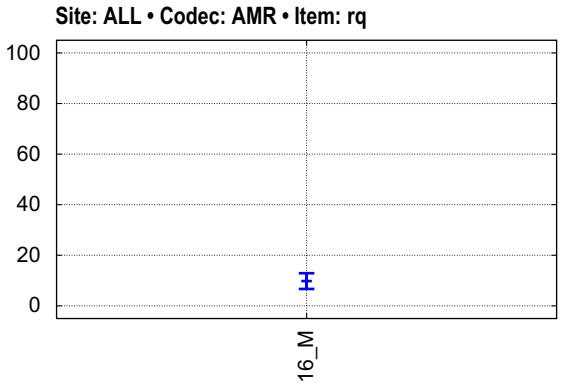
Site: ALL • Codec: AAC • Item: rq

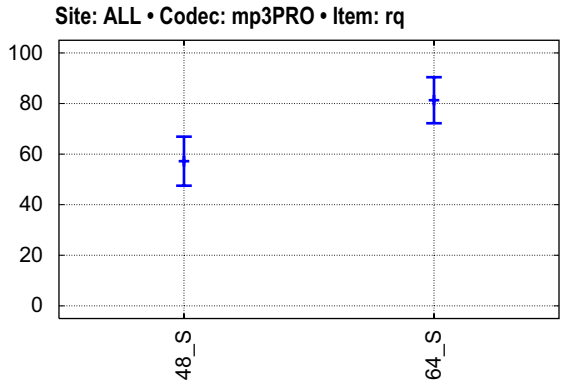


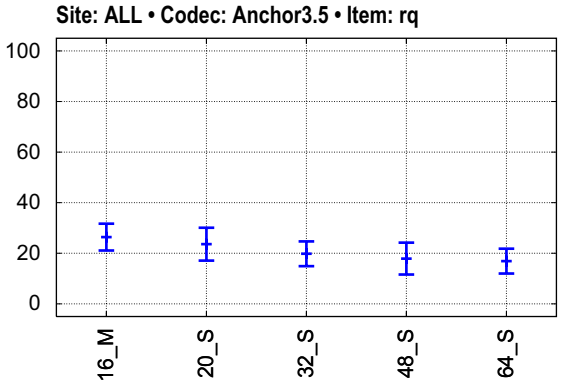
Appendix A — Codec / Item

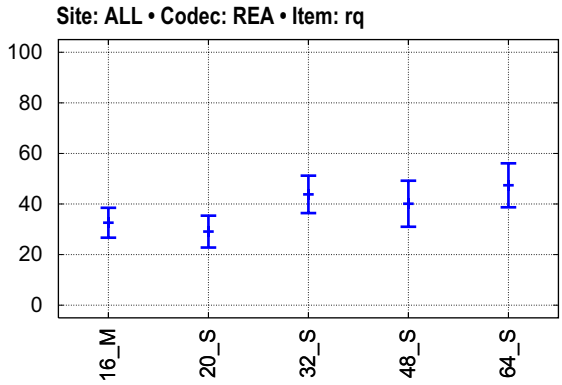


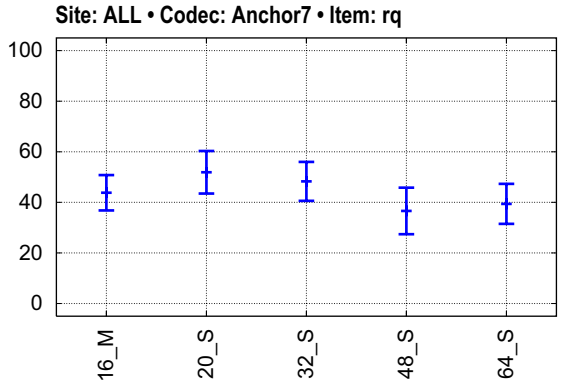


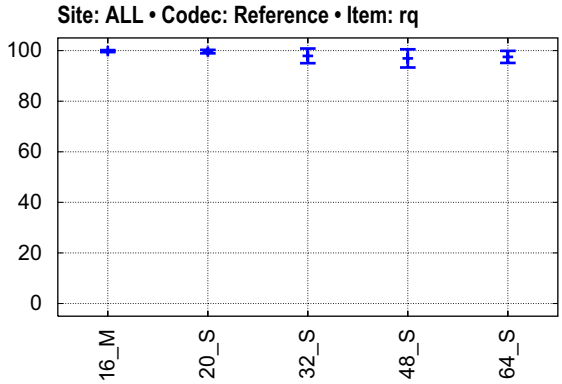






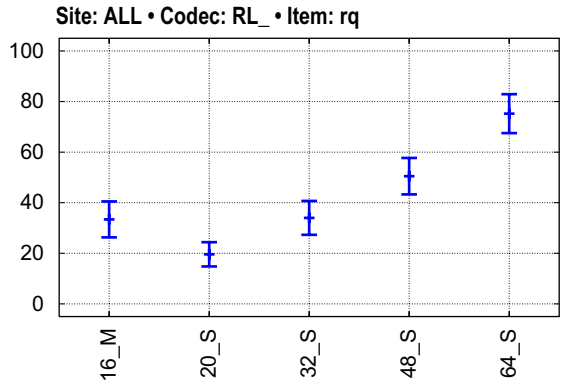


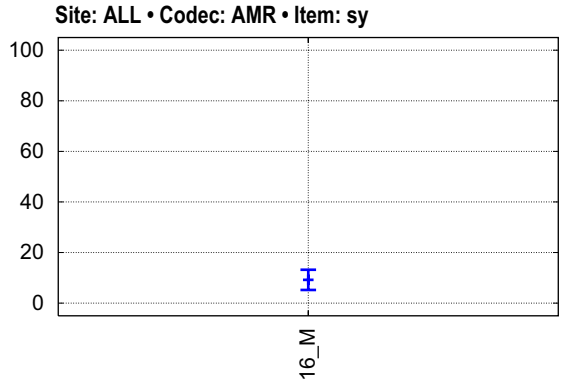


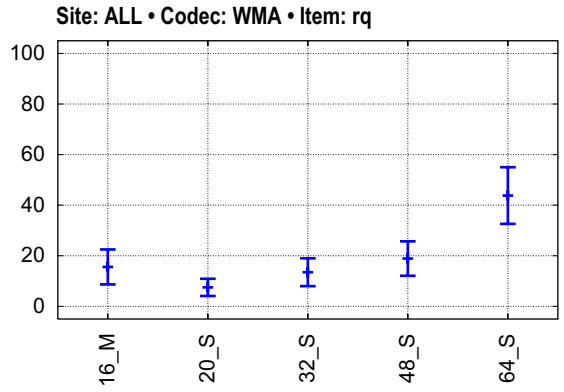


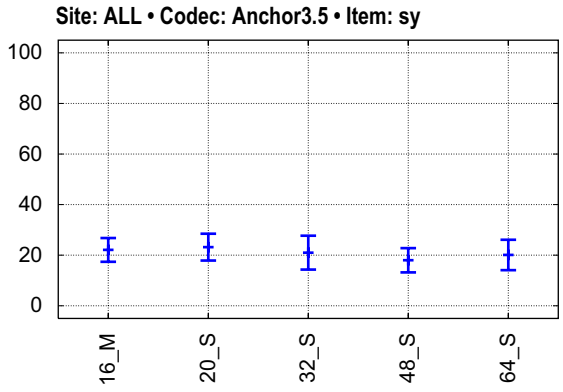
Ex. 1055-036

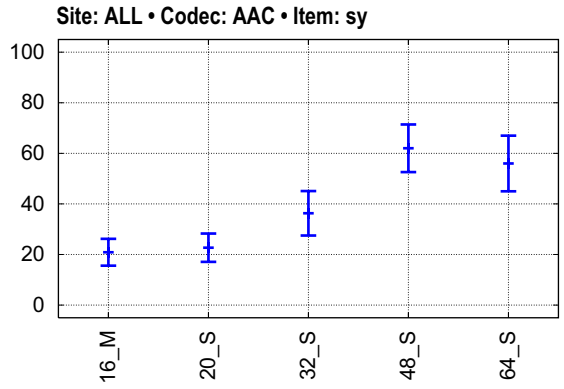
Appendix A — Codec / Item

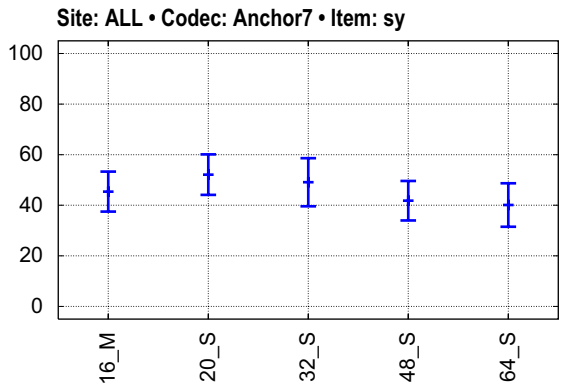


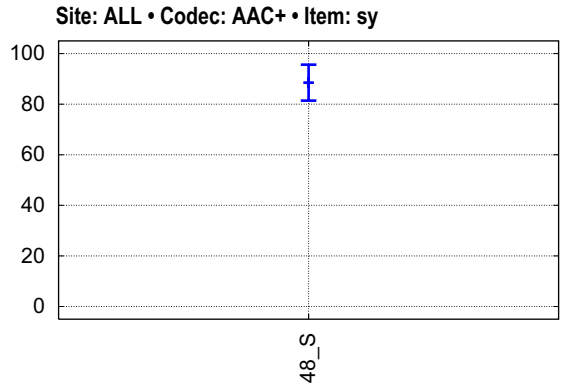


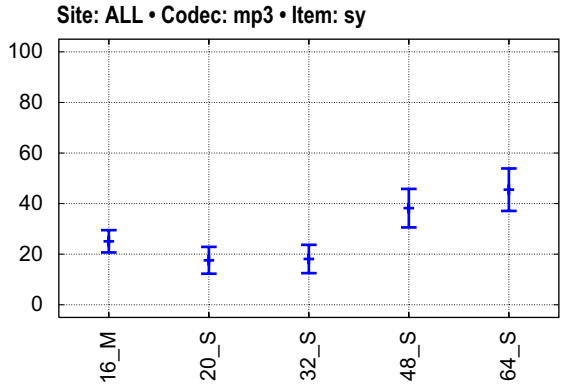






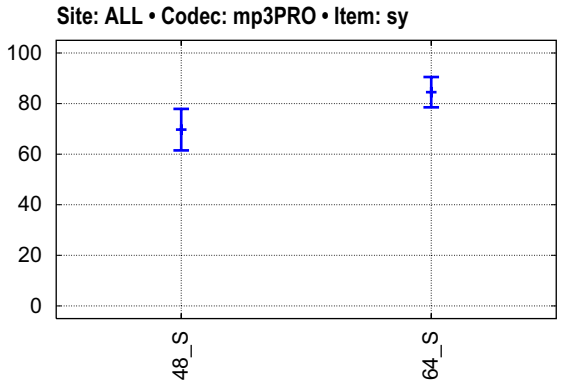


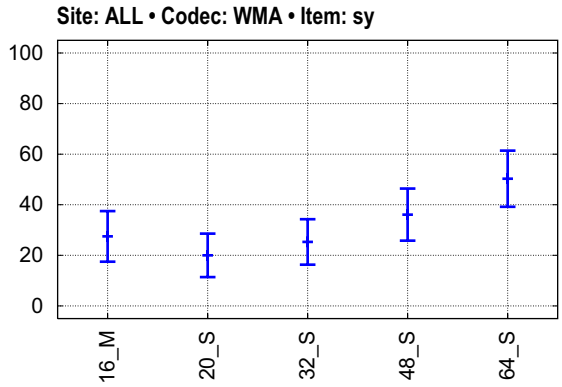


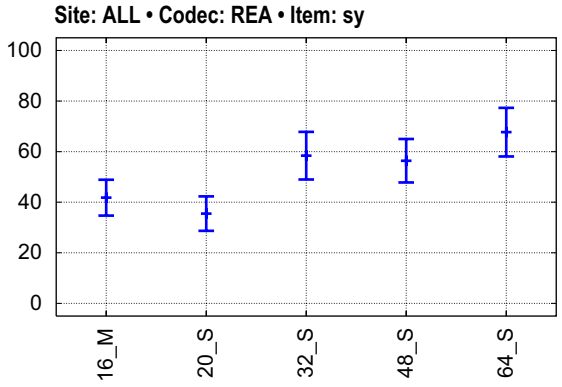


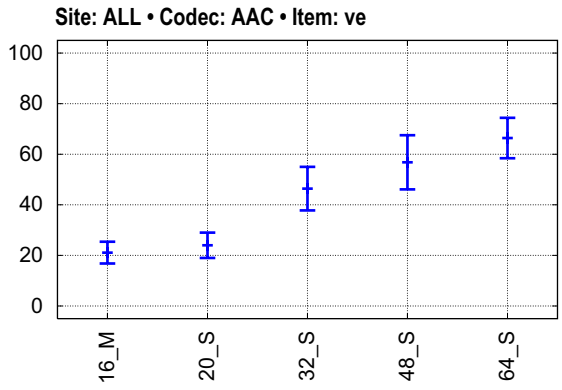
Ex. 1055-037

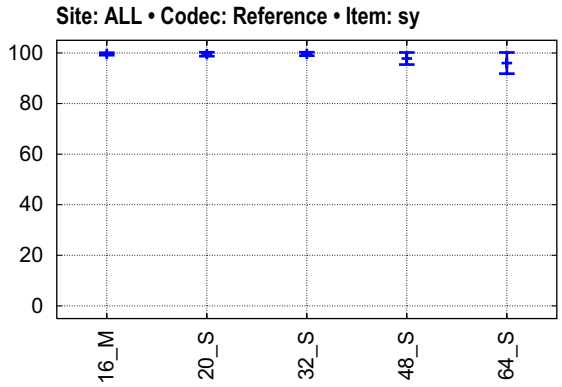
Appendix A — Codec / Item

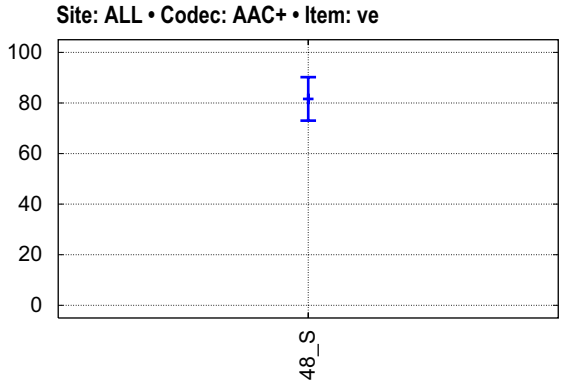


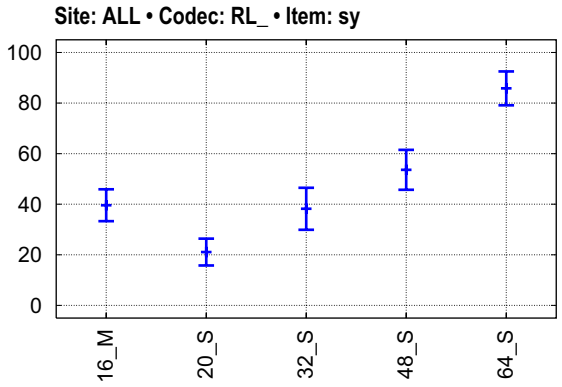


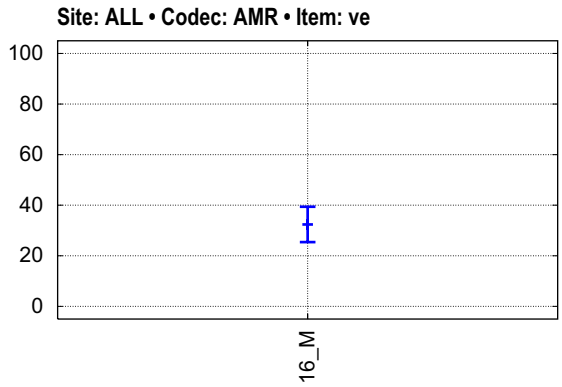








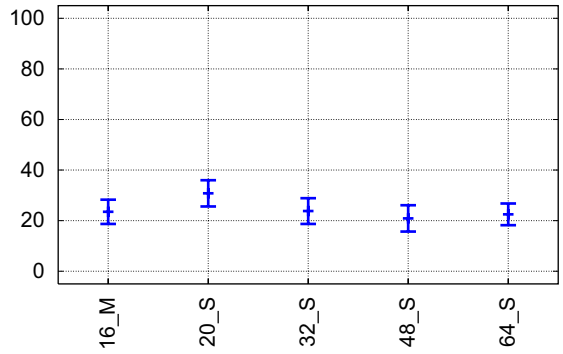




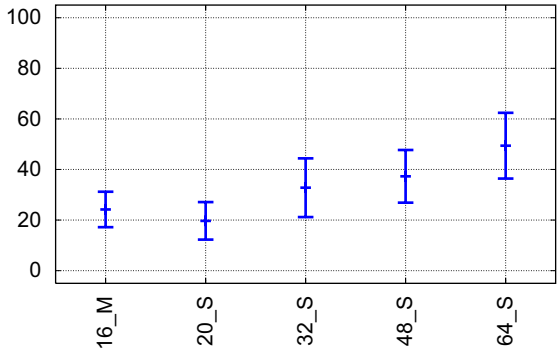
Ex. 1055-038

Appendix A — Codec / Item

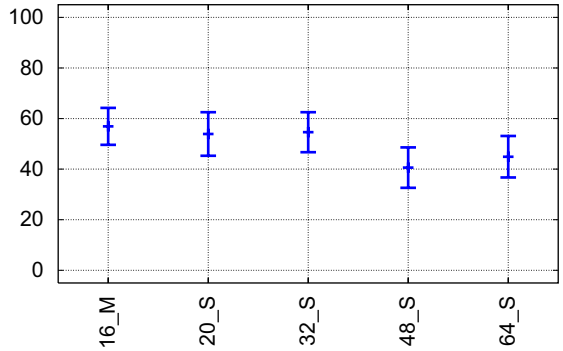
Site: ALL • Codec: Anchor3.5 • Item: ve



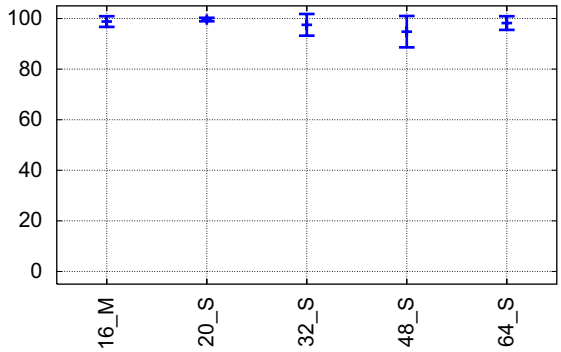
Site: ALL • Codec: REA • Item: ve



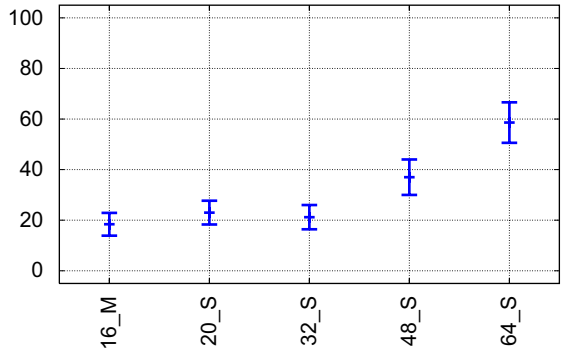
Site: ALL • Codec: Anchor7 • Item: ve



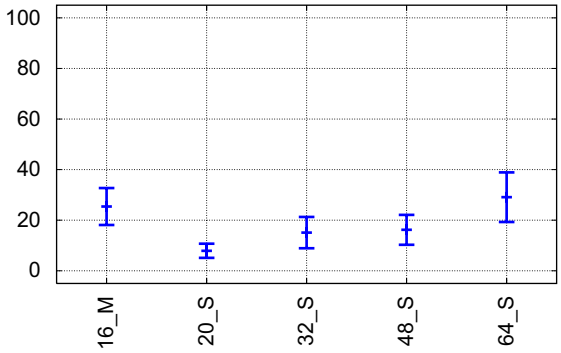
Site: ALL • Codec: Reference • Item: ve



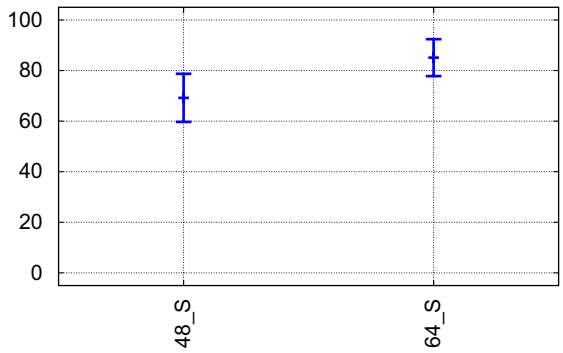
Site: ALL • Codec: mp3 • Item: ve



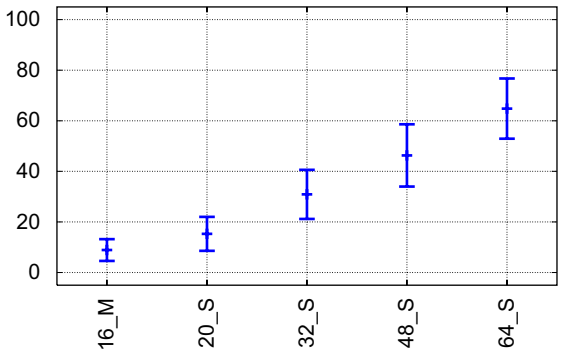
Site: ALL • Codec: RL_ • Item: ve



Site: ALL • Codec: mp3PRO • Item: ve

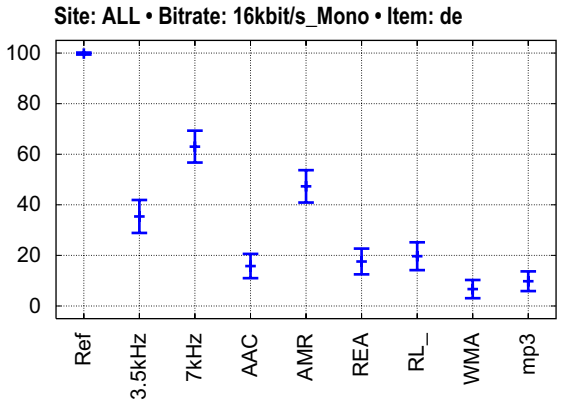


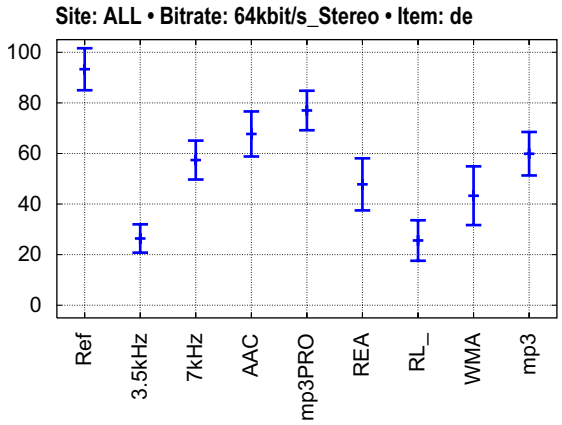
Site: ALL • Codec: WMA • Item: ve

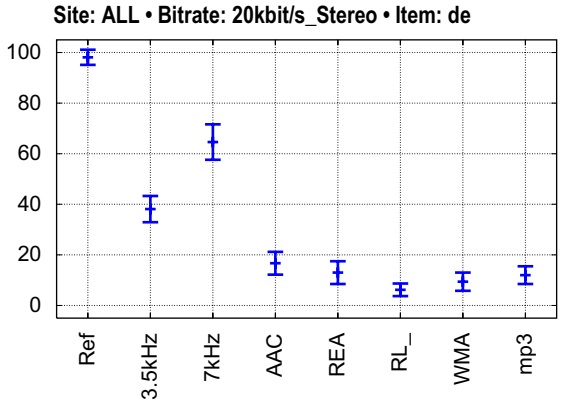


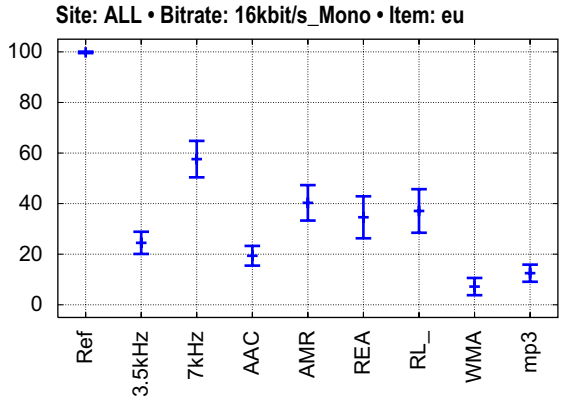
Ex. 1055-039

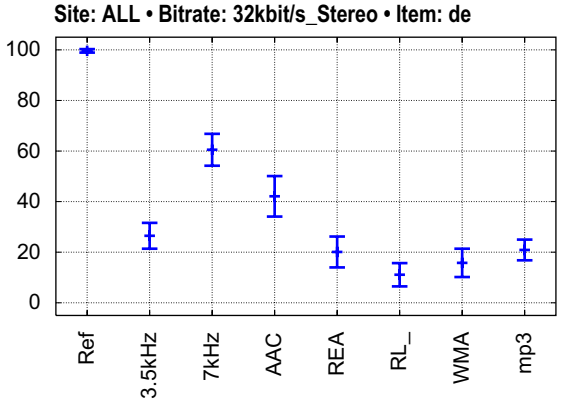
Appendix A — Item / Bitrate

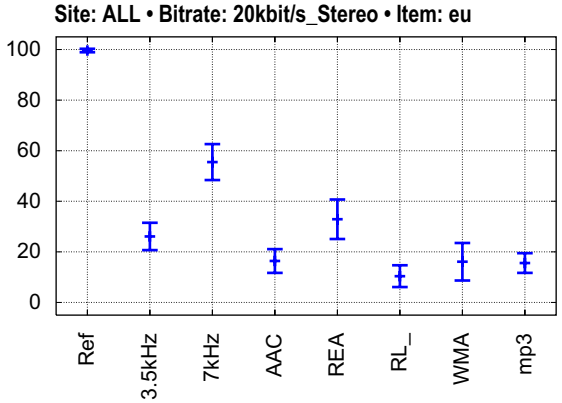


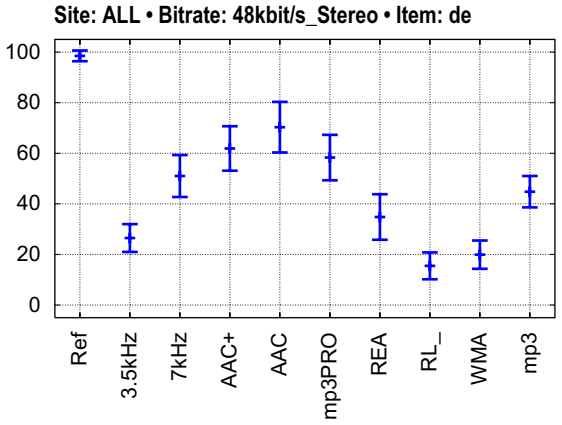


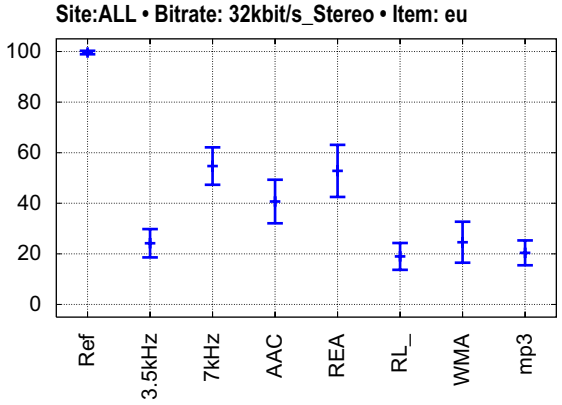






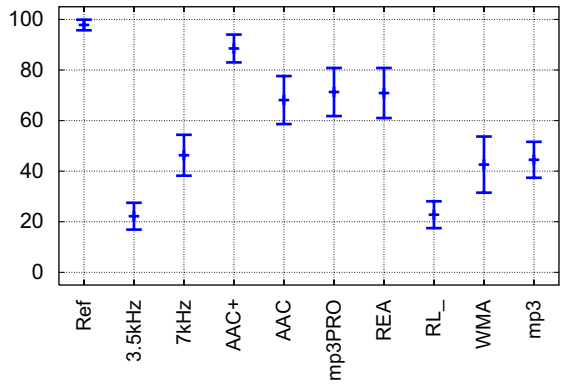




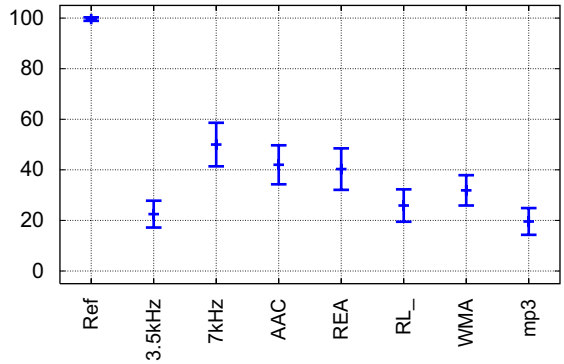


Appendix A — Item / Bitrate

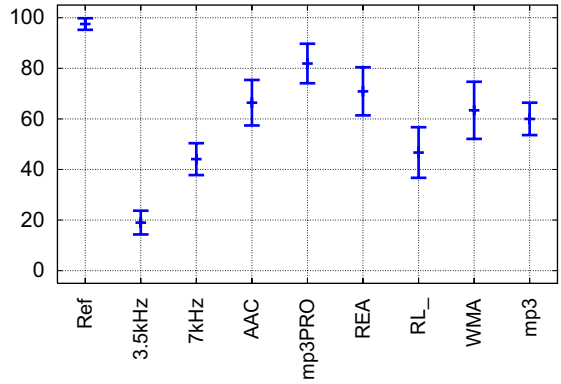
Site: ALL • Bitrate: 48kbit/s_Stereo • Item: eu



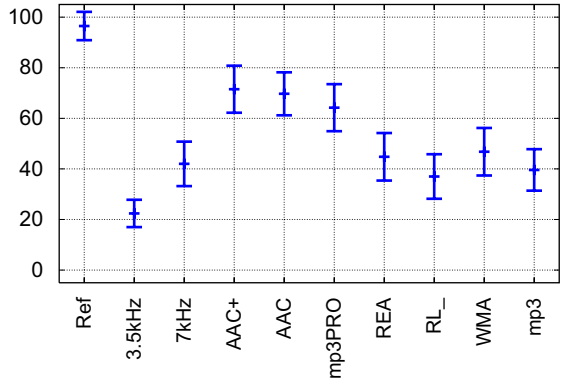
Site: ALL • Bitrate: 32kbit/s_Stereo • Item: ff



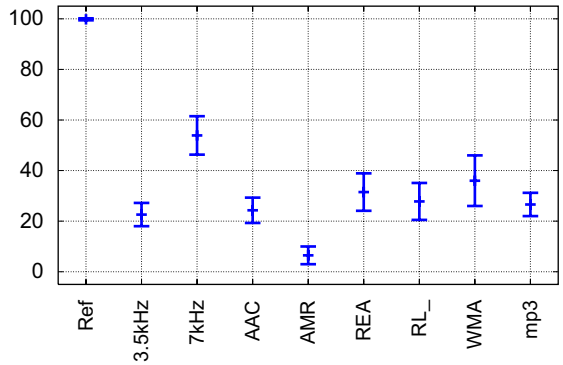
Site: ALL • Bitrate: 64kbit/s_Stereo • Item: eu



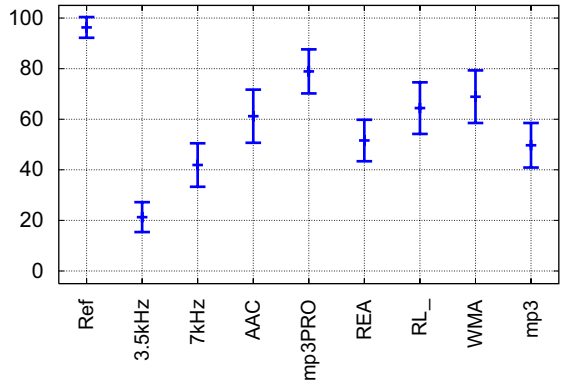
Site: ALL • Bitrate: 48kbit/s_Stereo • Item: ff



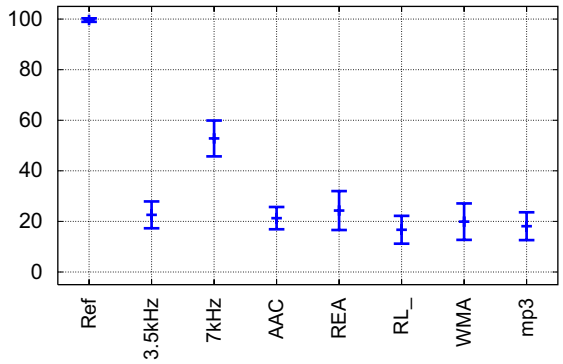
Site: ALL • Bitrate: 16kbit/s_Mono • Item: ff



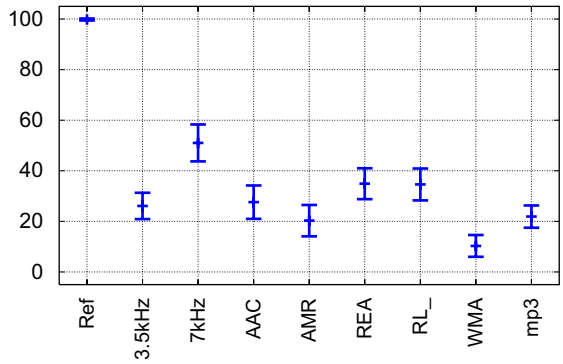
Site: ALL • Bitrate: 64kbit/s_Stereo • Item: ff



Site: ALL • Bitrate: 20kbit/s_Stereo • Item: ff



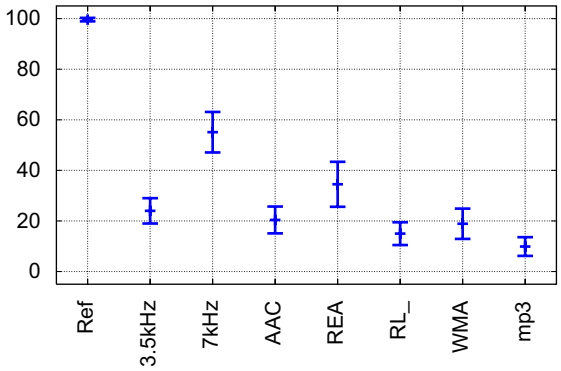
Site: ALL • Bitrate: 16kbit/s_Mono • Item: gr



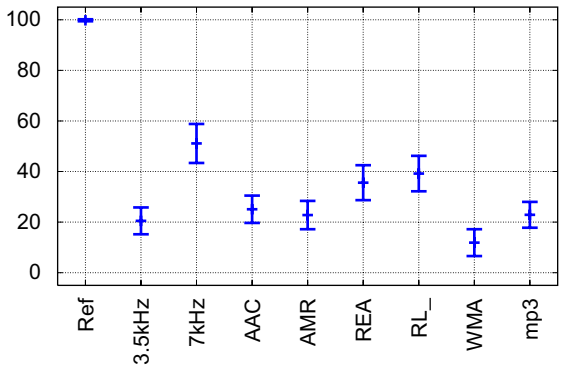
Ex. 1055-041

Appendix A — Item / Bitrate

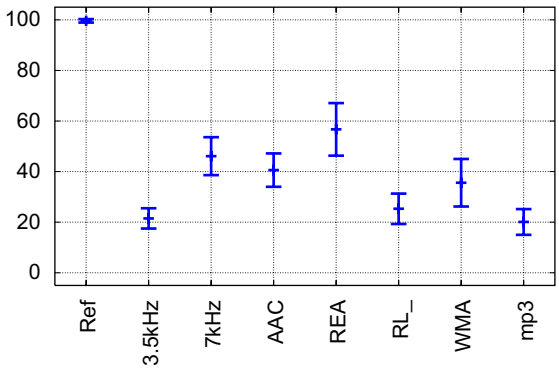
Site: ALL • Bitrate: 20kbit/s_Stereo • Item: gr



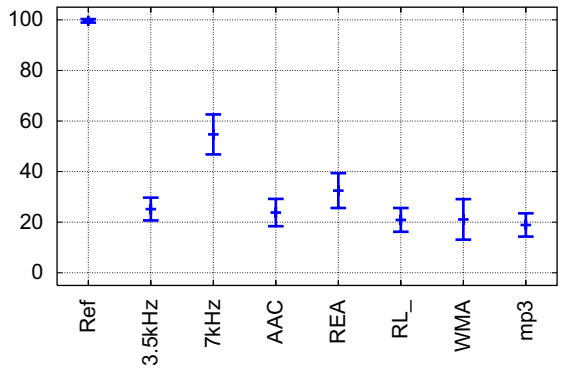
Site: ALL • Bitrate: 16kbit/s_Mono • Item: pp



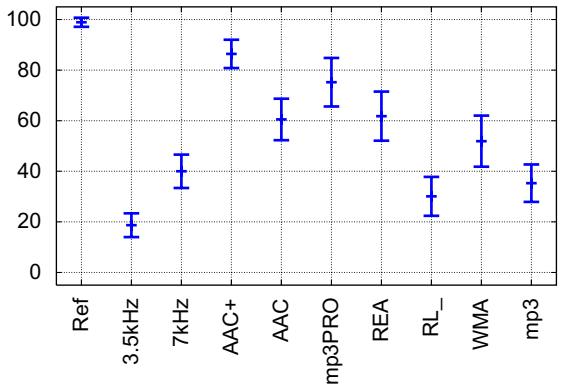
Site: ALL • Bitrate: 32kbit/s_Stereo • Item: gr



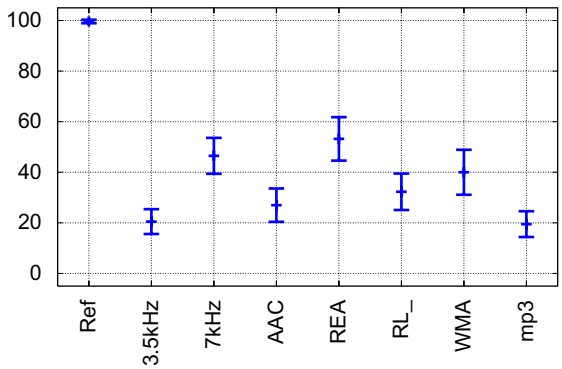
Site: ALL • Bitrate: 20kbit/s_Stereo • Item: pp



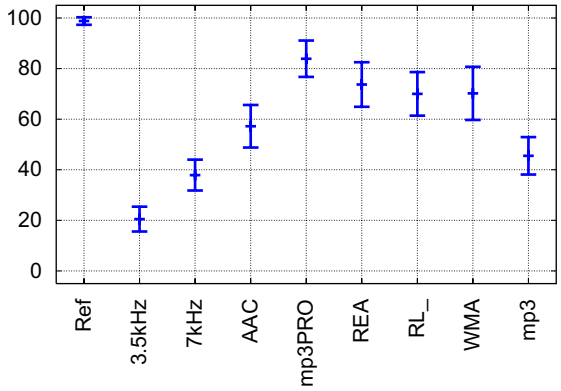
Site: ALL • Bitrate: 48kbit/s_Stereo • Item: gr



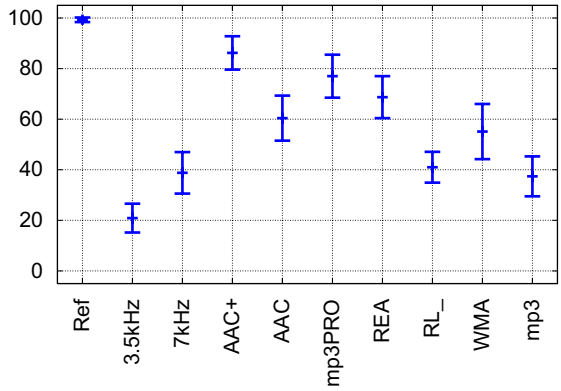
Site: ALL • Bitrate: 32kbit/s_Stereo • Item: pp



Site: ALL • Bitrate: 64kbit/s_Stereo • Item: gr



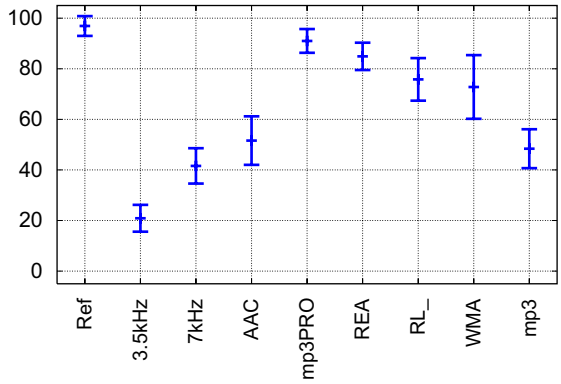
Site: ALL • Bitrate: 48kbit/s_Stereo • Item: pp



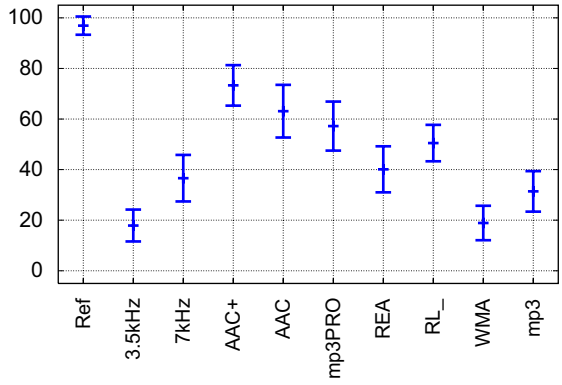
Ex. 1055-042

Appendix A — Item / Bitrate

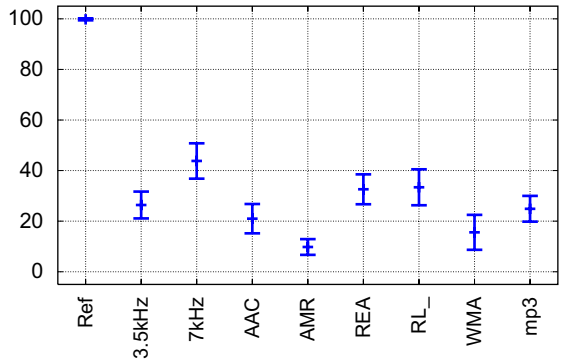
Site: ALL • Bitrate: 64kbit/s_Stereo • Item: pp



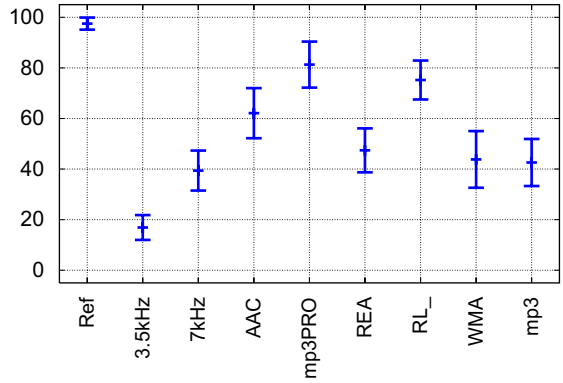
Site: ALL • Bitrate: 48kbit/s_Stereo • Item: rq



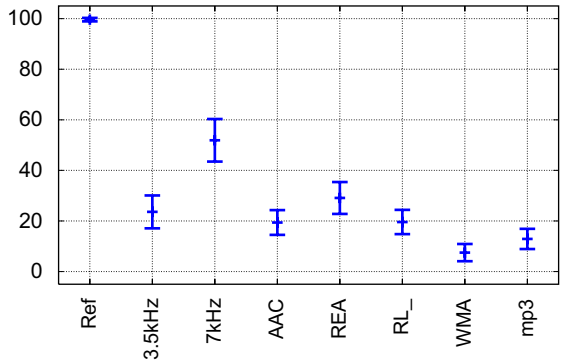
Site: ALL • Bitrate: 16kbit/s_Mono • Item: rq



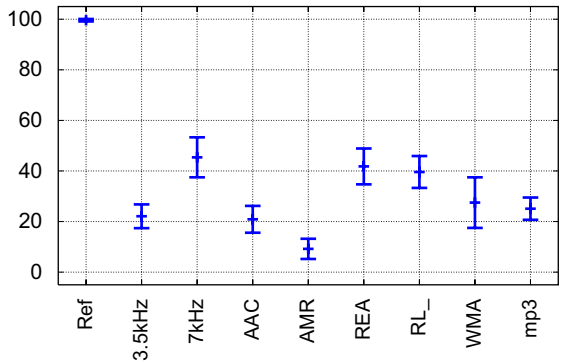
Site: ALL • Bitrate: 64kbit/s_Stereo • Item: rq



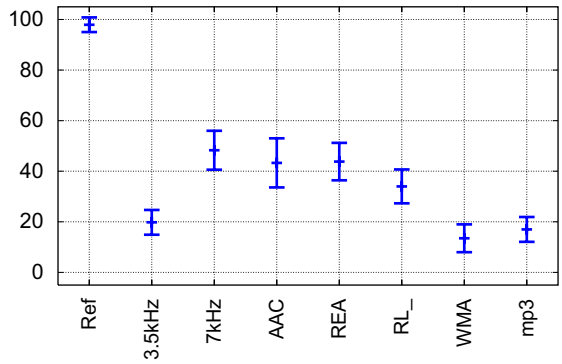
Site: ALL • Bitrate: 20kbit/s_Stereo • Item: rq



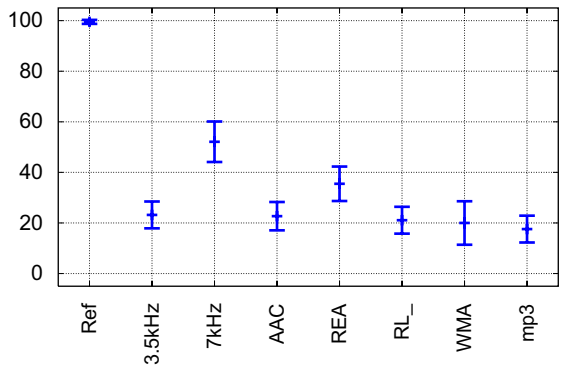
Site: ALL • Bitrate: 16kbit/s_Mono • Item: sy



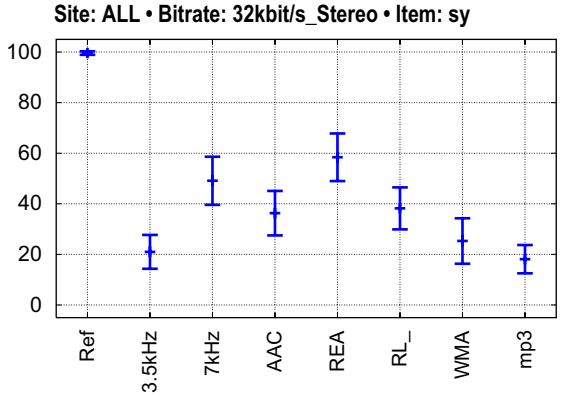
Site: ALL • Bitrate: 32kbit/s_Stereo • Item: rq

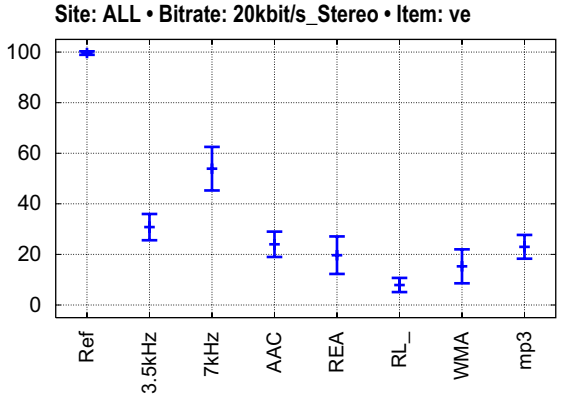


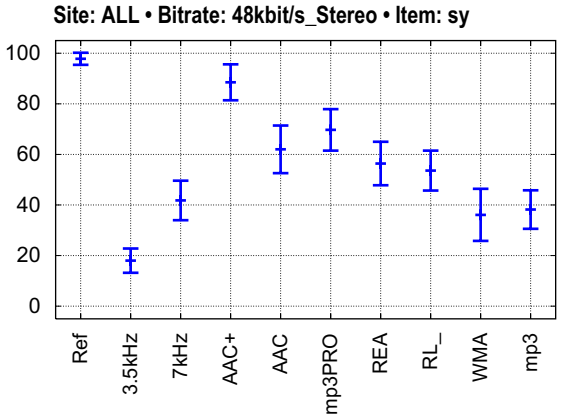
Site: ALL • Bitrate: 20kbit/s_Stereo • Item: sy

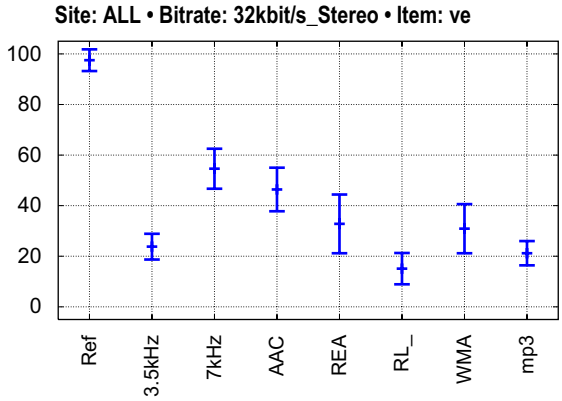


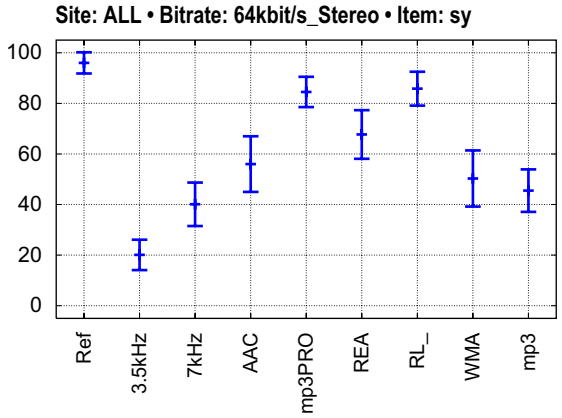
Appendix A — Item / Bitrate

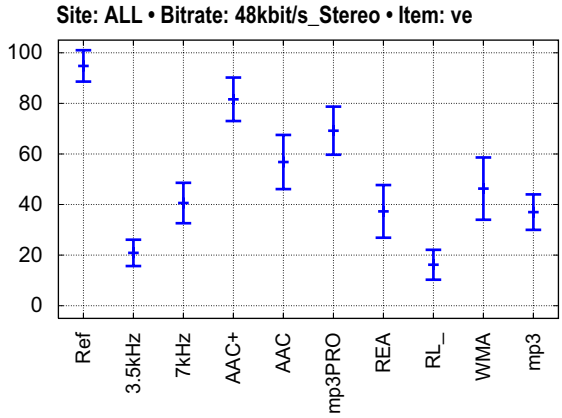


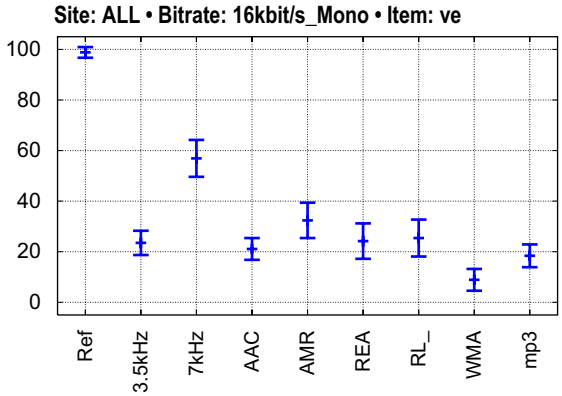


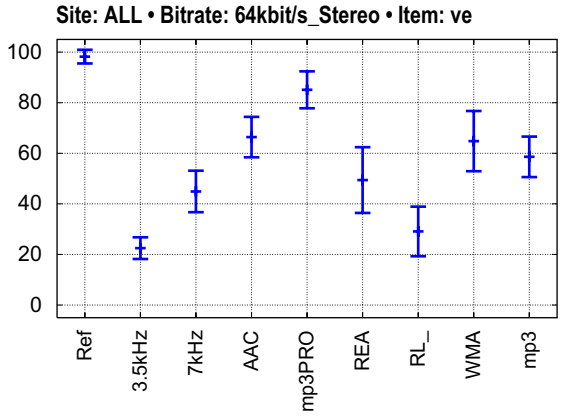












EBU•UER

Last revision: 03 July 2003