

TS 26.074 V 2.0.0 (1999-12)

Technical Specification

**3rd Generation Partnership Project (3GPP);
TSG-SA Codec Working Group
Mandatory Speech Codec speech processing functions;
AMR Speech Codec Test Sequences**



Reference

TSG-SA4-W1 (<Shortfilename>.PDF)

Keywords

Adaptive Multi-Rate, Mandatory speech coder

3GPP

Postal address

Office address

Internet

secretariat@3gpp.org
Individual copies of this deliverable
can be downloaded from
<http://www.3gpp.org>

Copyright Notification

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.
© 1999, 3GPP Organizational Partners (ARIB, CWTS, ETSI, T1, TTA, TTC).
All rights reserved.

Contents

Intellectual Property Rights	4
Foreword.....	4
1 Scope	5
2 Normative references.....	5
3 Definitions and abbreviations	5
3.1 Definitions	5
3.2 Abbreviations.....	5
4 General.....	6
5 Test sequence format	6
5.1 File format	6
5.2 Codec homing.....	6
6 Speech codec test sequences.....	7
6.1 Codec configuration.....	7
6.2 Speech codec test sequences.....	7
6.2.1 Speech encoder test sequences.....	7
6.2.2 Speech decoder test sequences.....	8
6.2.3 Codec homing sequence.....	8
7 Test sequences for source controlled rate operation	8
7.1 Codec configuration.....	9
7.2 Test Sequences	9
7.2.1 Test sequences for background noise estimation	9
7.2.2 Test sequences for pitch, tone and complex signal detection.....	10
7.2.3 Real speech and tones	10
7.2.4 Test sequence for signal-to-noise ratio estimation	10
8 Sequences for finding the 20 ms framing of the adaptive multi-rate speech encoder	10
8.1 Bit synchronisation.....	10
8.2 Frame synchronisation.....	11
8.3 Formats and sizes of the synchronisation sequences	11
9 Trau Testing with 8 Bit A- and μ -law PCM Test Sequences	12
History.....	13

Intellectual Property Rights

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects, Working Group 4 (Codec).

The contents of this TS may be subject to continuing work within the 3GPP and may change following formal TSG-S4 approval. Should the TSG-S4 modify the contents of this TS, it will be re-released with an identifying change of release date and an increase in version number as follows:

Version m.t.e

where:

- m indicates [major version number]
- x the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- y the third digit is incremented when editorial only changes have been incorporated into the specification.

1 Scope

The present document specifies the digital test sequences for the adaptive multi-rate (AMR) speech codec. These sequences test for a bit exact implementation of the adaptive multi-rate speech transcoder (TS 26.090 [2]), voice activity detection (TS 26.094 [5]), comfort noise (TS 26.092 [3]), and source controlled rate operation (TS 26.092 [4]).

2 Normative references

This TS incorporates by dated and undated reference, provisions from other publications. These normative references are cited at the appropriate places in the text and the publications are listed hereafter. For dated references, subsequent amendments to or revisions of any of these publications apply to this TS only when incorporated in it by amendment or revision. For undated references, the latest edition of the publication referred to applies.

- [1] TS 26.071 : "AMR Speech Codec; General Description"
- [2] TS 26.090 : "AMR Speech Codec; Speech Transcoding Functions"
- [3] TS 26.092 : "AMR Speech Codec; Comfort Noise Aspects"
- [4] TS 26.093 : "AMR Speech Codec; Source Controlled Rate Operation"
- [5] TS 26.094 : "AMR Speech Codec; Voice Activity Detector"
- [6] TS 26.091 : "AMR Speech Codec; Error Concealment of Lost Frames"
- [7] TS 26.073 : "AMR Speech Codec; ANSI C-code"

3 Definitions and abbreviations

3.1 Definitions

Definition of terms used in the present document can be found in TS 26.090 [2], TS 26.091 [6], TS 26.092 [3], TS 26.093 [4] and TS 26.094 [5].

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

[T.B.A]

4 General

Digital test sequences are necessary to test for a bit exact implementation of the adaptive multi-rate speech transcoder (TS 26.090 [2]), voice activity detection (TS 26.094 [5]), comfort noise generation (TS 26.092 [3]), and source controlled rate operation (TS 26.092 [4]).

The test sequences may also be used to verify installations of the ANSI C code in TS 26.073 [7].

Clause 5 describes the format of the files which contain the digital test sequences. Clause 6 describes the test sequences for the speech transcoder. Clause 7 describes the test sequences for the VAD, comfort noise and source controlled rate operation.

Clause 8 describes the method by which synchronisation is obtained between the test sequences and the speech codec under test.

[Clause 9 describes the alternative acceptance testing of the speech encoder and decoder in the TRAU by means of 8 bit A-law compressed test sequences on the A-Interface.]

5 Test sequence format

This clause provides information on the format of the digital test sequences for the GSM adaptive multi-rate speech transcoder (TS 26.090 [2]), voice activity detection (TS 26.094 [5]), comfort noise generation (TS 26.092 [3]), and source controlled rate operation (TS 26.092 [4]).

5.1 File format

The test sequence files in PC (little-endian) byte order are provided in archive files (ZIP format) which accompany the present document.

Following decompression, three types of file are provided:

- Files for input to the speech encoder: *.INP
- Files for comparison with the encoder output and for input to the speech decoder: *.COD
- Files for comparison with the decoder output: *.OUT
- One mode control file for the mode switching test T21.MOD

All file formats are described in TS 26.073 [7].

5.2 Codec homing

Each *.INP file includes two homing frames (see TS 26.073 [7]) at the start of the test sequence. The function of these frames is to reset the speech encoder state variables to their initial value. In the case of a correct installation of the ANSI-C simulation (TS 26.073 [7]), all speech encoder output frames shall be identical to the corresponding frame in the *.COD file. In the case of a correct hardware implementation undergoing testing, the first speech encoder output frame is undefined and need not be identical to the first frame in the *.COD file, but all remaining speech encoder output frames shall be identical to the corresponding frames in the *.COD file.

The function of the two homing frames in the *.COD files is to reset the speech decoder state variables to their initial value. In the case of a correct installation of the ANSI-C simulation (TS 26.073 [7]), all speech decoder output frames shall be identical to the corresponding frame in the *.OUT file. In the case of a correct hardware implementation undergoing testing, the first speech decoder output frame is undefined and need not be identical to first frame in the *.OUT file, but all remaining speech decoder output frames shall be identical to the corresponding frames in the *.OUT file.

6 Speech codec test sequences

This clause describes the test sequences designed to exercise the adaptive multi-rate speech transcoder (TS 26.090 [2]).

6.1 Codec configuration

The speech encoder shall be configured not to operate in the source controlled rate mode.

6.2 Speech codec test sequences

6.2.1 Speech encoder test sequences

Twenty-two encoder input sequences are provided. Note that for the input sequences T00.INP to T03.INP, the amplitude figures are given in 13-bit precision. The active speech levels are given in dBov.

- T00.INP - Synthetic harmonic signal. The pitch delay varies slowly from 18 to 143.5 samples. The minimum and maximum amplitudes are -997 and +971.
- T01.INP - Synthetic harmonic signal. The pitch delay varies slowly from 144 down to 18.5 samples. Amplitudes at saturation point -4096 and +4095.
- T02.INP - Sinusoidal sweep varying from 150 Hz to 3400 Hz. Amplitudes ± 1250 .
- T03.INP - Sinusoidal sweep varying from 150 Hz to 3400 Hz. Amplitudes ± 4000 .
- T04.INP - Female speech, active speech level: -19.4 dBov, flat frequency response.
- T05.INP - Male speech, active speech level: -18.7 dBov, flat frequency response.
- T06.INP - Female speech, ambient noise, active speech level: -35.0 dBov, flat frequency response.
- T07.INP - Female speech, ambient noise, active speech level: -25.0 dBov, flat frequency response.
- T08.INP - Female speech, ambient noise, active speech level: -15.6 dBov, flat frequency response.
- T09.INP - Female speech, car noise, active speech level: -35.5 dBov, flat frequency response.
- T10.INP - Female speech, car noise, active speech level: -26.1 dBov, flat frequency response.
- T11.INP - Female speech, car noise, active speech level: -15.8 dBov, flat frequency response.
- T12.INP - Male speech, ambient noise, active speech level: -34.9 dBov, flat frequency response.
- T13.INP - Male speech, ambient noise, active speech level: -24.8 dBov, flat frequency response.
- T14.INP - Male speech, ambient noise, active speech level: -15.0 dBov, flat frequency response.
- T15.INP - Male speech, babble noise, active speech level: -34.1 dBov, flat frequency response.
- T16.INP - Male speech, babble noise, active speech level: -24.3 dBov, flat frequency response.
- T17.INP - Male speech, babble noise, active speech level: -14.4 dBov, flat frequency response.
- T18.INP - Female speech, ambient noise, active speech level: -26.0 dBov, modified IRS frequency response, with many zero frames.
- T19.INP - Male speech, ambient noise, active speech level: -36.0 dBov, modified IRS frequency response, with many zero frames.
- T20.INP - Sequence for exercising the LPC vector quantisation codebooks and ROM tables of the codec.
- T21.INP - Speech sequence for mode switching test.

The output using these input sequences will be different depending on the tested adaptive multi-rate mode. In the notation used below <mode> should be changed to the number of the tested mode, i.e. one of 122, 102, 795, 74, 67, 59, 515, or 475.

The T00.INP and T01.INP sequences were designed to test the pitch lag of the GSM adaptive multi-rate speech encoder. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the T00_<mode>.COD and T01_<mode>.COD sequences, respectively.

The T02.INP and T03.INP sequences are particularly suited for testing the LPC analysis, as well as for finding saturation problems. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the T02_<mode>.COD and T03_<mode>.COD sequences, respectively.

The T04.INP and T05.INP sequences contain a lot of low-frequency components. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the T04_<mode>.COD and T05_<mode>.COD sequences, respectively.

The T18.INP and T19.INP sequences contain some “all zeros” frames (silence) in between segments of speech. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the T18_<mode>.COD and T19_<mode>.COD sequences, respectively.

The T20.INP sequence was designed to exercise the LPC code indices and the ROM table indices of the codec.

The sequences T06.INP to T17.INP were selected on the basis of bringing various input characteristics (background noise) and levels to the test sequence set. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the T06_<mode>.COD to T17_<mode>.COD sequences, respectively.

The T21.INP sequence was designed to test mode switching in the encoder. For testing mode switching this sequence is used together with the mode control file T21.MOD. See TS 26.073 [7] for the format of the mode control file. In a correct implementation, the resulting speech encoder output parameters shall be identical to those specified in the sequence T21.COD. Note that T21.COD contains parameter frames in different codec modes.

6.2.2 Speech decoder test sequences

Twenty-one times eight speech decoder input sequences TXX_<mode>.COD (XX = 00..20, <mode> = {122, 102, 795, 74, 67, 59, 515, or 475}) are provided for the static mode tests. These are the output of the corresponding TXX.INP sequences, one set per mode. In a correct implementation, the resulting speech decoder output shall be identical to the corresponding TXX_<mode>.OUT sequences.

The switching test decoder input T21.COD shall result in decoder output identical to the T21.OUT sequence. For the decoder switching test no special mode control file is needed since the mode information is included in the .COD file according to the file format (see TS 26.073 [7]).

6.2.3 Codec homing sequence

In addition to the test sequences described above, the homing sequences are provided to assist in codec testing. T22.INP contains one encoder-homing-frame. The sequences T22_<mode>.COD (<mode> = {122, 102, 795, 74, 67, 59, 515, or 475}) contain one decoder-homing-frame each for the corresponding mode. The use of these sequences is described in TS 26.071 [1].

All files are contained in the archive T.TGZ which accompanies the present document.

7 Test sequences for source controlled rate operation

This subclause describes the test sequences designed to exercise the VAD algorithm options 1 and 2 (TS 26.094), comfort noise (TS 26.092), and source controlled rate operation (TS 26.093).

Test sequences DTX*.* are to be used with VAD option 1. DTX1.*, DTX2.*, and DTX4.* shall be run only with speech codec mode MR122. Test sequence DTX3.* shall be run for all the speech codec modes (MR122, MR102, MR795, MR67, MR59, MR515 and MR475).

Test sequences DT2.* are to be used with VAD option 2. DT21.*, DT23.*, and DT24.* shall be run only with speech codec mode MR122. Test sequence DT22.* shall be run for all the speech codec modes (MR122, MR102, MR795, MR67, MR59, MR515 and MR475).

7.1 Codec configuration

The VAD, comfort noise and source controlled rate operation shall be tested in conjunction with the speech coder (TS 26.090). The speech encoder shall be configured to operate in the source controlled rate mode, with either VAD option 1 or VAD option 2.

7.2 Test Sequences

Each DTX test sequence consists of three files:

- Files for input to the speech encoder: *.INP
- Files for comparison with the encoder output and input to the speech decoder: *.COD
- Files for comparison with the decoder output: *.OUT

The *.COD and *.OUT file names has the format DTxA_<mode>.*, where "x" is the VAD option (X for option 1 and 2 for option 2), "A" is the test case number (1, 2, 3 or 4) and <mode> is the speech codec mode.

In a correct implementation, the speech encoder parameters generated by the *.INP file shall be identical to those specified in the *.COD file; and the speech decoder output generated by the *.COD file shall be identical to that specified in the *.OUT file.

Sequence name	No. of frames	Size (bytes)		
		*.INP	*.COD	*.OUT
DTX1	710	227 200	355 000	227 200
DTX2	898	287 360	449 000	287 360
DTX3	1620	518 400	810 000	518 400
DTX4	1188	380 160	594 000	380 160
DT21	938	300 160	469 000	300 160
DT22	616	197 280	308 000	197 120
DT23	938	300 320	469 000	300 160
DT24	1188	380 160	594 000	380 160

7.2.1 Test sequences for background noise estimation

Background noise estimation algorithm is tested by the following test sequences:

DTX1.*

DTX2.*

DT21.*

DT22.*

(The sequence DTX1.INP is the same as in ETS 300 725 (GSM 06.54) sequence DTX01.INP)

7.2.2 Test sequences for pitch, tone and complex signal detection

Pitch, tone and complex signal detection algorithms are tested by the following test sequence:

DTX3.*

7.2.3 Real speech and tones

This test sequence consists of very clean speech, barely detectable speech and a swept frequency tone (The sequences DTX4.INP and DT24.INP are the same as in ETS 300 725 (GSM 06.54) sequence DTX07.INP):

DTX4.*

DT24.*

7.2.4 Test sequence for signal-to-noise ratio estimation

The full range of SNR estimates are tested by the following test sequence:

DT23.*

8 Sequences for finding the 20 ms framing of the adaptive multi-rate speech encoder

[This section needs further adaptation of the text and terminology from GSM to 3G context.]

When testing the decoder, alignment of the test sequences used to the decoder framing is achieved by the air interface (testing of MS) or can be reached easily on the Abis-interface (testing on network side).

When testing the encoder, usually there is no information available about where the encoder starts its 20 ms segments of speech input to the encoder.

In the following, a procedure is described to find the 20 ms framing of the encoder using special synchronisation sequences. This procedure can be used for MS as well as for network side.

Synchronisation can be achieved in two steps. First, bit synchronisation has to be found. In a second step, frame synchronisation can be determined. This procedure takes advantage of the codec homing feature of the adaptive multi-rate codec, which puts the codec in a defined home state after the reception of the first homing frame. On the reception of further homing frames, the output of the codec is predefined and can be triggered to.

8.1 Bit synchronisation

The input to the speech encoder is a series of 13 bit long words (104 kbits/s, 13 bit linear PCM). When starting to test the speech encoder, no knowledge is available on bit synchronisation, i.e., where the encoder expects its least significant bits, and where it expects the most significant bits.

The encoder homing frame consists of 160 samples, all set to zero with the exception of the least significant bit, which is set to one (0 0000 0000 0001 binary, or 0x0008 hex if written into 16 bit words left justified). If two such encoder homing frames are input to the encoder consecutively, the corresponding decoder homing frame of the used codec mode is expected at the output as a reaction of the second encoder homing frame.

Since there are only 13 possibilities for bit synchronisation, after a maximum of 13 trials bit synchronisation can be reached for each codec mode. In each trial three consecutive encoder homing frames are input to the encoder. If the corresponding decoder homing frame is not detected at the output, the relative bit position of the three input frames is shifted by one and another trial is performed. As soon as the decoder homing frame of the used codec mode is detected at the output, bit synchronisation is found, and the first step can be terminated.

The reason why three consecutive encoder homing frames are needed is that frame synchronisation is not known at this stage. To be sure that the encoder reads two complete homing frames, three frames have to be input. Wherever the encoder has its 20 ms segmentation, it will always read at least two complete encoder homing frames.

An example of the 13 different frame triplets is given in sequence BITSYNC.INP.

8.2 Frame synchronisation

Once bit synchronisation is found, frame synchronisation can be found by inputting two identical frames consecutively to the encoder. There exist 160 different output sequences depending on the 160 different positions that the beginning of this sequence of frames can possibly have with respect to the encoder framing.

Before inputting this special synchronisation sequence to the encoder, again the encoder has to be reset by one encoder homing frame. A second encoder homing frame is needed to provoke a decoder homing frame at the output that can be triggered to. And since the framing of the encoder is not known at that stage, three encoder homing frames have to precede the special synchronisation sequence to ensure that the encoder reads at least two homing frames, and at least one decoder homing frame is produced at the output, serving as a trigger for recording.

After the last decoder homing frame of the used codec mode it is required to detect two consecutive output frames that are different from the preceding decoder homing frame. To achieve this in the 12.2 kbit/s mode (no lookahead in the linear prediction analysis [4]), the last 40 samples of the third encoder homing frame shall be different from 0x0008 hex. Only the first 120 samples of this frame were set to 0x0008 hex in this mode.

The special synchronisation sequence preceded by three encoder homing frames are given in SEQSYNC.INP. For the 12.2 kbit/s mode this sequence is different in the third frame and is given in SEQSYNC_122.INP.

Generally, the output sequences will be different depending on the tested adaptive multi-rate mode. In the notation below <mode> should be changed to the number of the tested mode, i.e. one of 122, 102, 795, 74, 67, 59, 515 or 475.

In all 160 output sequences only the second frame after the last decoder homing frame is given in SYNC000_<mode>.COD through SYNC159_<mode>.COD. These output frames were calculated by shifting the sequence SEQSYNC.INP respectively SEQSYNC_122.INP through the positions 0 to 159, where the samples at the beginning were set to zero. For each codec mode it was finally verified that the last frame in each of the 160 output sequences is different to all other last frames.

The three digit number in the filenames above indicates the number of samples by which the input was retarded with respect to the encoder framing. By a corresponding shift in the opposite direction, alignment with the encoder framing for the used codec mode can be reached.

8.3 Formats and sizes of the synchronisation sequences

BITSYNC.INP:

This sequence consists of 13 frame triplets. It has the format of the speech encoder input test sequences (13 bit left justified with the three least significant bits set to zero).

The size of it is therefore:

$$\text{SIZE (BITSYNC.INP)} = 13 * 3 * 160 * 2 \text{ bytes} = 12480 \text{ bytes}$$

SEQSYNC.INP/SEQSYNC_122.INP:

This sequence consist of a 3 frame header (see clause 8.2 for details) and the special synchronisation sequence, consisting of two identical frames. It has the format of the speech encoder input test sequences (13 bit left justified with the three least significant bits set to zero).

The size of it is therefore:

SIZE (SEQSYNC.INP/SEQSYNC_122.INP) = 5 * 160 * 2 bytes = 1600 bytes

SYNCXXX_<mode>.COD:

These sequences consists of 1 encoder output frame each. They have the format of the speech encoder output test sequences (16 bit words right justified). In these frames the values of the FRAME_TYPE and MODE_INFO fields are set to the transmit frame type and to the corresponding encoding mode information [3].

The size of them is therefore:

SIZE (SYNCXXX_<mode>.COD) = (244 + 6) * 2 bytes = 500 bytes

All files are contained in the archive S.TGZ which accompanies the present document.

9 Trau Testing with 8 Bit A- and μ -law PCM Test Sequences

[This section needs further adaptation of the text and terminology from GSM to 3G context.]

In the previous clauses, tests for the transcoder in the TRAU are described, using 13 bit linear test sequences. However, these 13 bit test sequences require a special interface in the TRAU and do not allow testing in the field. In most cases the TRAU has to be set in special mode before testing.

The 'Y' in the file names below stands for A (A-law) and U (μ -law), respectively.

As an alternative, the speech codec tests in the TRAU can be performed using A- or μ -law compressed 8 bit PCM test sequences on the A interface. For this purpose modified input test sequences (TXX_Y.INP) are generated from the original sequences (see clause 6) by A or μ law compression. As an input to the encoder they result in modified encoder output sequences (TXX.COD). These modified (TXX.COD) sequences are used as decoder input sequences. The decoder will then produce the output sequences TXX_Y.OUT, which are A- or μ compressed.

The A- and μ -law compression and decompression does not change the homing frames at the encoder input. The format of all A- and μ -law PCM files TXX_Y.INP and TXX_Y.OUT is one sample (8 bit) written into 16 bit words. The format of the modified TXX.COD files is as described in clause 5.

All files are contained in the archives T_A.TGZ (for the A-law sequences) and T_U.TGZ (for the μ -law sequences) which accompany the present document.

In addition to the test sequences above, special input (SEQSYNC_Y.INP/SEQSYNC_122_Y.INP) and output (SEQSYNC000_<mode>.COD through SEQSYNC159_<mode>.COD) sequences for frame synchronisation are provided. The Y again stands for A and μ law compressed PCM and <mode> is described in clause 6.

All files are contained in the archives S_A.TGZ (for the A-law sequences) and S_U.TGZ (for the μ -law sequences) which accompany the present document. The synchronization procedure is described in clause 8.

History

Document history		
V. 0.1.0	June 1999	First version directly based on GSM 06.74 (TSGS4#5(99)197)
V 1.0.0	June 1999	Presented for information to TSG SA#4 meeting.
V 1.1.0	October 1999	Test sequences for VAD option 2 included.
V 2.0.0	December 1999	Approval at TSG SA