ETSI TS 103 624 V1.1.1 (2019-11)

TECHNICAL SPECIFICATION

Characterization Methodology and Requirement Specifications for the ETSI LC3plus codec

Reference
DTS/STQ-279

Keywords
codec, listening quality, speech

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights ... 4
Foreword ... 4
Modal verbs terminology ... 4
1 Scope ... 5
2 References ... 5
2.1 Normative references ... 5
2.2 Informative references ... 6
3 Definition of terms, symbols and abbreviations ... 6
3.1 Terms ... 6
3.2 Symbols ... 6
3.3 Abbreviations ... 6
4 Introduction ... 7
5 Characterization methodologies ... 7
5.1 Overview ... 7
5.2 Experiments ... 8
5.3 Item processing ... 8
5.4 Subjective methodologies ... 8
5.5 Objective methodologies ... 9
6 Characterization test plan ... 9
6.1 Testing Conventions ... 9
6.1.1 Introduction ... 9
6.1.2 Software versions ... 9
6.1.3 Test condition numbering ... 9
6.2 Characterization test plan for clean channels with application in DECT scenarios ... 10
6.2.1 Overview ... 10
6.2.2 NB conditions ... 10
6.2.3 WB conditions ... 11
6.2.4 SWB conditions ... 12
6.2.5 FB conditions ... 13
6.3 Characterization plan for error prone channels with application in DECT scenarios ... 13
6.3.1 Overview ... 13
6.3.2 NB conditions ... 14
6.3.3 WB conditions ... 14
6.3.4 SWB conditions ... 14
6.4 Characterization test plan for clean channels with application in VoIP scenarios ... 14
6.4.1 Overview ... 14
6.4.2 NB conditions ... 15
6.4.3 WB conditions ... 15
6.4.4 SWB conditions ... 16
6.5 Characterization plan for Packet Loss Concealment (PLC) with application in VoIP scenarios ... 17
6.5.1 Overview ... 17
6.5.2 NB conditions ... 18
6.5.3 WB conditions ... 18
6.5.4 SWB conditions ... 19
7 Requirement verification ... 19
7.1 Requirement verification for subjective tests ... 19
7.2 Requirement verification for objective tests ... 19
Annex A (normative): Conditions for the P.800 experiments ... 20
Annex B (informative): Bibliography ... 21
History ... 22

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

Foreword

This Technical Specification (TS) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ).

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

1 Scope

The present document specifies the subjective and objective methodologies developed in cooperation between TC STQ and TC DECT for the characterization of the Low Complexity Communication Codec Plus (LC3plus). It describes experimental tests and conditions used for subjective and objective testing. Based on these methodologies the performance requirements for this codec are specified.

The requirements in the present document are specified to characterize a high-quality codec for use in modern telecommunication networks, including but not limited to DECT and VoIP. A special focus is placed on the fact that end-to-end connections are often of hybrid nature concatenating different technologies and thus tandeming (i.e. transcoding) different codecs.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at https://docbox.etsi.org/Reference/.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are necessary for the application of the present document.

[1] Recommendation ITU-T P.800 (08/1996): "Methods for subjective determination of transmission quality".

[2] Recommendation ITU-T P.863 (03/2018): "Perceptual objective listening quality prediction".

[3] Recommendation ITU-T G.722 (09/2012): "7 kHz audio-coding within 64 kbit/s".

[4] Recommendation ITU-T G.726 (12/1990): "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)".

[5] ETSI TS 103 634: "Digital Enhanced Cordless Telecommunications (DECT); Low Complexity Communication Codec plus (LC3plus)".

[6] Recommendation ITU-T G.191 (01/2019): "Software tools for speech and audio coding standardization".

[7] ETSI TS 126 442: "Universal Mobile Telecommunications System (UMTS); LTE; Codec for Enhanced Voice Services (EVS); ANSI C code (fixed-point) (3GPP TS 26.442)".

[8] ETSI TS 126 173: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; ANSI-C code for the Adaptive Multi-Rate - Wideband (AMR-WB) speech codec (3GPP TS 26.173)".

[9] ETSI TS 126 073: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; ANSI-C code for the Adaptive Multi Rate (AMR) speech codec (3GPP TS 26.073)".

[10] Recommendation G.711 Appendix I (09/1999): "A high quality low-complexity algorithm for packet loss concealment with G.711".

[11] IETF RFC 8251: "Update to the Opus Audio Codec".

[12] Recommendation ITU-T G.711 (11/1988): "Pulse code modulation (PCM) of voice frequencies".

[13] Recommendation ITU-T G.722 (11/2006) "7 kHz audio-coding within 64 kbit/s; Appendix IV "A low-complexity algorithm for packet loss concealment with G.722".

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ETSI TR 103 590: "Digital Enhanced Cordless Telecommunications (DECT); Study of Super Wideband Codec in DECT for narrowband, wideband and super-wideband audio communication including options of low delay audio connections (<= 10 ms framing)".

[i.2] IETF RFC 6716: "Definition of the Opus Audio Codec".

[i.3] 3GPP, S4-141392: "EVS-7c Processing functions for characterization phase", TSG S4#81.

[i.4] 3GPP, S4-141319: "EVS-8b EVS Permanent Document EVS-8b: Test plans for selection phase including lab task specification", TSG S4#81.

[i.5] 3GPP, S4-141372: "EVS-8c EVS Permanent Document EVS-8c: Test plans for characterization phase including lab task specification", TSG S4#81.

[i.6] IEEE: "A method for comparing the performance of EVS and other voice codecs under bursty packet loss", IPTcomm, 2018.

3 Definition of terms, symbols and abbreviations

3.1 Terms

Void.

3.2 Symbols

Void.

3.3 Abbreviations

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

ACR	Absolute Category Rating
AMR-NB	Adaptive Multirate speech codec Narrowband
AMR-WB	Adaptive Multirate speech codec Wideband
BER	Bit Error Rate
CBR	Constant Bitrate
CELT	Constrained Energy Lapped Transform
CuT	Codec under Test
DCR	Degradation Category Rating
DECT	Digital Enhanced Cordless Telecommunications
DP	DECT Profile
EVS	codec for Enhanced Voice Services
EVS-WB	EVS WideBand

FB	FullBand
FEC	Forward Error Correction
FER	Forward Error Correction RealTime Protocol
FP	Fixed Part
LC3plus	Low Complexity Communication Codec Plus
NB	NarrowBand
PLC	Packet Loss Concealment
PLP	Packet Loss Profile
PLR	Packet Loss Rate
PP	Portable Part
RF	Radio Frequency
RSSI	Received Signal Strength Indicator
RTP	RealTime Protocol
STL	Standard Template Library
SWB	Super WideBand
VoIP	Voice over IP
WB	WideBand

4 Introduction

The present document defines characterization methodologies as well as the performance requirements to be evaluated for the ETSI Low Complexity Communication Codec Plus (LC3plus) [5]. The performance of the codec was initially studied by the TC DECT group in ETSI TR 103 590 [i.1] which is considered as qualification of the codec.

The purpose of the characterization phase experiments is to demonstrate the performance of the codec over a set of conditions and the following use cases:

- Voice services in DECT and VoIP
- Interworking VoIP scenarios between different networks
- Music streaming as additional interesting use case

The characterization utilizes the set of characterization methodologies and configurations of subjective and objective experiments defined in clause 5. The experiments are designed in order to evaluate whether LC3Plus achieves the following codec objectives:

- Introduction of Super-Wideband (SWB) quality in voice services
- Increased capacity of DECT systems when compared to legacy DECT codecs
- Improved robustness for packet loss and bit errors
- Ensure suitable performance in case of transcoding or self-tandeming conditions

All details on the definition of codec objectives for DECT and VoIP and the derived performance requirements and performance objectives are specified in clause 6.

Clause 7 defines the statistical analysis to be conducted on the subjective results to verify that the performance of the Codec under Test (CuT) is sufficient in comparison to the specified performance requirement or performance objectives. In the present document, CuT always means ETSI LC3plus [5].

5 Characterization methodologies

5.1 Overview

The present clause describes the experiment design and the subjective and objective methodologies. The aim of the characterization test is to assess the clean channel performance, self-tandeming capabilities, cross-tandeming, as well as rate switching conditions and variation of the input speech level.

The characterization tests shall be conducted in the same way as the 3GPP EVS selection/characterization process [i.4] and [i.5].

5.2 Experiments

All test conditions shall be separated according to the category audio bandwidth and channel conditions. This results in six experiments, i.e. 3x audio bandwidth times 2x channel conditions. Additionally, one multi-bandwidth experiment shall be conducted in order to provide a quality overview.

Each experiment is evaluated using subjective and objective methodologies described in clauses 5.4 and 5.5.

Table 1 outlines the experiment setup:

Table 1: Experiment overview

Experiment number	Experiment label	Max. bandwidth of input	Channel conditions	Estimated number of conditions
1	NB clean	4 000 Hz	No error	40
2	NB error	4 000 Hz	Bit error & packet loss	32
3	WB clean	8 000 Hz	No error	60
4	WB error	8 000 Hz	Bit error & packet loss	32
5	SWB clean	16 000 Hz	No error	40
6	SWB error	16 000 Hz	Bit error & packet loss	35
7	M fullscale (see note)	20 000 Hz	No error & bit error & packet loss	44
NOTE: The M fullscale experiment contains all bandwidth conditions to span the complete P.800 quality range [1].				

A complete list of all experiments and conditions describing the exact configuration for each condition and the relevant comparison points are contained in archive ts_103624v010101p0.zip which accompanies the present document.

5.3 Item processing

The test items shall be processed according to the EVS processing plan [i.3]. For transcoding, no frame synchronization between the codecs shall be applied. The frequency masks used by 3GPP EVS characterization tests shall be applied to the input signals. The items shall be processed and prepared for the experiments using the STL 2009 [6] tools.

5.4 Subjective methodologies

All subjective experiments shall be conducted using the Recommendation ITU-T P.800 [1] procedure using speech material. Subjects shall be naïve listeners and native speakers. Experiments should be conducted in different languages and labs.

Table 2 shows the P.800 experiment configurations.

Table 2: P.800 experiment configuration

Parameter	Experiment						
	1	2	3	4	5	6	7
Rating scale	ACR	ACR	ACR	ACR	DCR	ACR	ACR
Min. number of listeners	24	24	24	24	24	24	24
Min. num. of talkers	4	4	4	4	4	4	4
Min. num. of samples per talker	6	6	6	6	6	6	6
Min. number of votes per sample	4	4	4	4	4	4	4
Min. number of votes per condition	96	96	96	96	96	96	96
Est. test duration in min. (see note)	47	41	63	41	68	38	45
NOTE: Estimation calculation contained in archive ts_103624v010101p0.zip.							

5.5 Objective methodologies

All experiments listed in Table 1 shall be assessed by the objective quality evaluation using the perceptual objective listening quality prediction tool standardized by ITU-T also known as Recommendation ITU-T P.863 [2].

Tests shall be run in the full band mode with full band reference files and appropriate degraded files.

6 Characterization test plan

6.1 Testing Conventions

6.1.1 Introduction

The following clauses specify performance requirements and conditions to be evaluated for the following use cases:

- DECT with clean channel conditions.
- DECT with error prone channel conditions.
- VoIP without packet loss conditions.
- VoIP including packet loss conditions.

Besides performance requirements, performance objectives are specified. The performance objectives are only foreseen as informative comparison conditions.

6.1.2 Software versions

The following software version for the different codecs shall be used:

- G.711 A-law: Recommendation ITU-T G.711 [12] and G.711 Appendix. I (PLC) [10].
- IETF RFC 8251 [11] OPUS: V1.1.3 (deployed) or V1.3.0 (latest), fix-point.

NOTE: OPUS is a codec in accordance with IETF RFC 6716 [i.2] and IETF RFC 8251 [11].

- EVS: EVS Codec ETSI TS 126 442 [7] V12.7.0 and V13.2.0 or latest one; EVS Codec ETSI TS 126 442 [7]. V12.12.0 and 13.7.0.
- LC3plus: Latest.
- G.722: Recommendation ITU-T G.722 [3] and G.722 Appendix IV [13] or Recommendation ITU-T G.722 [3] + Appendix IV.
- AMR-WB (G.722.2): ETSI TS 126 173 [8] V15.1.0 (latest).
- AMR-NB: ETSI TS 126 073 [9] V15.0.0 (latest).
- G.726: Recommendation ITU-T G.726 [4].

6.1.3 Test condition numbering

The test conditions are numbered according the scheme given in Table 3.

Table 3: Test condition numbering

NB	WB	SWB
DECT with clean channel conditions		
1xx	2xx	3xx
DECT with error prone channel conditions		
4xx	5xx	6xx
VoIP without packet loss conditions		
7xx	8xx	9xx
VoIP including packet loss conditions		
10xx	11xx	12xx

6.2 Characterization test plan for clean channels with application in DECT scenarios

6.2.1 Overview

CuT in DECT shall provide the same or better voice quality than the VoIP network provides and guarantees higher efficiency than DECT audio codecs used today, meaning same quality at lower bit rates to allow better DECT slot exploitation in conjunction with channel coding to provide better protection for bit errors and packet loss concealment.

As network interworking scenarios, the following cases shall be evaluated:

- Voice calls from legacy VoIP to DECT
- Voice calls from DECT to legacy VoIP
- Voice calls from DECT over legacy VoIP to DECT

DECT uses today G.726 (NB) and G.722 (WB). Today's VoIP terminals utilize G.711 (NB) and G.722 (WB).

6.2.2 NB conditions

The test shall verify the performance of the CuT in NB mode. Speech coding for narrowband speech connections using a normal 32 kbit/s payload DECT RF slot shall not be worse than what is achieved by Recommendation ITU G.726 [4]. The CuT shall enable the same range where communication is possible between DECT PP and FP as achieved at the date of publication of the present document by DECT-G.726 connections.

The voice quality by transcoding between VoIP G.711 to/from CuT shall not be worse than connections between VoIP-G711 and DECT-G.726.

Additional performance objectives should be defined in comparison to OPUS (CELT mode, constant bitrate mode (CBR), 32 kbit/s, complexity=0, FEC off, NB mode, 10 ms framing).

The following NB conditions shall be included into the test (Input speech levels to be applied are -16 dBov, -26 dBov, -36 dBov):

100. Direct reference conditions with limited audio bandwidth (cut off frequency of 4 kHz) but no speech coding.

CuT:

101. LC3plus 32 kbit/s, 10 ms framing.

Requirement:

102. G.726, 32kbit/s with G.711 Appendix I PLC.

Performance objective:

103. OPUS, CELT mode, CBR, 32 kbit/s, complexity = 0, FEC off, NB mode, 10 ms framing.

The following transcoding scenarios shall be tested:

CuT:

104. G.711->LC3plus (32 kbit/s).
105. LC3plus (32 kbit/s)-> G.711.
106. LC3plus (32 kbit/s)-> G.711-> LC3plus (32 kbit/s).

Requirement:

107. G.711->G.726 (32 kbit/s).
108. G.726 (32 kbit/s)-> G.711.
109. G.726 (32 kbit/s)-> G.711->G.726 (32 kbit/s).

Performance objective:

110. G.711 -> OPUS (32 kbit/s).
111. OPUS (32 kbit/s) -> G.711.
112. OPUS (32 kbit/s) -> G.711 -> OPUS (32 kbit/s).

The following codecs shall be tested for self-tandeming (double and triple):

113. LC3plus (32 kbit/s).
114. G.726 (32 kbit/s).
115. OPUS (32 kbit/s).
116. G.711 (64 kbit/s).

6.2.3 WB conditions

The test shall verify the performance of the candidate codec in WB mode for DECT scenarios. Speech coding for wideband speech connections using a 32 kbit/s payload for normal DECT RF slots shall not be worse than what is achieved today by Recommendation ITU G.722 [3] using a 64 kbit/s payload for long DECT RF slots. The DECT evolution RF connection shall enable at least the same range where communication is possible between DECT PP and FP compared to today's G.722 DECT connections. It is envisioned that the range can be further extended.

The voice quality by transcoding between VoIP networks using G.722 to/from the DECT evolution speech codec shall not be worse than connections between VoIP-G.722 and DECT-G.722.

Additional performance objectives should be defined in comparison to OPUS (CELT mode, CBR, 32 kbit/s, complexity=0, FEC off, WB mode, 10 ms framing).

The following WB conditions shall be included into the test (input speech levels which shall be used are -16 dBov, -26 dBov and -36 dBov):

200. Direct reference condition with limited audio bandwidth with cut off frequency of 8 kHz, but no speech coding.

CuT:

201. LC3plus, 32 kbit/s, 16 kHz, 10 ms framing, 16 bits per audio sample.

Requirement:

202. G.722, 64 kbit/s.

Performance objective:

203. OPUS, CELT mode, constant bitrate (CBR): 32 kbit/s, complexity=0, FEC off, WB mode, 10 ms framing.

To be characterized:

204. LC3plus for bitrates: 32 kbit/s, 48 kbit/s. Sampling rate of 16 kHz and nominal speech level. Short frame size (5 ms frame size) against regular frame size LC3plus 32 kbit/s codec (10 ms frame size).
205. LC3plus for bitrates: 64 kbit/s, 96 kbit/s. Sampling rate of 16 kHz and nominal speech level. Short frame size (2,5 ms frame size) against regular frame size LC3plus 32 kbit/s codec (10 ms frame size).

The following transcoding scenarios shall be tested:

CuT:

206. LC3plus (32 kbit/s)-> G.722 (64 kbit/s).
207. G.722 (64 kbit/s)-> LC3plus (32 kbit/s).
208. LC3plus (32 kbit/s)->G.722 (64 kbit/s)-> LC3plus (32 kbit/s).

Requirement:

209. G.722 (64 kbit/s)-> G.722 (64 kbit/s).
210. G.722 (64 kbit/s)-> G.722 (64 kbit/s) -> G.722 (64 kbit/s).

Performance objective:

211. OPUS (32 kbit/s) -> G.722 (64 kbit/s).
212. G.722 (64 kbit/s) -> OPUS (32 kbit/s).
213. OPUS (32 kbit/s) -> G.722 (64 kbit/s) -> OPUS (32 kbit/s).

The following codecs shall be tested for self-tandeming (double, triple):

214. LC3plus (32 kbit/s).
215. OPUS (32 kbit/s).
216. G.722 (64 kbit/s).

6.2.4 SWB conditions

Speech coding for super-wideband speech connections using a long 64 kbit/s payload DECT RF slot shall not be worse than what is achieved by EVS-SWB at 13,2 kbit/s and better than what is achieved by ITU G.722 at 64 kbit/s. The DECT evolution RF connection shall enable the same range where communication is possible between DECT PP and FP as achieved today by G.722 DECT connections.

The voice quality degradation by transcoding between VoIP networks using OPUS (fullband mode) or EVS to/from DECT evolution speech codec shall be characterized.

Additional objectives should be defined in comparison to OPUS (CELT mode, CBR, 64 kbit/s, complexity=0, FEC off, FB mode, 10 ms framing).

The following conditions shall be tested (Input speech levels to be applied are -16 dBov, -26 dBov, -36 dBov):

300. Direct reference conditions with limited audio bandwidth but no speech coding. Lowpass cutoff frequency of 16 kHz shall be used.

CuT:

301. LC3plus, 64 kbit/s at sampling rate of 32 kHz. 10 ms framing, 16 bits per audio sample.

Requirement:

302. G.722 (64 kbit/s, WB, with Appendix IV PLC).
303. EVS (SWB at 13.2 kbit/s). No channel aware mode used.

Performance Objective:

304. OPUS, CELT mode, CBR, 64 kbit/s, complexity=0, FEC off, fullband, mode, 10 ms framing.

To be characterized:

305. LC3plus for bitrates: 64 kbit/s, 96 kbit/s. Sampling rate of 32 kHz and nominal speech level. Short frame size (5 ms frame size) against regular frame size LC3plus 64 kbit/s codec (10 ms frame size).
306. LC3plus for bitrates: 96 kbit/s, 128 kbit/s. Sampling rate of 16 kHz and nominal speech level. Short frame size (2,5 ms frame size) against regular frame size LC3plus 64 kbit/s codec (10 ms frame size).

The following transcoding scenarios shall be tested:

307. EVS (SWB at 13,2 kbit/s) -> LC3plus (64 kbit/s).
308. LC3plus (64 kbit/s) -> EVS (SWB at 13,2 kbit/s).
309. LC3plus (64 kbit/s) -> EVS (SWB at 13,2 kbit/s) -> LC3plus (64 kbit/s).

310. EVS (SWB at 13,2 kbit/s) -> OPUS (64 kbit/s).
311. OPUS (64 kbit/s) -> EVS (SWB at 13,2 kbit/s).
312. OPUS (64 kbit/s) -> EVS (SWB at 13,2 kbit/s) -> OPUS (64 kbit/s).

313. LC3plus (64 kbit/s) -> OPUS (64 kbit/s).
314. OPUS (64 kbit/s) -> LC3plus (64 kbit/s).
315. LC3plus (64 kbit/s) -> OPUS (64 kbit/s) -> LC3plus (64 kbit/s).

The following codec shall be tested for self-tandeming (double, triple and quadruple):

316. LC3plus (64 kbit/s).
317. Opus (64 kbit/s).
318. EVS (SWB at 13,2 kbit/s).

6.2.5 FB conditions

DECT evolution shall have the option to allocate a service for high quality FB music streaming.

Music audio quality of CuT coded originals and the impact by transcoding between music coder (e.g. MP3) and CuT shall be characterized.

Performance objective:

OPUS (CELT mode, CBR, 115 kbit/s, complexity=0, FEC off, FB mode, 10 ms framing, stereo)

DECT foresees the following operation points for stereo music transmission:

- 7a (64 kbps per channel, i.e. 128 kbps total)
- 7b (96 kbps per channel, i.e. 192 kbps total)
- 7c (128 kbps per channel, i.e. 256 kbps total)

However, the characterization of the CuT for FB for music conditions should not be part of the same test as the NB, WB and SWB conditions, since the rest of the conditions are meant to be used for the assessment of the speech quality.

6.3 Characterization plan for error prone channels with application in DECT scenarios

6.3.1 Overview

The packet loss concealment performance of CuT shall be evaluated compared to G.726 in NB and G.722 in WB and to SWB codec.

Table 3: DECT Packet loss and bit error rates

Packet Loss Profile	Normalized Averaged signal strength (RSSI)	PLR [%]	BER rounded [%]
DP0	1 (136 dB)	0	0
DP1	0,41 (56 dB)	0,99	0,01
DP2	0,35 (48 dB)	0,88	0,31
DP3	0,29 (40 dB)	7,39	2,92

The CuT shall be compared to the requirement condition under four typical signal strengths representing DECT packet loss and bit error profiles for a 10 ms framing. The configurations of the codecs include a specific setup to adapt to the DECT channel characteristic, e.g. configuration of the channel coder or addition of parity bit. The DECT error profiles are labelled DP0, DP1, DP2 and DP3 (see Table 3).

Switching of channel coder configurations (including rate switching) shall be tested as well. Switching within all possible channel configurations shall not be worse than only operating in the mode with highest protection.

For completeness, random patterns of 1,43 %, 3 % and 6 % for a 10 ms frame size shall be tested. The random patterns are label as FER 1,4 %, 3 % and 6 %.

6.3.2 NB conditions

The error profiles DP0, DP1, DP2, DP3 (see Table 4) and random FER 1,4 %, 3 % and 6 % shall be applied for:

CuT:

400. LC3plus, 32 kbit/s, sampling rate 8 kHz, 10 ms framing.

Requirement:

401. G.726 32 kbit/s with G.711 Appendix I PLC.

6.3.3 WB conditions

The error profiles DP0, DP1, DP2, DP3 (see Table 4) and random FER 1,4 %, 3 % and 6 % shall be applied to:

CuT:

500. LC3plus, 32 kbit/s, sampling rate 16 kHz, 10 ms framing.

Requirement:

501. G.722 64 kbit/s with G.722 Appendix IV PLC.

6.3.4 SWB conditions

The error profiles DP0, DP1, DP2, DP3 (see Table 4) and random FER 1,4 %, 3 % and 6 % shall be applied for:

CuT:

600. LC3plus, bitrate 64 kbit/s, sampling rate 32 kHz, 10 ms framing.

Requirement:

601. G.722 64 kbit/s with G.722 Appendix IV PLC (G.722 conditions required to check that SWB connections achieve the same DECT distance of portable and fix part compared to legacy WB connections).

6.4 Characterization test plan for clean channels with application in VoIP scenarios

6.4.1 Overview

VoIP networks today use mainly G.711 for NB, G.722 for WB and OPUS for SWB. As SWB services are currently deployed in VoLTE, EVS-SWB may serve as the alternative reference point. However, OPUS is used in the following proposal.

A new ETSI VOIP codec shall provide the same or better speech quality than today's VoIP network provides. It is envisioned that the new codec provides a Packet Loss Concealment (PLC) better than the currently provided PLC for G.711 and G.722 for narrowband and wideband calls.

For VoIP, the network interworking scenario with mobile phones shall be of main focus, leading to the transcoding conditions: VoIP to mobile terminals and vice versa.

As relevant mobile codecs AMR-NB, AMR-WB, EVS-WB and EVS-SWB shall be considered operating at the most commonly used configurations as outlined in the following clauses 6.4.2 to 6.4.4.

6.4.2 NB conditions

For narrowband VoIP network speech coding connections using up to 64 kbit/s payload, the CuT shall not be worse than G.711 coding.

The following conditions shall be included into the test (Input speech levels to be applied are -16 dBov, -26 dBov, -36 dBov):

CuT:

700. LC3plus, bitrate of 32 kbit/s, sampling rate 8 kHz, 10 ms framing, 16 bits per audio sample.

Requirement:

701. G.711 64 kbit/s, with Appendix I PLC.

Performance objective:

702. OPUS, CELT mode, CBR, 32 kbit/s, complexity=0, FEC off, NB mode, 10 ms framing.
703. AMR-NB 12,2 kbit/s.

The following transcoding scenarios shall be tested:

CuT:

704. AMR-NB (12,2 kbit/s) -> LC3plus (32 kbit/s).
705. LC3plus (32 kbit/s) -> AMR-NB (12,2 kbit/s).

Requirement:

706. AMR-NB (12,2 kbit/s) -> G.711.
707. G.711 -> AMR-NB (12,2 kbit/s).

Performance objective:

708. AMR-NB (12,2 kbit/s) -> OPUS (32 kbit/s).
709. OPUS (32 kbit/s) -> AMR-NB (12,2 kbit/s).

6.4.3 WB conditions

VoIP network speech coding done by CuT for wideband speech connections using up to 64 kbit/s payload shall not be worse than G.722 coding.

The following WB conditions shall be included into the test (Input speech levels to be applied are -16 dBov, -26 dBov, -36 dBov):

CuT:

800. LC3plus, 32 kbit/s, sampling rate 16 kHz, 10 ms framing, 16 bits per audio sample.

Requirement:

801. G.722, 64 kbit/s with Appendix IV PLC will be used.

Performance objective:

802. OPUS, CELT mode, CBR, 32 kbit/s, complexity=0, FEC off, WB mode, 10 ms framing.

The following transcoding scenarios shall be tested:

CuT:

803. AMR-WB (23,85 kbit/s) -> LC3plus (32 kbit/s).
804. LC3plus (32 kbit/s) -> AMR-WB (23,85 kbit/s).
805. AMR-WB (12,65 kbit/s) -> LC3plus (32 kbit/s).
806. LC3plus (32 kbit/s) -> AMR-WB (12,65 kbit/s).
807. EVS-WB (24,4 kbit/s) -> LC3plus (32 kbit/s).
808. LC3plus (32 kbit/s) -> EVS-WB (24,4 kbit/s).
809. EVS-WB (13,2 kbit/s) -> LC3plus (32 kbit/s).
810. LC3plus (32 kbit/s) -> EVS-WB (13,2 kbit/s).

Requirement:

811. AMR-WB (23,85 kbit/s) -> G.722 (64 kbit/s).
812. G.722 (64 kbit/s) -> AMR-WB (23,85 kbit/s).
813. AMR-WB (12,65 kbit/s) -> G.722 (64 kbit/s).
814. G.722 (64 kbit/s) -> AMR-WB (12,65 kbit/s).
815. EVS-WB (24,4 kbit/s) -> G.722 (64 kbit/s).
816. G.722 (64 kbit/s) -> EVS-WB (24,4 kbit/s).
817. EVS-WB (13,2 kbit/s) -> G.722 (64 kbit/s).
818. G.722 (64 kbit/s) -> EVS-WB (13,2 kbit/s).

Performance objective:

819. AMR-WB (23,85 kbit/s) -> OPUS (32 kbit/s).
820. OPUS (32 kbit/s) -> AMR-WB (23,85 kbit/s).
821. AMR-WB (12,65 kbit/s) -> OPUS (32 kbit/s).
822. OPUS (32 kbit/s) -> AMR-WB (12,65 kbit/s).
823. EVS-WB (24,4 kbit/s) -> OPUS (32 kbit/s).
824. OPUS (32 kbit/s) -> EVS-WB (24,4 kbit/s).
825. EVS-WB (13,2 kbit/s) -> OPUS (32 kbit/s).
826. OPUS (32 kbit/s) -> EVS-WB (13,2 kbit/s).

The following codecs shall be tested for self-tandeming (double, triple):

827. LC3plus (32 kbit/s).
828. OPUS (32 kbit/s).
829. EVS-WB (13,2 kbit/s).

6.4.4 SWB conditions

VoIP network speech coding for super wideband speech connection typically uses a payload of 64 kbit/s payload. The CuT shall not be worse than OPUS ((CELT mode, constant bitrate mode (CBR), 64 kbit/s, complexity=0, FEC off, fullband mode, 10 ms framing) for coding.

The following SWB conditions shall be included into the test (Input speech levels to be applied are -16 dBov, -26 dBov, -36 dBov):

CuT:

900. LC3plus, bitrate 64 kbit/s, sampling rate 32 kHz, 10 ms framing, 16 bits per audio sample.

Requirement:

901. OPUS, CELT mode, CBR, 64 kbit/s, complexity=0, FEC off, fullband mode, 10 ms framing.

The following transcoding scenarios shall be tested:

902. EVS-SWB (24,4 kbit/s) -> LC3plus (64 kbit/s).
903. LC3plus (64 kbit/s) -> EVS-SWB (24,4 kbit/s).
904. EVS-SWB (13,2 kbit/s) -> LC3plus (64 kbit/s).
905. LC3plus (64 kbit/s) -> EVS-SWB (13,2 kbit/s).

906. EVS-SWB (24,4 kbit/s) -> OPUS (64 kbit/s).
907. OPUS (64 kbit/s) -> EVS-SWB (24,4 kbit/s).
908. EVS-SWB (13,2 kbit/s) -> OPUS (64 kbit/s).
909. OPUS (64 kbit/s) -> EVS-SWB (13,2 kbit/s).

The following codecs shall be tested for self-tandeming (double, triple):

910. LC3plus (64 kbit/s).
911. OPUS (64 kbit/s).
912. EVS-WB (13,2 kbit/s).

6.5 Characterization plan for Packet Loss Concealment (PLC) with application in VoIP scenarios

6.5.1 Overview

The following Packet Loss Profiles (PLPs) shall be applied.

Table 4: Packet Loss Profiles for VoIP application testing

N	Packet Loss Profile	Burst Length (No. Packets)	Average Loss Rate (%)	Comments
1	PLP0	0	0	No loss
2	PLP1	1	1,43	Uniform loss
3	PLP2	2	2,87	Synchronized with PLP1
4	PLP4	4	5,74	Synchronized with PLP1
5	PLP7	7	10,05	Synchronized with PLP1

The PLPs have been designed to test the impact of different bursts of losses on the quality of voice when different codecs and packet loss mitigation and resilience mechanisms are used (such as Packet Loss Concealment and Forward Error Correction). As shown in the table above, each PLP has a distinct length of loss burst.

PLP1 was the first profile to be developed and it was designed by distributing individual losses (i.e. burst length of 1) randomly along the whole PLP. It was used to design the rest of the PLPs by increasing the size of the burst. For example, if an extract of packet statuses from PLP1 was 111110111111 (where 1 indicates a successfully received packet and 0 indicates a lost packet) PLP4 is then developed by switching the status of the 3 packets succeeding the lost one from 1 to 0 (i.e. 111110111111 becomes 111110000111). As a result, all PLP types have bursts that begin at the same location in the speech sample (see Figure 1 below for an example applied to a 12 second speech sample consisting of 3 sentences).

NOTE: This only represents one approach to testing conditions with bursty packet loss and the applied loss profiles can be changed (e.g. to replicate the characteristics of particular networks and network conditions).

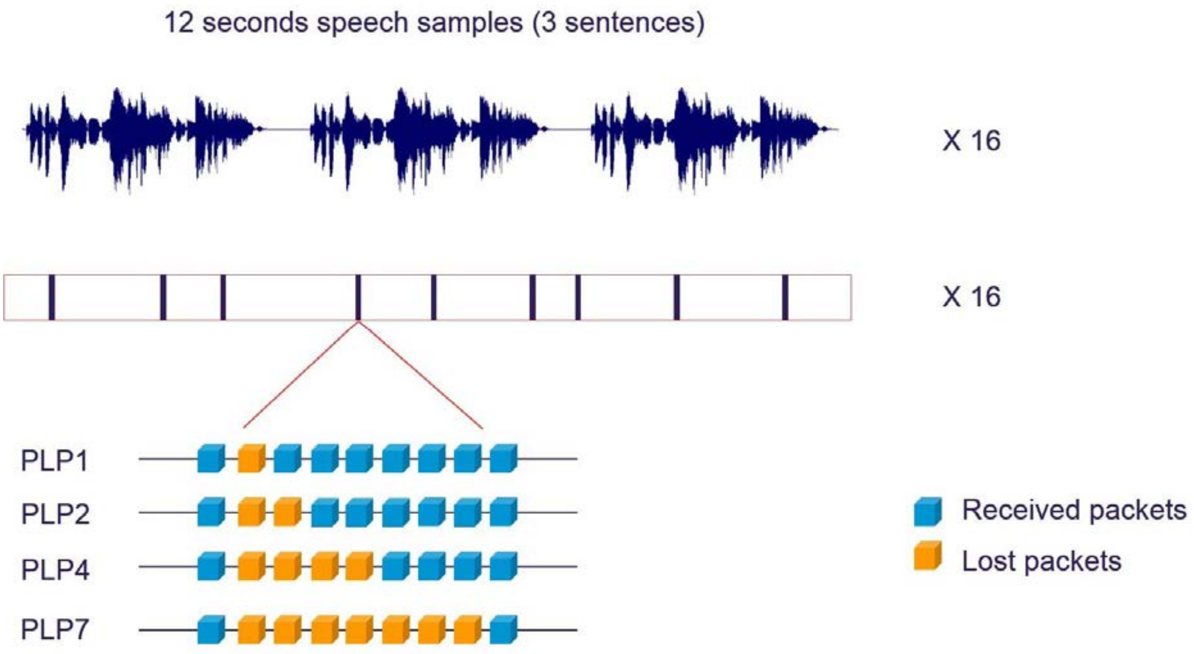

Figure 1

PLC frames should be aligned to the same frames for all speech input frames; For VoIP, only 20 ms PLC frames will be triggered. This means for 10 ms frame codec always 2 frames of PLC indication are in a row.

For completeness, random pattern of 1,43 %, 3 % and 6 % for a 20 ms frame size shall be tested. The random patterns are labeled as FER 1,4 %, 3 % and 6 %.

As LC3plus may not use the full bit rate of 64 kbps of the VoIP transmission, RTP based redundancy modes should be tested as well. Here, besides the main LC3plus frame of the current frame, an additional redundant LC3plus frame with an offset of X packets is transmitted in the same RTP payload. In this scenario, the playout is delayed by X packet, but the redundant LC3plus frame can be used to handle packet losses. The assessment described in the IEEE paper [i.6] mentioned above uses an offset of 3 packets for the EVS codec. The same value should be used for LC3plus. The processed condition may be based on simulations on bit stream level.

It is envisioned that the CuT provides a Packet Loss Concealment (PLC) better than the currently provided PLC for G.711, G.722 and OPUS for NB, WB and SWB calls respectively.

6.5.2 NB conditions

The CuT PLC shall be as good as, or better than G.711 appendix I for random packet losses at 20 ms packet size.

The Packet Loss Profiles PLP0, PLP1, PLP2, PLP4 and PLP7 (see Table 4) and FER 1,4 %, 3 % and 6 % shall be applied for:

CuT:

1000. LC3plus, bitrate of 32 kbit/s, sampling rate 8 kHz, 10 ms framing, 16 bits per audio sample.
1001. LC3plus, 64 kbit/s with RTP redundancy mode, sampling rate 8 kHz, 10 ms framing, 16 bits per audio sample (only relevant for PLP4 and PLP7; due to a redundancy configuration with offset=3, single and double losses in PLP1 and PLP2 can be completely compensated).

Requirement:

1002. G.711 64 kbit/s, with Appendix I PLC (Packet Loss Concealment).

6.5.3 WB conditions

The CuT PLC shall be as good as or better than G.722 Appendix IV for random packet losses at 20 ms packet size.

The Packet Loss Profiles PLP0, PLP1, PLP2, PLP4 and PLP7 (see Table 4) and FER 1,4 %, 3 % and 6 % shall be applied for:

CuT:

1100. LC3plus, bitrate 32 kbit/s, sampling rate 16 kHz, 10 ms framing, 16 bits per audio sample.
1101. LC3plus, bitrate 64 kbit/s with RTP redundancy mode, sampling rate 16 kHz, 10 ms framing, 16 bits per audio sample (only relevant for PLP4 and PLP7, due to a redundancy configuration with offset=3, single and double losses in PLP1 and PLP2 can be completely compensated).

Requirement:

1102. G.722 64 kbit/s with Appendix IV PLC (Packet Loss Concealment).

6.5.4 SWB conditions

The Packet Loss Profiles PLP0, PLP1, PLP2, PLP4 and PLP7 (see Table 4) and FER 1,4 %, 3 % and 6 % shall be applied for:

CuT:

1200. LC3plus, bitrate 64 kbit/s, sampling rate 32 kHz, 10 ms framing, 16 bits per audio sample.
1201. LC3plus, bitrate 64 kbit/s with RTP redundancy mode, sampling rate 32 kHz, 10 ms framing, 16 bits per audio sample (relevant for all PLPs).

Requirement:

1202. OPUS, CELT mode, CBR, 64 kbit/s, complexity=0, FEC off, FB mode, 10 ms framing.

7 Requirement verification

7.1 Requirement verification for subjective tests

Each condition in clauses 6.1 to 6.4 labelled as requirement shall be compared to the corresponding CuT condition using a Student's Dependent Groups t-test (single-sided at 95 % confidence level) on the subjective scores. This data will be the base for verifying that the CuT meets or exceeds the requirement. Additionally, the same rule shall be applied to the performance objectives, whereas the comparison between CuT and performance objective have only informative character.

A complete report for each requirement and objective will be provided in a future revision of the present document.

7.2 Requirement verification for objective tests

Each condition in clauses 6.2 to 6.5 labelled as requirement shall be compared to the corresponding CuT condition using a Student's Dependent Groups t-test (single-sided at 95 % confidence level) on the objective scores generated as described in clause 5.3. Additionally, the same rule shall be applied to the performance objectives.

The objective requirement verification is only for information. A complete report for each requirement and objective in comparison to the subjective results will be provided in a future revision of the present document.

Annex A (normative): Conditions for the P.800 experiments

The conditions for the seven P.800 experiments are contained in archive ts_103624v010101p0.zip which accompanies the present document.

Annex B (informative): Bibliography

- ETSI TS 126 071: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Mandatory speech CODEC speech processing functions; AMR speech Codec; General description (3GPP TS 26.071)".
- ETSI TS 126 441: "Universal Mobile Telecommunications System (UMTS); LTE; Codec for Enhanced Voice Services (EVS); General overview (3GPP TS 26.441)".

History

Document history		
V1.1.1	November 2019	Publication