

TECHNICAL REPORT

ISSUE 2

CLIENT: Sony Europe BV

This report details listening tests and technical measurements carried out on three digital audio compression / decompression (CODEC) systems.

The tests were specifically designed to compare a new Sony CODEC system with two other established CODEC systems.

The purpose of the tests was to verify the findings of the Sony Corporation

www.its-rtc.co.uk
www.itsglobal.com

AUTHORISED
FOR ISSUE:

Ken Biswell
Business Manager

DATE: 28-Feb-03

REPORT AUTHOR: Roy Brooker, Fraser Henderson

This report shall not be reproduced except in full without the written approval of ITS Research & Testing Centre.

A.61346 Issue 2

Audio Compression

Contents

| | |
|--------------------------------------|---|
| Introduction | 3 |
| 1. Listening Tests | 4 |
| 1.1 Test method | 4 |
| 1.2 Equipment used | 5 |
| 1.3 Listening test material creation | 6 |
| 1.4 Listening test procedure | 6 |
| 1.5 Statistical analysis of results | 6 |
| 2. Results | 7 |
| 3. Technical Tests | 9 |
| 3.1 Signal analysis | 9 |
| 3.2 Frequency response plots | 9 |
| Appendix I | : Raw results from listening tests |
| Appendix II | : Signal analysis |
| Appendix III | : Frequency response plots |

Introduction

ITS-RTC were requested by Sony Europe BV to conduct some listening tests to compare a new Sony digital audio CODEC system (ATRAC3plus) with two established CODEC systems: MP3 and WMA (version9). The tests were to mirror tests already carried out by Sony in order to provide independent conformation of their findings. The tests were designed specifically to compare the Sony system against the other two systems at different bit-rates. 132Kbps ATRAC3 was also compared to the two other CODEC systems at a high data rate only.

The tests were in two parts: Subjective listening and technical measurement. However, emphasis was put on the subjective tests.



Above: Listening test conducted at ITS-RTC facilities. Image available on file.

1. Listening tests

1.1 Test Method

Sony identified two generic methods of listening. Using expert listeners to give an absolute rating of quality or paired comparisons. The paired comparison methods had been chosen by Sony as the preferred method for this comparison test and after due consideration, the audio engineers at ITS-RTC agreed that this was the best approach for this particular test.

Ten listeners (two sessions of 5) were used. These comprised a mix of ITS-RTC staff with basic listening experience and some external 'experienced' listeners. The listeners were also of mixed gender and ages.

Ten program sequences were selected. ITS auditioned the material that Sony had used in their tests and also five additional tracks of our own choice. After due consideration we felt that the Sony tracks, which comprised a mix of complex music, speech and solo instruments, to be most appropriate for this test, however it was agreed that we substituted one of our tracks which the listeners were familiar with.

The paired comparison method used was the RAB method. A reference track (R) uncompressed WAV file is played followed by two test tracks to be compared (A and B). The listeners are asked to mark which they prefer A or B on a zero, one or two scale where zero means no preference, one means slight preference and two means significant preference. The order in which the comparisons were presented to the listeners and the order in which the tracks were presented to the listeners were randomised. Each comparison was listened to on two occasions as AB and then in the randomised order as BA.

Four tests were carried out, comparing the Sony products to two other products at four different data rates. Table 1 below details the four tests.

Table 1 The four subjective test sequences

| Session | Type | CODECs | Number of runs |
|---------|-----------------------------|---|----------------------------|
| 1 | 64Kbps comparison | ATRAC3plus versus MP3 and WMA | 40 (+5 dummy) of 10 tracks |
| 2 | 48Kbps comparison | ATRAC3plus versus MP3 and WMA | 40 (+5 dummy) of 10 tracks |
| 3 | ATRAC3 (132Kbps) comparison | ATRAC3 (132Kbps) versus MP3 (128Kbps) and WMA (128Kbps) | 40 (+5 dummy) of 10 tracks |
| 4 | Half bit rate comparison | ATRAC3plus (64Kbps) versus MP3 (128Kbps) | 20 (+5 dummy) of 10 tracks |

The 'program segment order' for each of the four test runs was randomly generated using a random pattern generator program, supplied by Sony Corporation. A different random 'seed' was used for each of the sections.

Each test run included 5 'dummy' runs at the beginning of the playback order that were not used in the final statistical calculations. These were included to stabilise the marking scheme of the listener.

The agreed music types are detailed in table 2, below: -

Table 2 Listening test material

| Sound type | Classification |
|--------------------------------|-------------------------|
| Jazz (Sony) | Jazz music |
| Jazz (ITS) | Jazz music |
| Percussion | Instrument |
| Harpsichord | Instrument |
| Piano | Instrument |
| English female speech | Spoken word |
| Chart female vocal POP | POP music (Normal) |
| Mainstream UK POP | POP music (Normal) |
| Mainstream electronic POP | POP music (Synthesiser) |
| Mainstream electronic POP (#2) | POP music (Synthesiser) |

1.2 Equipment used.

Tests were conducted in the "listening room" at ITS-RTCs laboratory. This room complies with the IEC 268-13 standard for listening rooms. The listeners all used Sony MDR-CD780 headphones. These were selected as offering high quality sound reproduction together with a degree of acoustic isolation. (They were closed front/open back design). The laboratory's experienced listeners confirmed their suitability for the test.

It was decided that all the formats to be tested should be created with their respective CODECs but then saved as WAV files. This would enable the programme segments to be stored onto audio CDs and replayed through the laboratory's high quality playback system. The playback system consisted of a "Meridian 596" CD/DVD player connected to a "Linn Kolektor" pre-amplifier. These are acknowledged as top-end products on the British Hi-fi market. The parallel line outputs from the Linn were fed to headphone amplifiers. Listening levels were individually adjusted for each listener but not altered once set.

1.3 Listening test material creation

All of the ATRAC3 and ATRAC3plus files were encoded and converted to WAV format by Sony Corporation. The laboratory takes on trust that these files were encoded with good intention and at the correct bit rates.

WMA files were encoded by ITS-RTC at a *constant* bit rate using the command-line interpreter of Windows Media Encoder version 9. MP3 files were encoded using MusicMatch Jukebox version 7.5 at *constant* bitrate.

The laboratory used Ahead Nero version 5.5.10.0 to compile audio CD's (44.1kHz, 16-bit) for both WMA and MP3 files. The CDDA information was then extracted using WinDAC32 to provide 44.1kHz 16-bit WAV files based on the MP3 and WMA tracks. The digital audio was extracted at a low speed (x4) from a Sony DRU-500A drive.

Each 'run' consisted of three segments according to the Sony RAB listening test method. Cool Edit Pro 2.0. was used to join the appropriate WAV files according to the random segment order such that an output WAV file then represented one 'run'.

The following rules were applied to each WAV file that made up a 'run': -

- i) Short tone at the start to indicate a new 'run'
- ii) 2 seconds into track the reference segment is played
- iii) Each subsequent segment has a gap of approximately 2 seconds between end and start of new segment
- iv) There is a 6 second pause between runs for the listener to mark their score-sheet.

These WAV files were then recorded onto a TDK metallic CD-R using Ahead Nero 5.5.10.0. A total of 4 CD's were produced, one for each of the listening test sections detailed in Table 1

A member of ITS-RTC quality control staff crosschecked the discs and run-order.

1.4 Listening Test Procedure

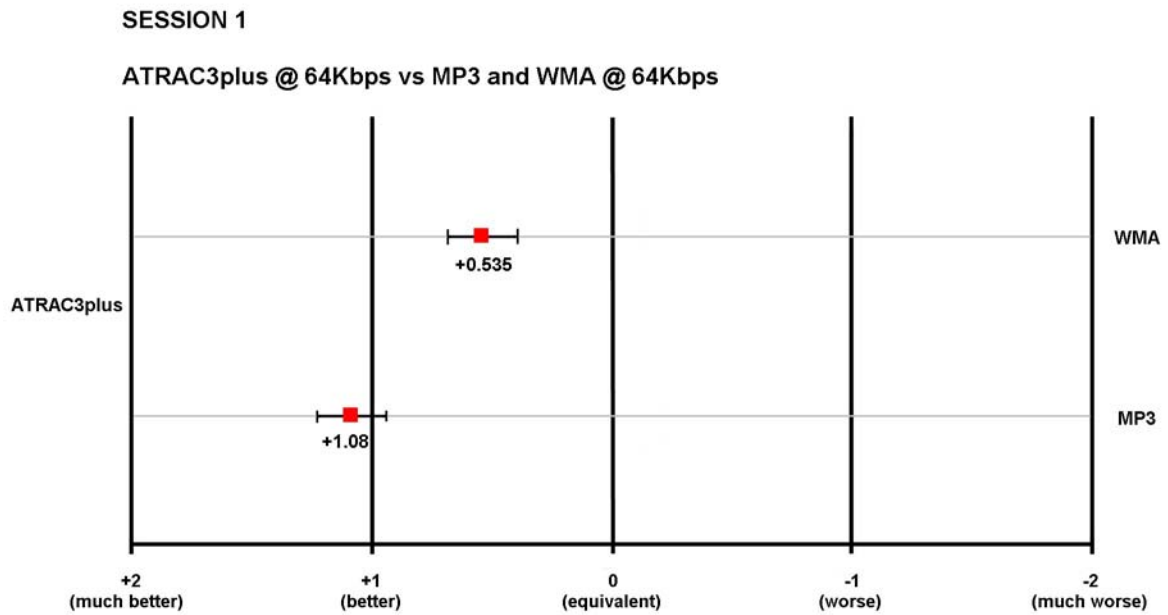
The listening test was performed over two days with 5 listeners on each day. Adequate breaks were provided for the listeners between each test CD.

The listeners rated each test on a paper score sheet

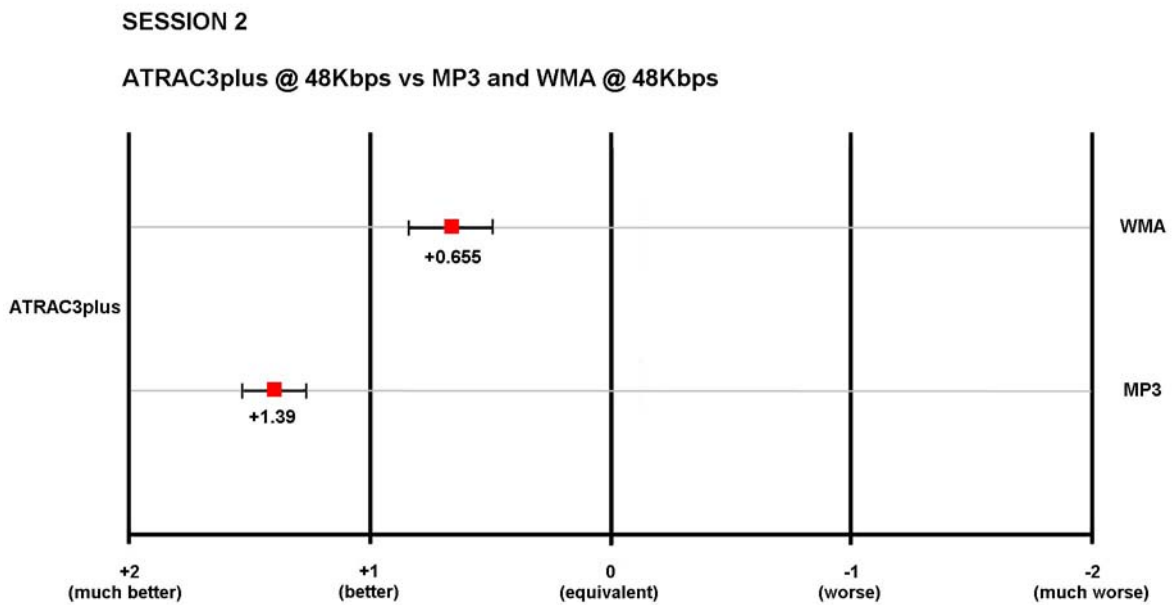
1.5 Statistical Analysis of results

The listening test results were analysed using a special software program, developed by Sony Corporation. The results from the statistical analysis and the raw data from the listeners' completed forms have been compared by a member of ITS-RTC statistics department who found the results to be consistent with expectations.

2. Results



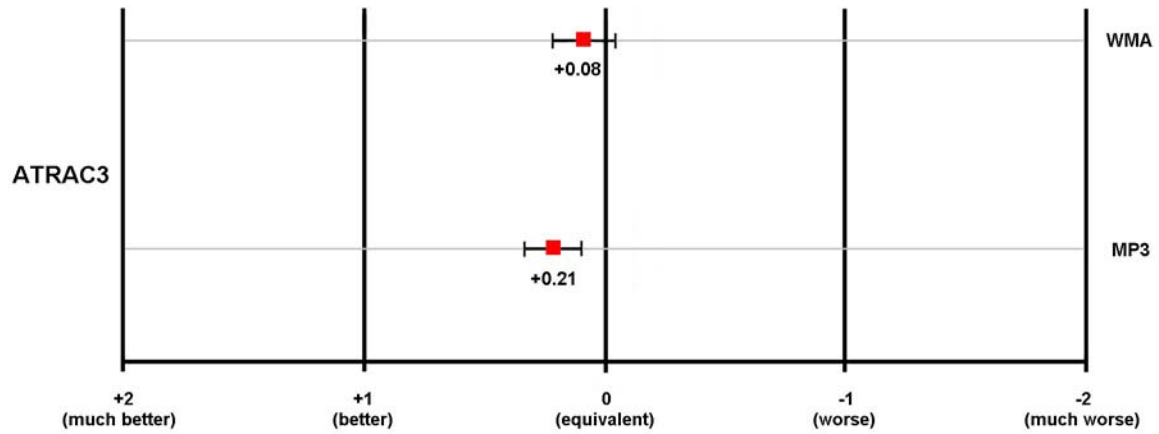
The results show that the ATRAC3plus CODEC system was favoured over the MP3 and WMA CODEC systems at 64Kbps.



The 48Kbps tests show that ATRAC3plus was again preferred, to a greater extent.

SESSION 3

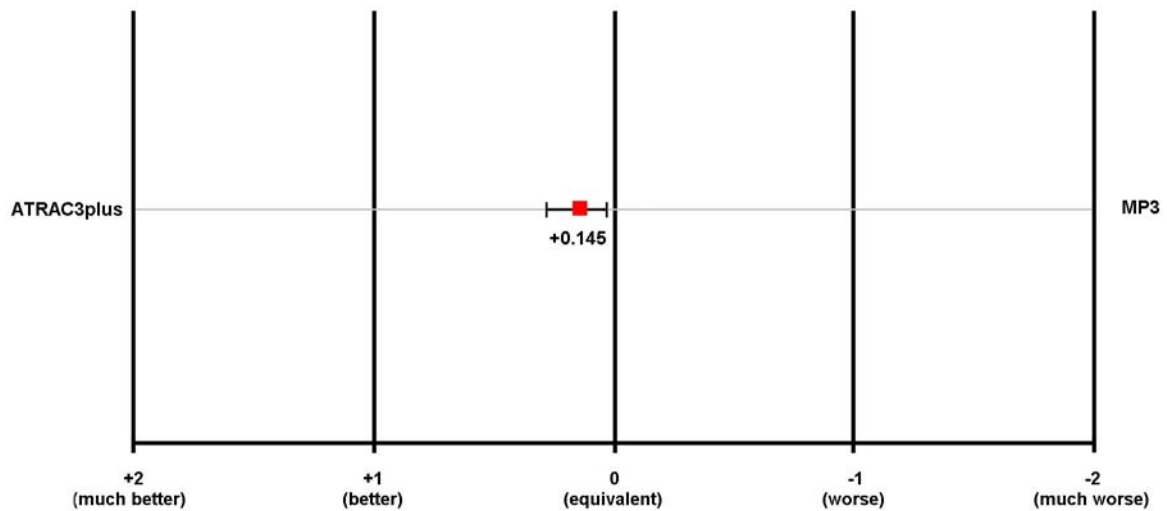
ATRAC3 @ 132Kbps vs MP3 and WMA @ 128Kbps



ATRAC3 (132Kbps) tracks were rated very slightly better than MP3 and WMA at 128Kbps, these CODEC systems being roughly equal when looking at the confidence levels.

SESSION 4

ATRAC3plus @ 64Kbps vs MP3 @ 128Kbps



ATRAC3plus 64Kbps was very close to MP3 128Kbps with a slight preference towards ATRAC3plus 64Kbps.

The raw results giving the 95% confidence levels are given in [Appendix 1](#)

3 Technical Tests

3.1 Spectrum Analysis

Spectrum analysis of test tones was made in the digital domain, using the Cool Edit software package. Two tones were used, a 1kHz at peak bits and a 1kHz tone at –60dB below peak bits. The full results are shown in **Appendix 2**. These tests are not designed to provide a comprehensive assessment or comparison of the products; they simply show that the CODEC systems are working, as expected, within certain defined parameters of noise, distortion and other spurious effects. The tests are limited to static tones and do not test the CODEC systems' dynamic properties.

The spectrums show that in a majority of cases ATRAC3plus has significantly lower sideband noise than the other CODEC systems for all bit rates tested.

3.2 Frequency response

A frequency sweep was recorded on each CODEC system and the decoded output plotted via the analogue output of a PC soundcard of known flat response. The resulting graphs are shown in **Appendix 3**. These plots simply show the potential of the CODEC system under very simple signal conditions, they do not necessarily show the response or bandwidth under complex signal conditions.

All graphs show flat responses but at lower bit rates both WMA and MP3 have reduced bandwidth at high frequencies. ATRAC3plus does not show this effect.

APPENDIX I

Raw Results from Listening Tests

Raw results from Listening tests

These results were generated using the statistical results software.

Section 1 (ATRAC3plus, 64Kbps comparison)

of subjects = 10 (gN = 400) (t = 2.093000 gt = 1.960000)

mean s.d 95% conf.intvl

CODEC_MP364Kbps : 1.0800 0.9687 [0.9366, 1.2234]

CODEC_WMA64Kbp : 0.5350 0.9765 [0.3905, 0.6795]

Section 2 (ATRAC3plus, 48Kbps comparison)

of subjects = 10 (gN = 400) (t = 2.093000 gt = 1.960000)

mean s.d 95% conf.intvl

CODEC_MP348Kbps: 1.3900 0.8785 [1.2600, 1.5200]

CODEC_WMA48Kbps: 0.6550 1.1324 [0.4874, 0.8226]

Section 3 (ATRAC3, 132Kbps comparison)

of subjects = 10 (gN = 400) (t = 2.093000 gt = 1.960000)

mean s.d 95% conf.intvl

CODEC_MP3128Kbps : 0.2100 0.7931 [0.0926, 0.3274]

CODEC_WMA128Kbp : 0.0800 0.8760 [-0.0497, 0.2097]

Section 4 (ATRAC3plus, 64Kbps comparison)

of subjects = 10 (gN = 200) (t = 2.093000 gt = 1.960000)

mean s.d 95% conf.intvl

CODEC_MP3128Kbps : 0.1450 0.8351 [0.0214, 0.2686]

APPENDIX II

Signal Analysis

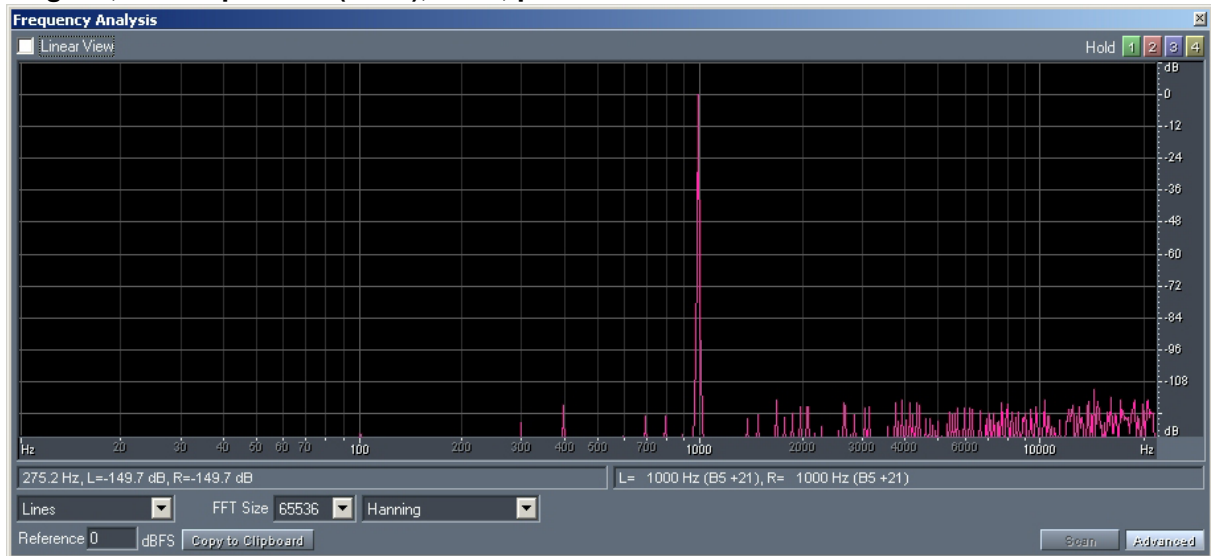
Technical Measurements

The frequency domain analysis of the encoded test signals was made in the digital domain using CoolEdit Pro version 2.0. Analysis WAV files are 44.1KHz, 16-bit audio.

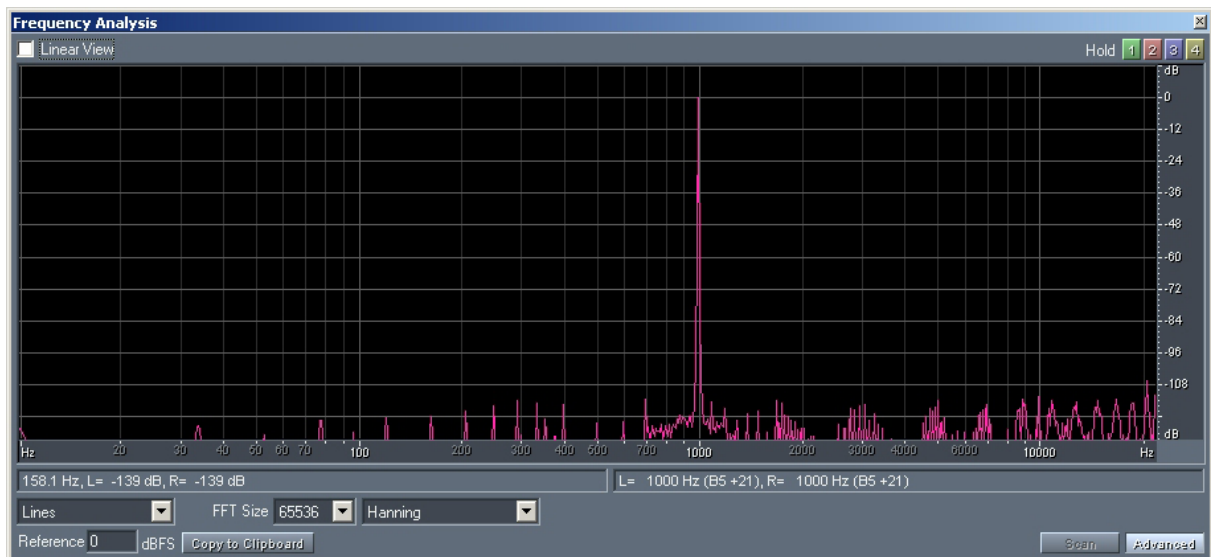
A 'Hanning' time window and 65536 point FFT algorithm was used in the calculation.

128Kbps, 1kHz tone analysis, peak bits

Original, Uncompressed (WAV), 1kHz, peak bits



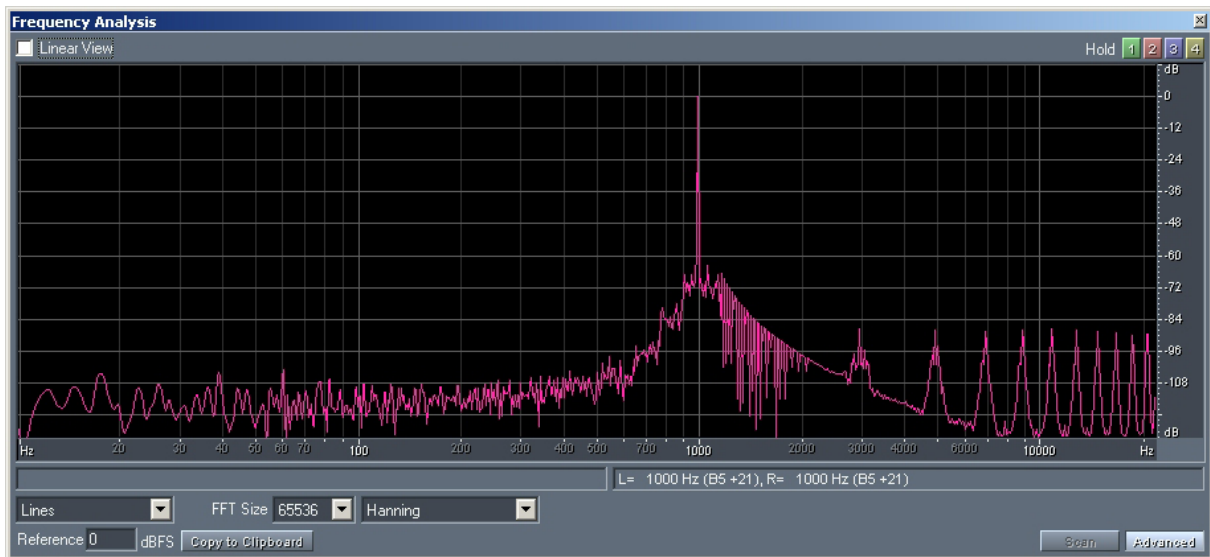
ATRAC3 132Kbps, 1kHz tone, peak bits



MP3 128Kbps, 1kHz tone, peak bits

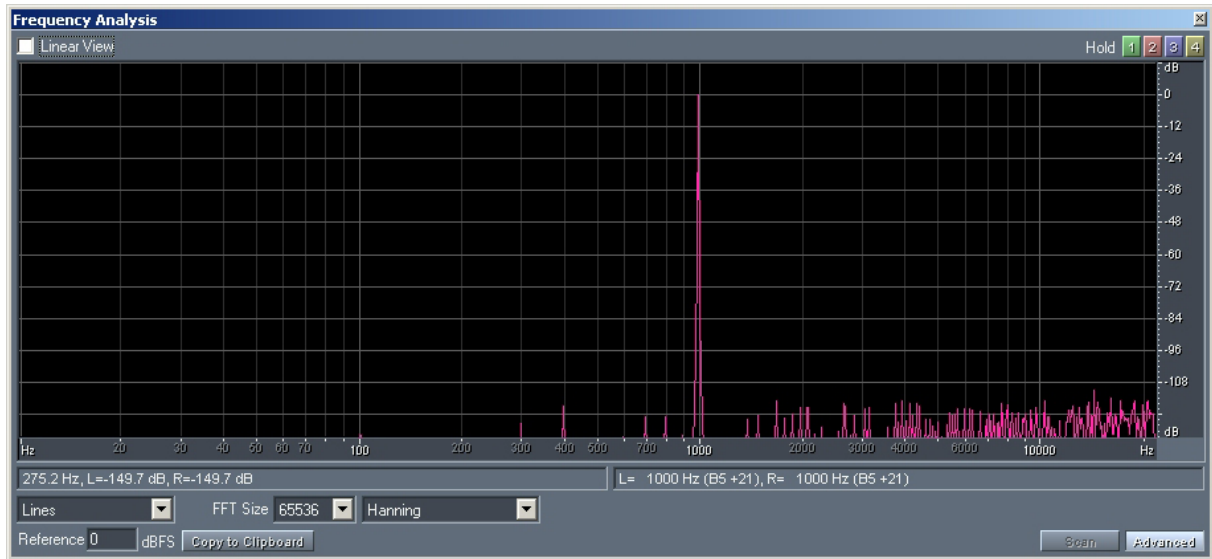


WMA 128Kbps, 1kHz tone, peak bits

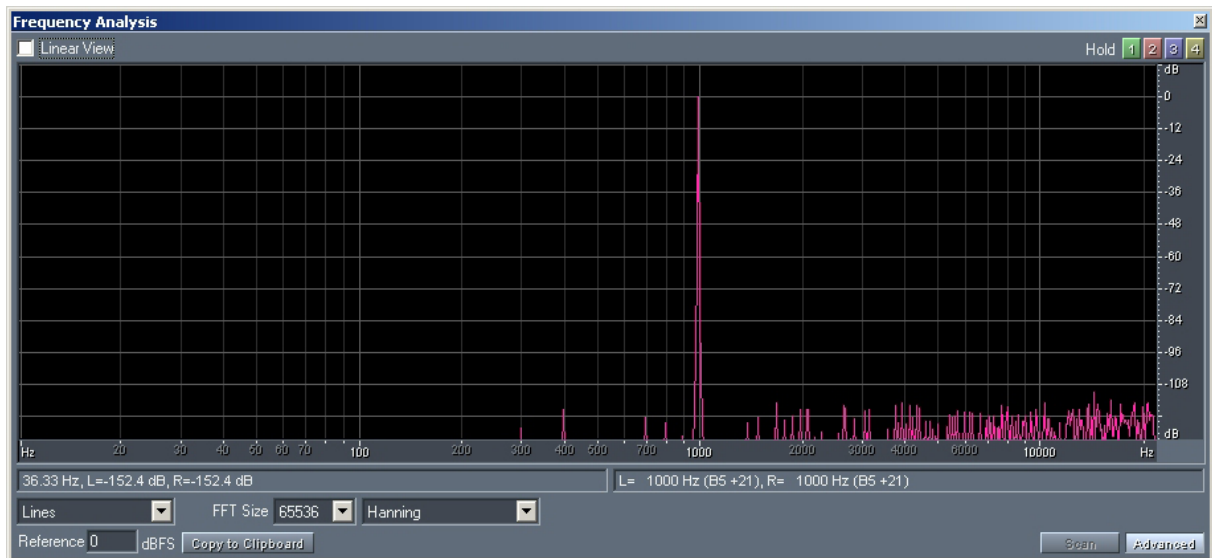


64Kbps, 1kHz tone analysis

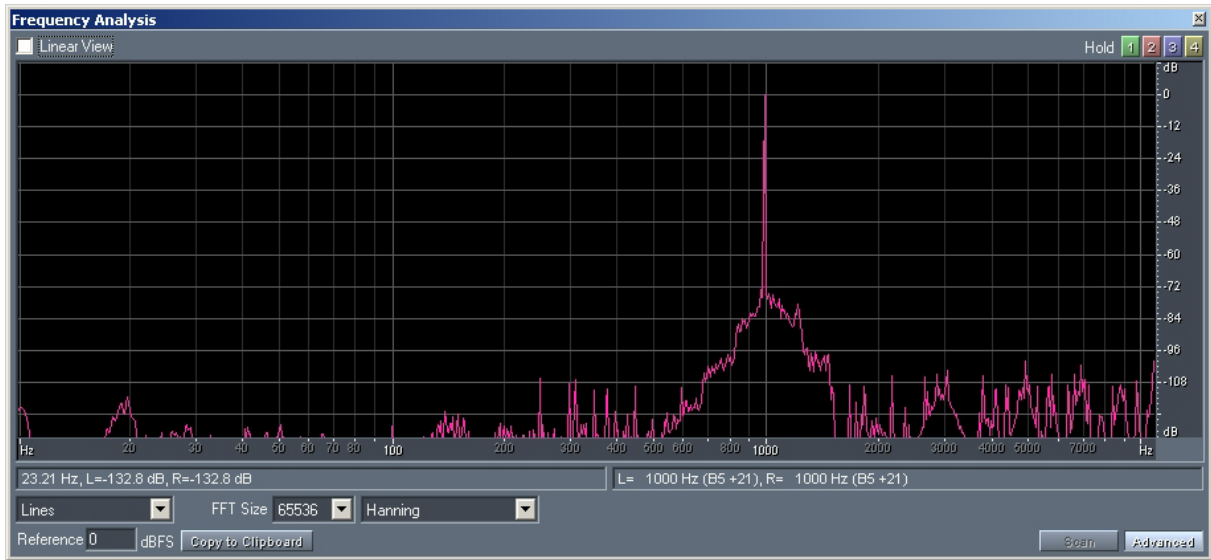
Original, uncompressed (WAV), 1kHz tone, peak bits



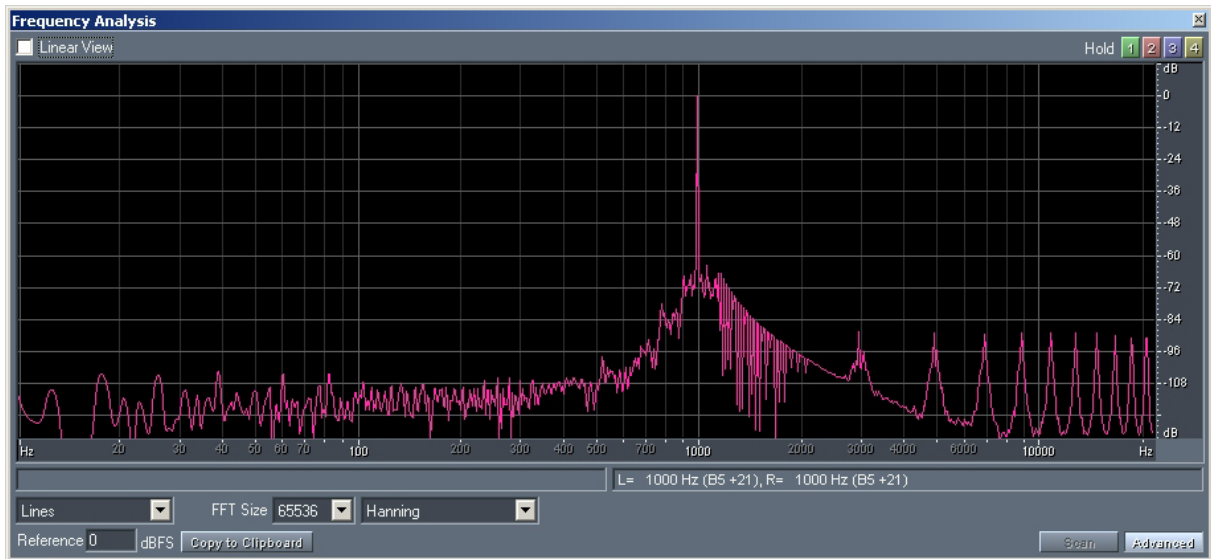
ATRAC3plus, 64Kbps, 1kHz tone, peak bits



MP3, 64Kbps, 1kHz tone, peak bits

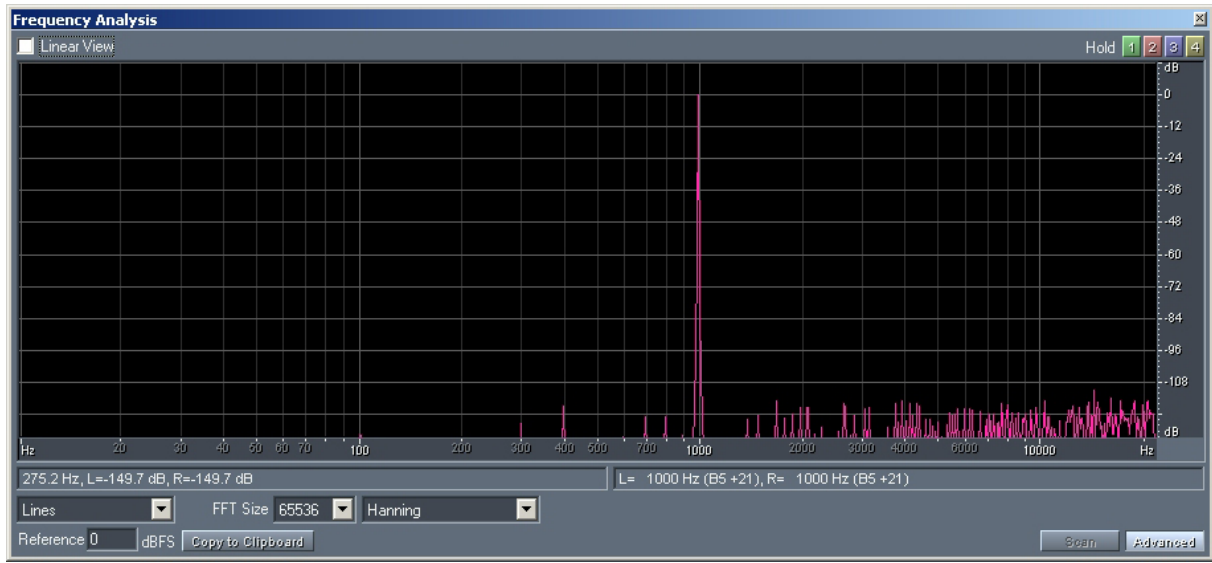


WMA, 64Kbps, 1kHz tone, peak bits

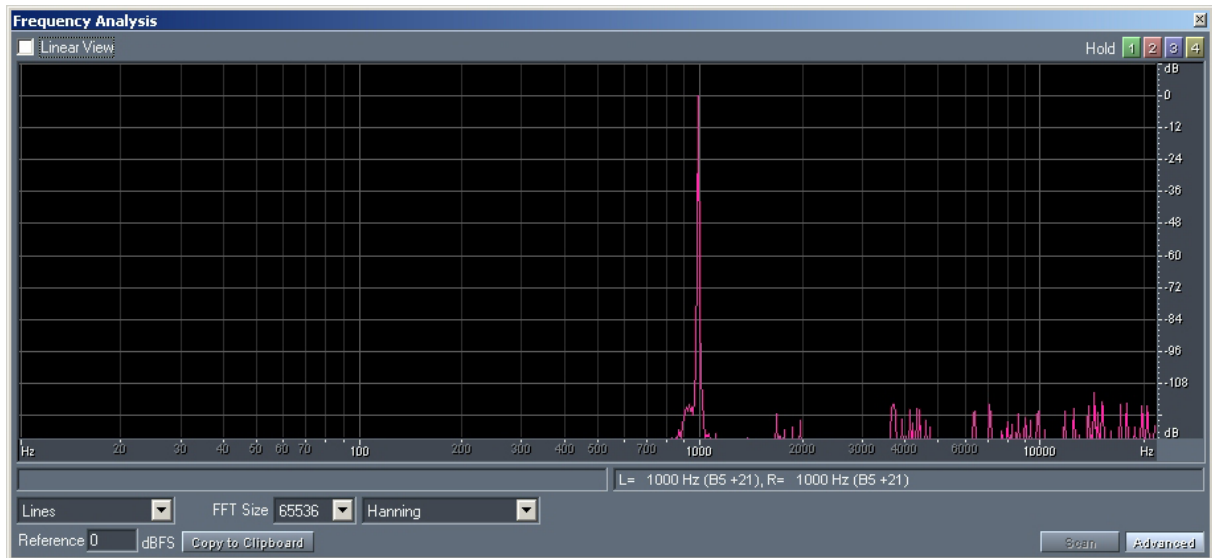


48Kbps, 1kHz tone analysis

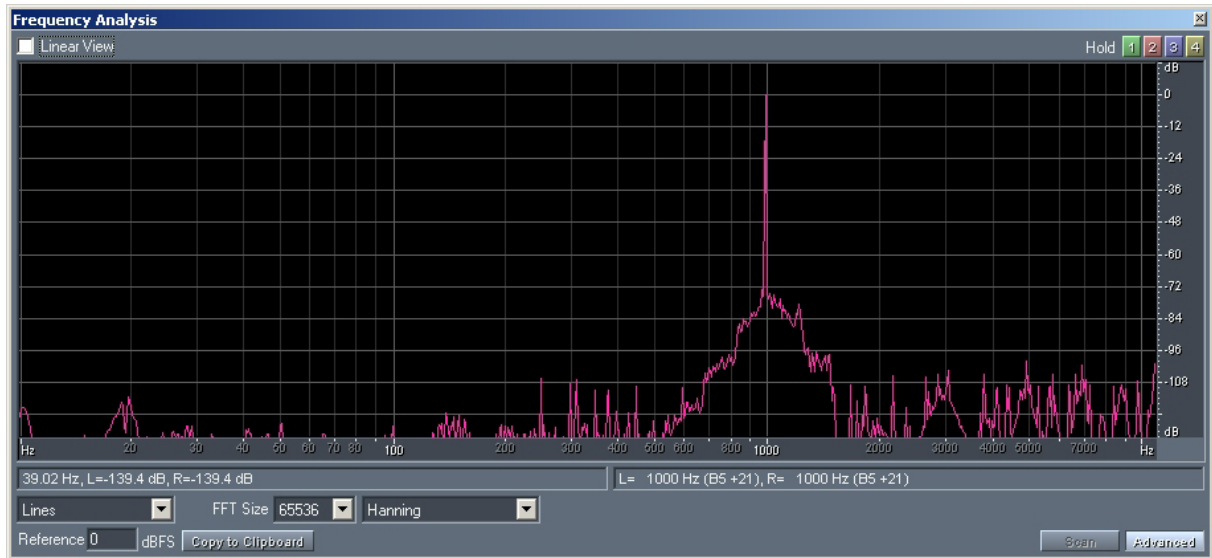
Original, uncompressed (WAV), 1kHz tone, peak bits



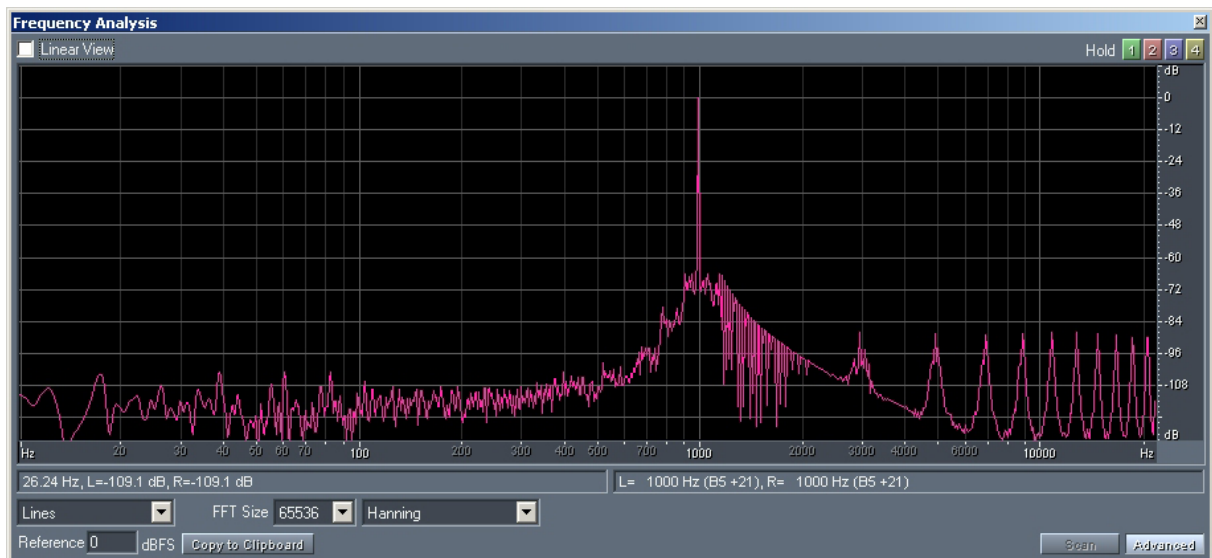
ATRAC3plus 48Kbps, 1kHz tone, peak bits



MP3 48Kbps, 1kHz tone, peak bits

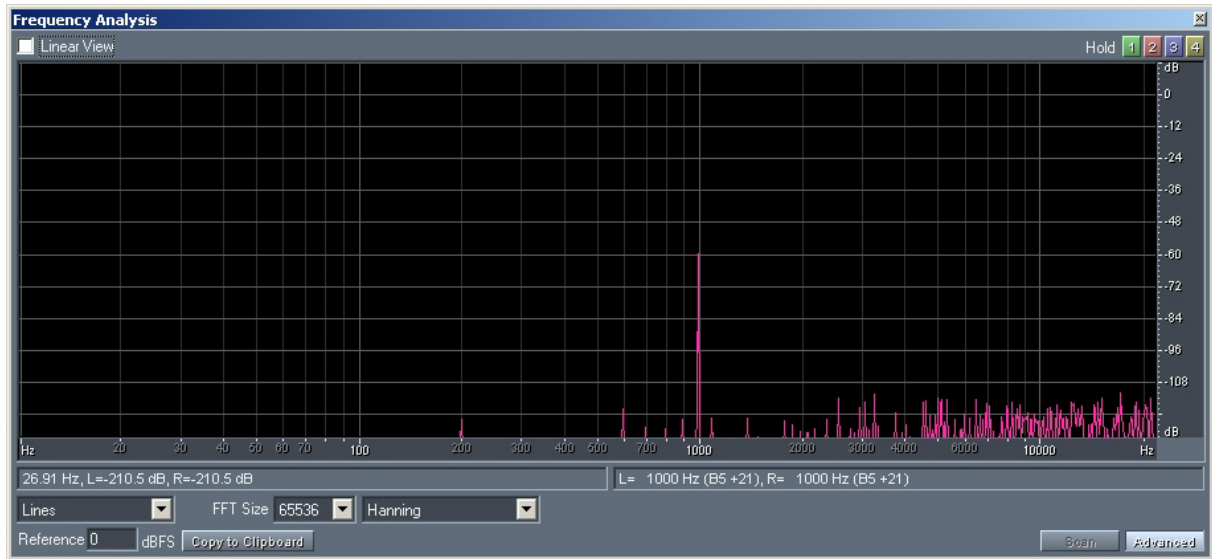


WMA 48Kbps, 1kHz tone, peak bits

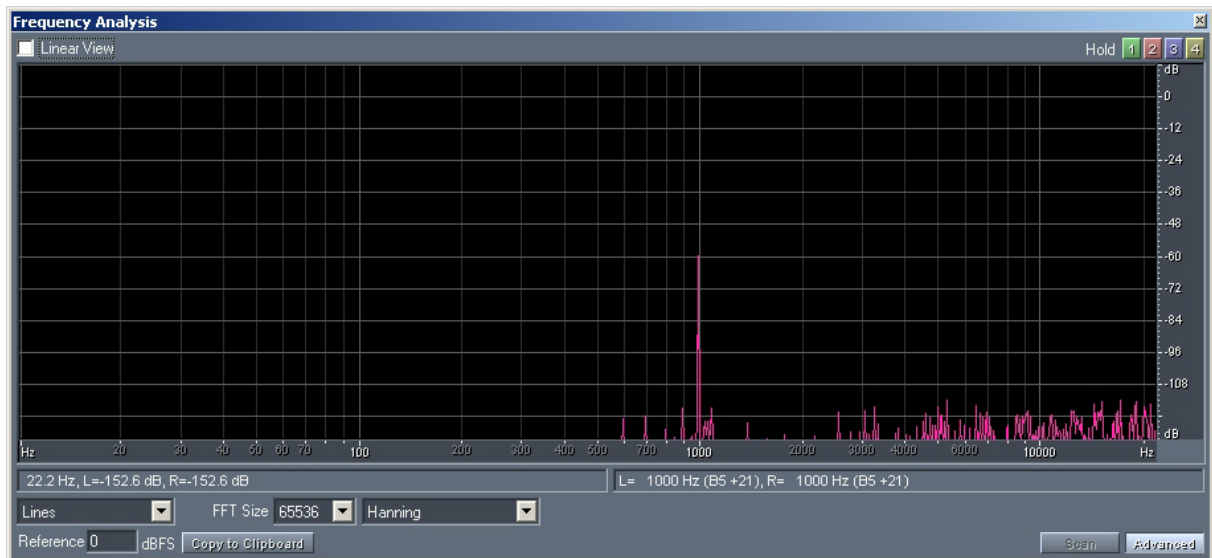


128Kbps 1kHz tone analysis at -60dB

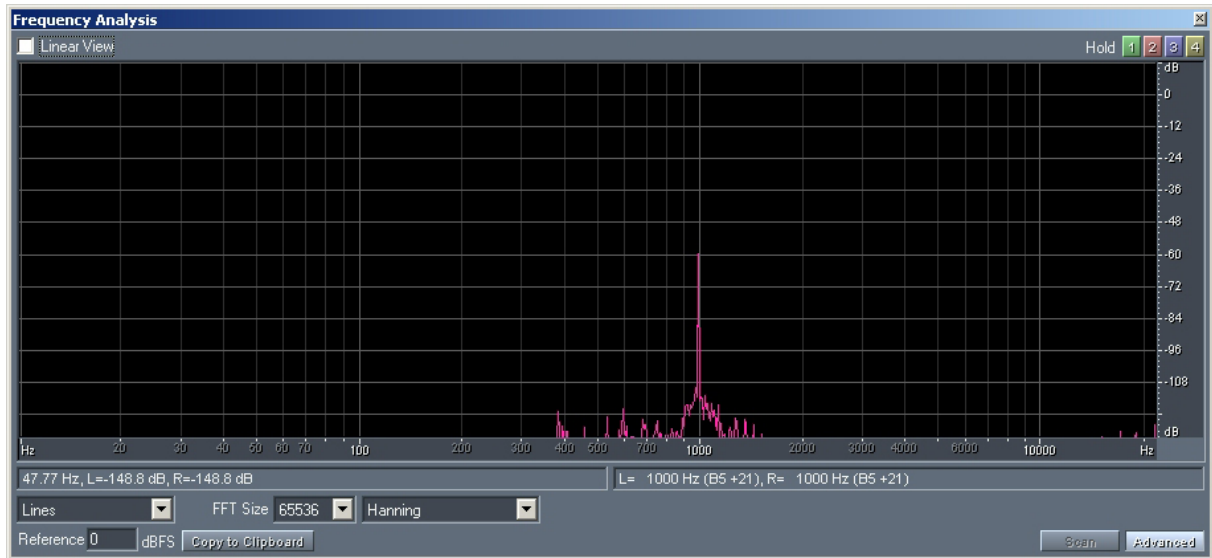
Original, Uncompressed (WAV), 1kHz, -60dB



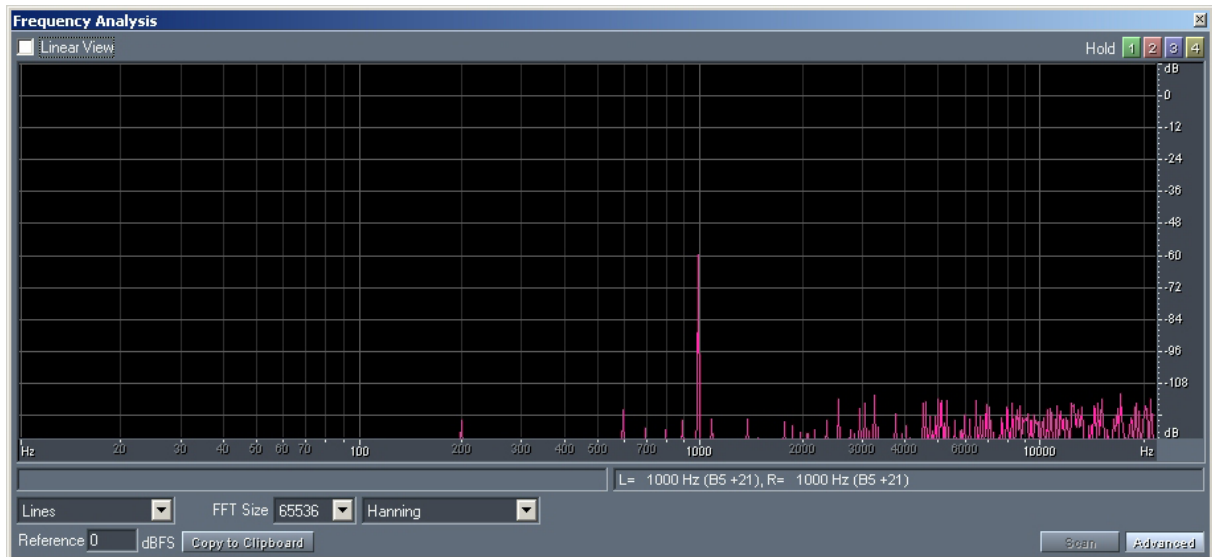
ATRAC3 132Kbps, 1kHz tone, -60dB



MP3 128Kbps, 1kHz tone, -60dB

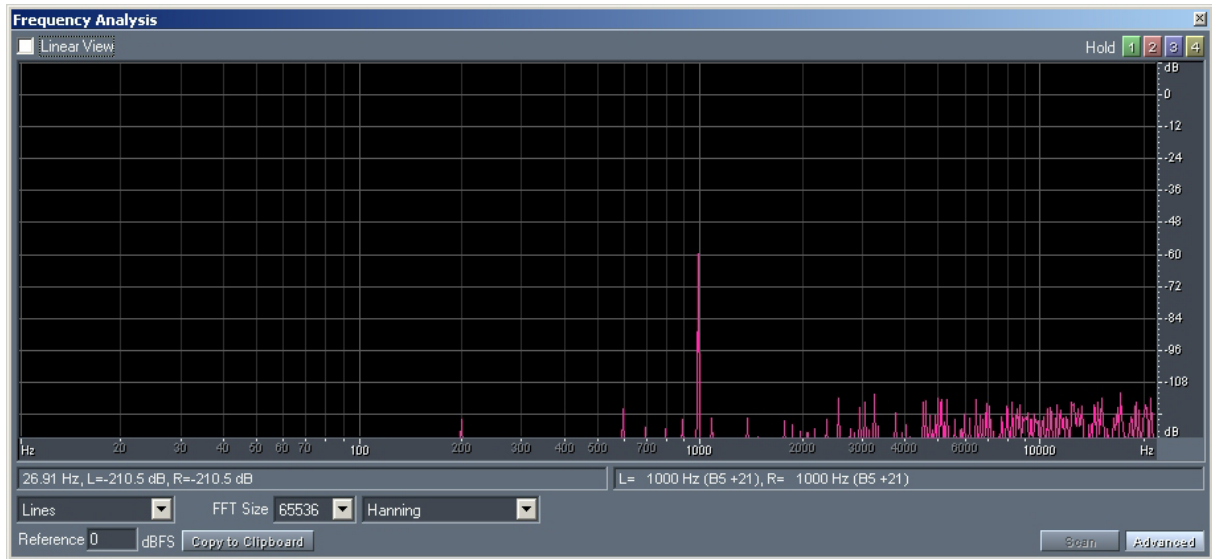


WMA 128Kbps, 1kHz tone, -60dB

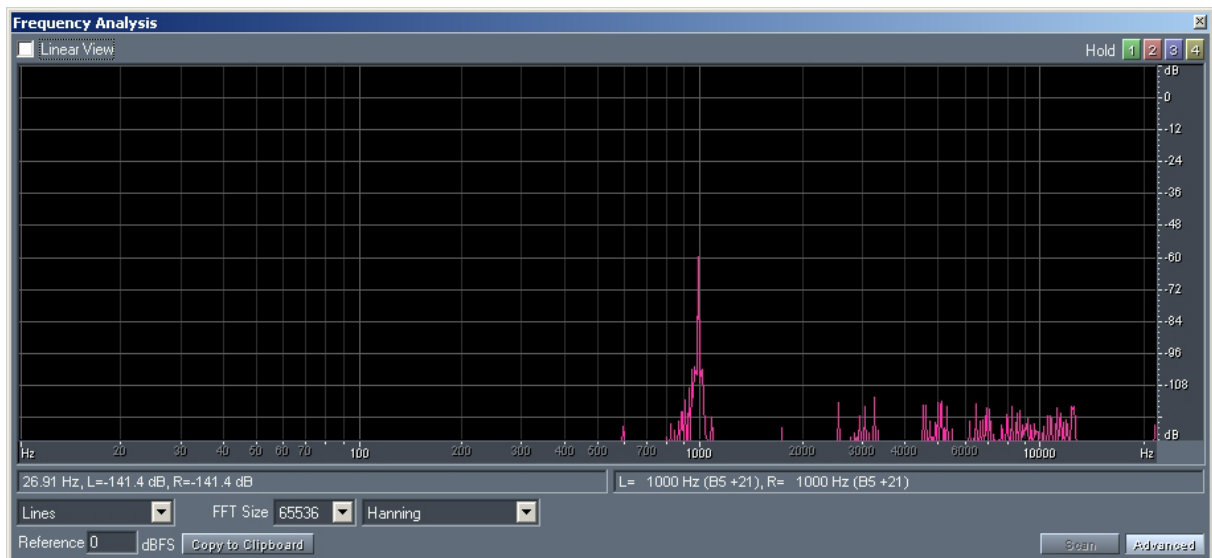


64Kbps 1kHz tone analysis at -60dB

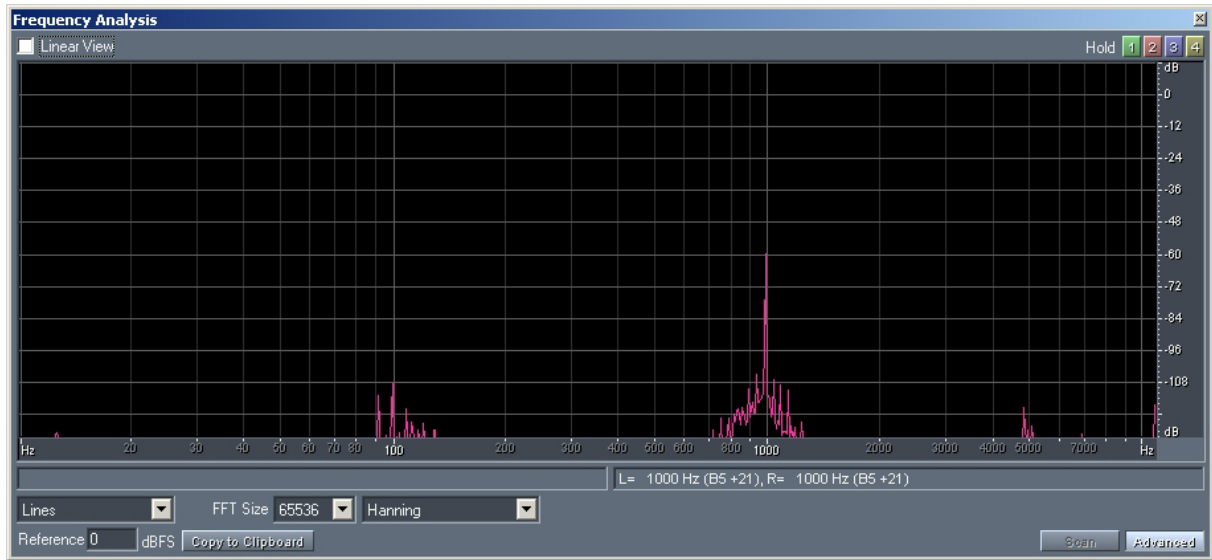
Original, Uncompressed (WAV), 1kHz, -60dB



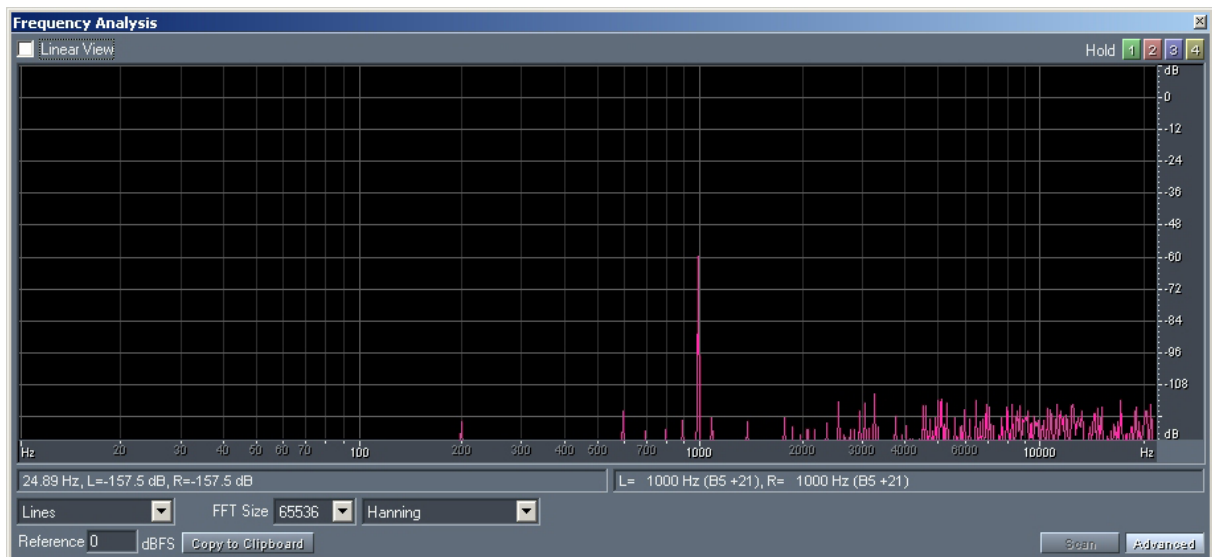
ATRAC3plus 64Kbps, 1kHz tone, -60dB



MP3 64Kbps, 1kHz tone, -60dB

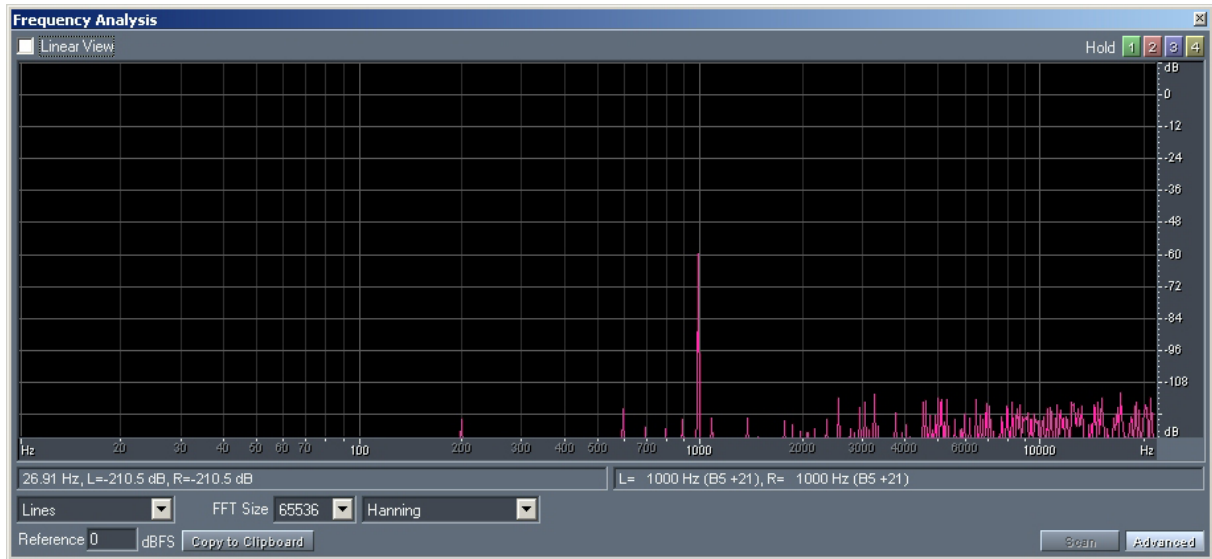


WMA 64Kbps, 1kHz tone, -60dB

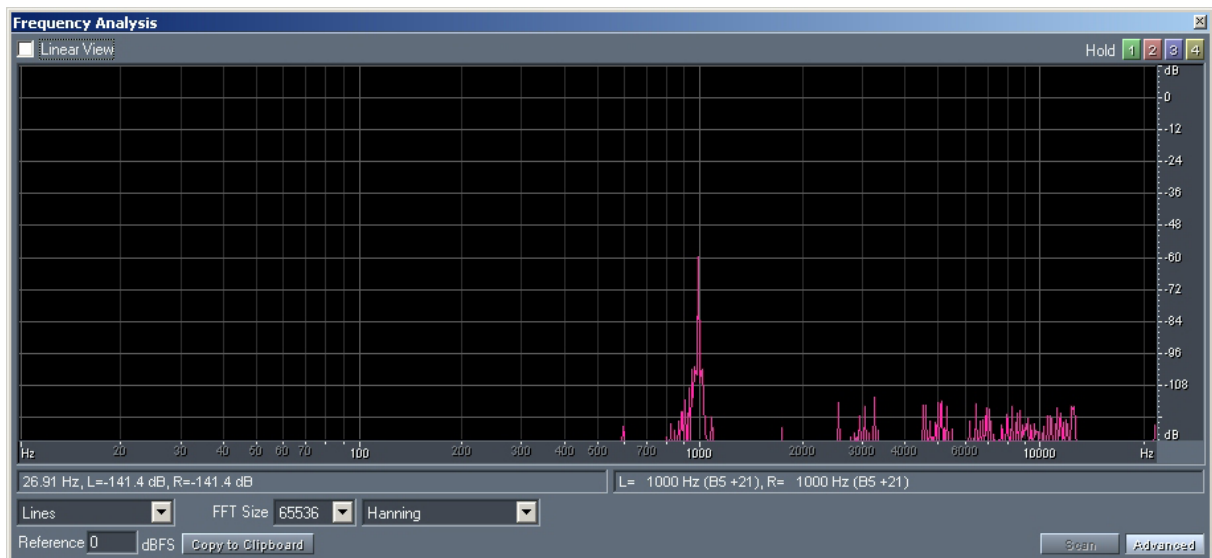


48Kbps 1kHz tone analysis at -60dB

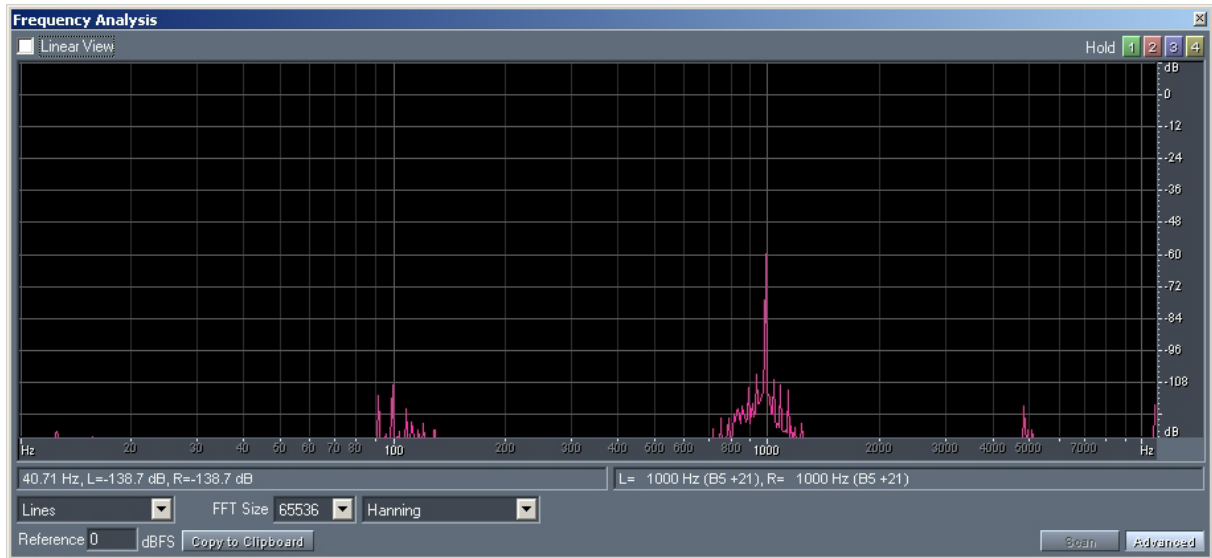
Original, Uncompressed (WAV), 1kHz, -60dB



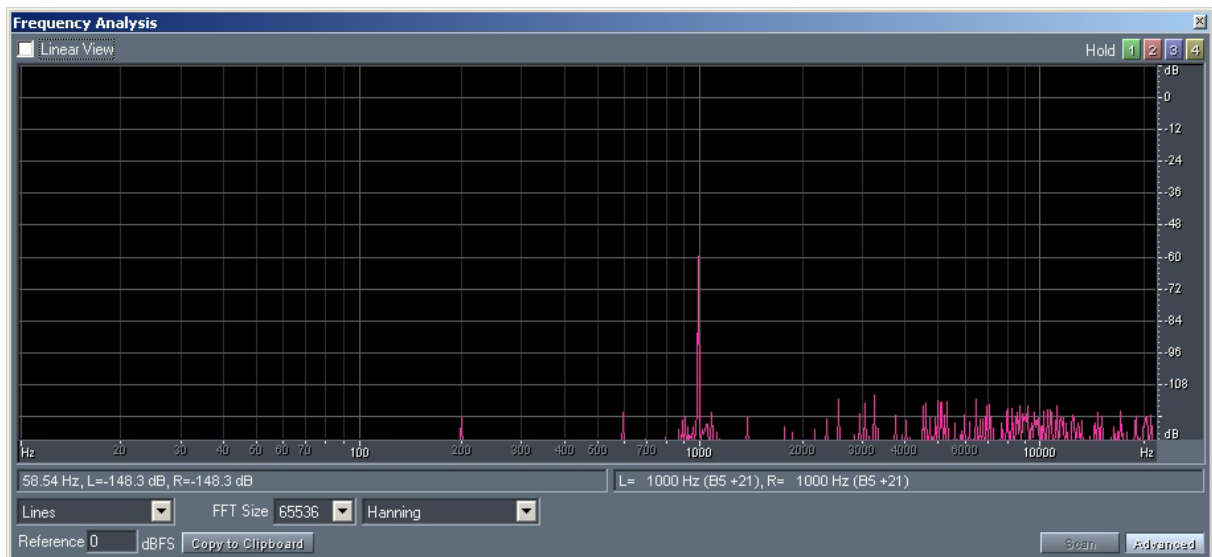
ATRAC3plus 48Kbps, 1kHz tone, -60dB



MP3 48Kbps, 1kHz tone, -60dB



WMA 48Kbps, 1kHz tone, -60dB



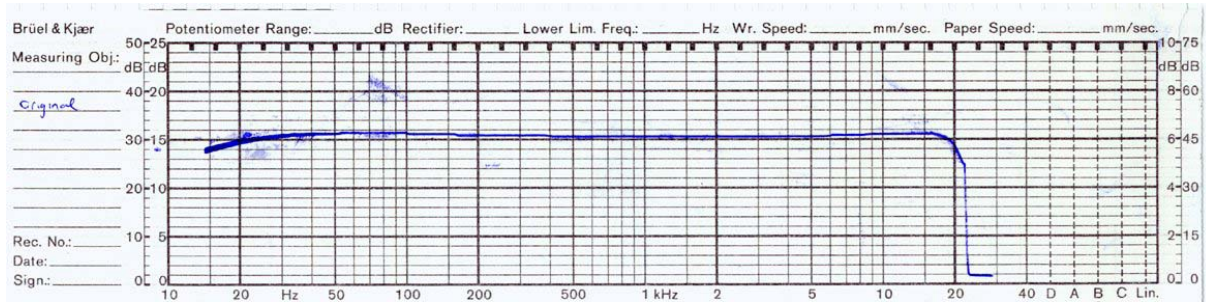
APPENDIX III

Frequency Response Plots

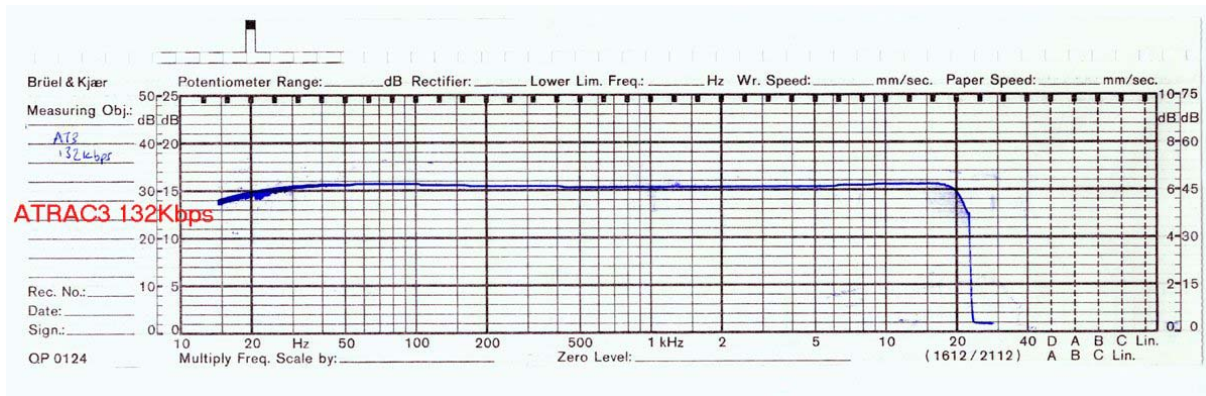
Technical Measurements – Frequency response plots

Original frequency response curves from approximately 10Hz to 22kHz (source: CBS test disc)

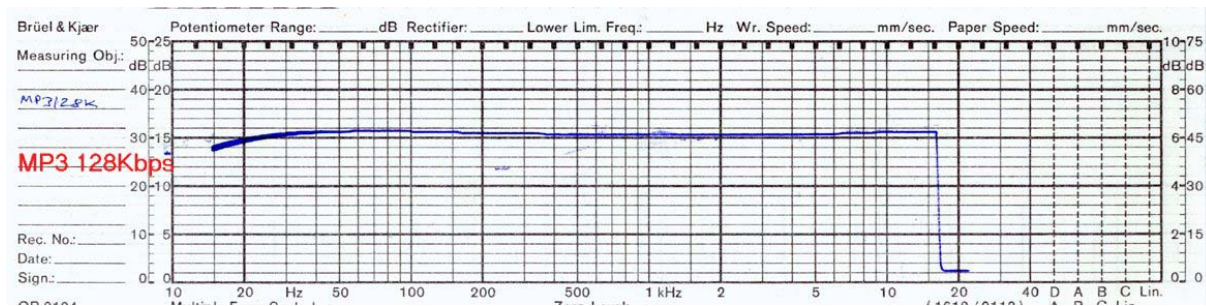
Original WAV, 44.1kHz sampling rate, 16-bit audio



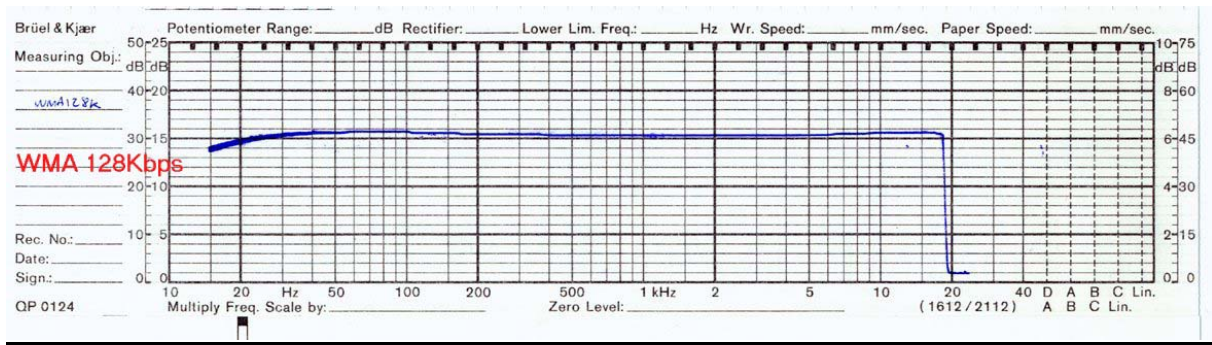
ATRAC3, 132 Kbps



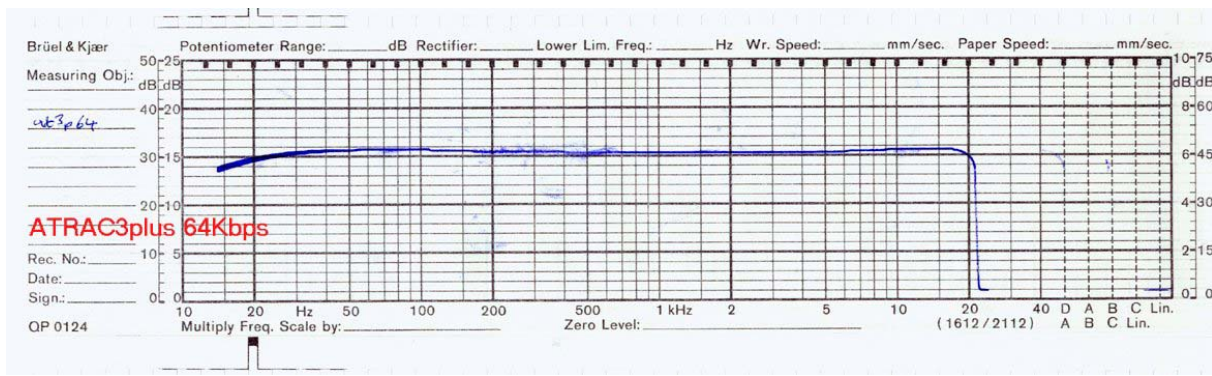
MP3, 128Kbps



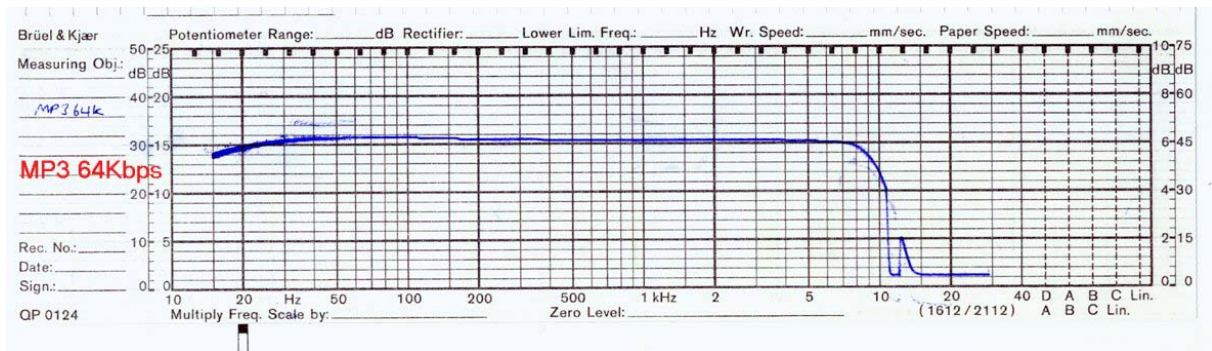
WMA, 128Kbps



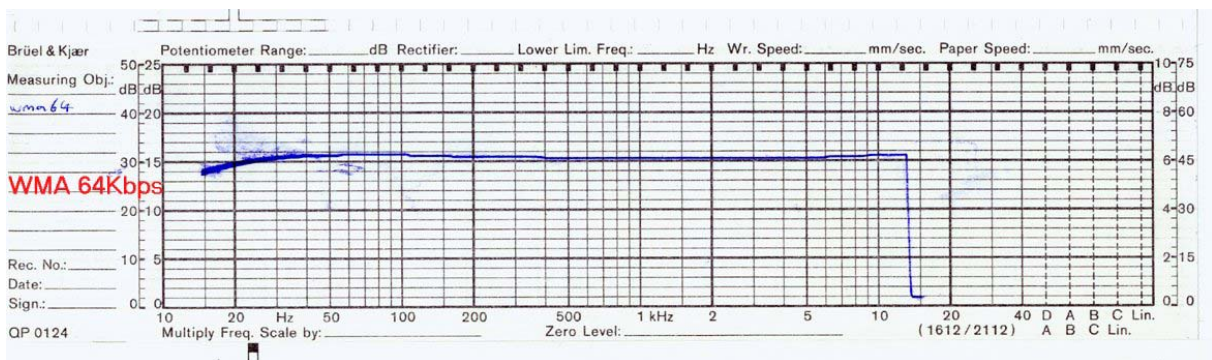
ATRAC3plus, 64Kbps



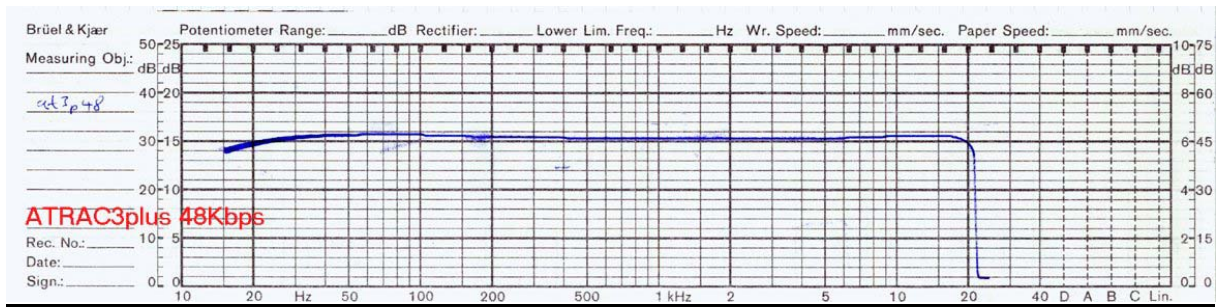
MP3, 64Kbps



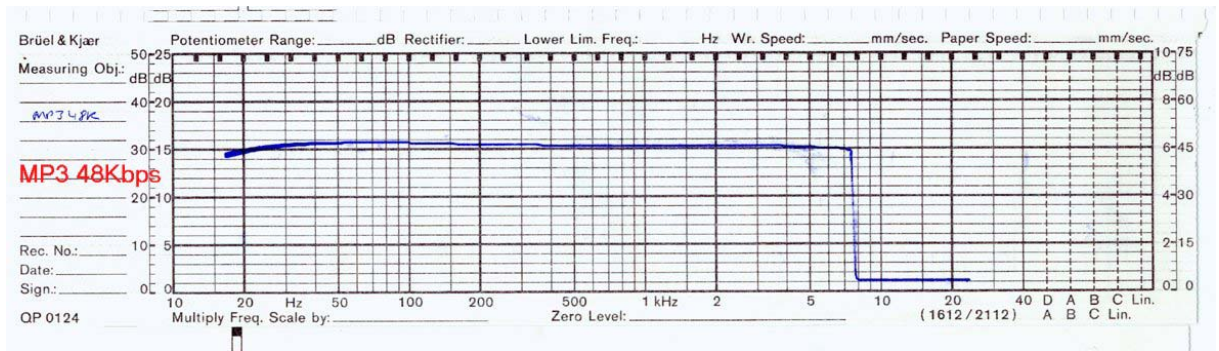
WMA, 64Kbps



ATRAC3plus, 48Kbps



MP3, 48Kbps



WMA, 48Kbps

