ETSI TR 126 989 V15.0.0 (2018-07)

TECHNICAL REPORT

LTE;
Mission Critical Push To Talk (MCPTT);
Media, codecs and Multimedia Broadcast/Multicast Service (MBMS) enhancements for MCPTT over LTE
(3GPP TR 26.989 version 15.0.0 Release 15)

Reference

RTR/TSGS-0426989vf00

Keywords

LTE

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

Intellectual Property Rights

Foreword

This Technical Report (TR) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under http://webapp.etsi.org/key/queryform.asp.

Modal verbs terminology

In the present document "**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.

Contents

Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

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

Version x.y.z

where:

- x the first digit:
  - 1 presented to TSG for information;
  - 2 presented to TSG for approval;
  - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

1 Scope

The present document covers the enhancement required to support MCPTT.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TS 22.179: "Mission Critical Push To Talk (MCPTT) over LTE; Stage 1"

[3] 3GPP TR 26.952: "Codec for Enhanced Voice Services (EVS);Performance Characterization".

[4] 3GPP TS 26.114: "IP Multimedia Subsystem (IMS); Multimedia Telephony; Media handling and interaction".

[5] ITU-T Technical Paper - GSTP-GVBR, Performance of ITU-T G.718 (http://www.itu.int/pub/T-TUT) (http://www.itu.int/pub/publications.aspx-lang=en&parent=T-TUT-ASC-2010).

[6] ETSI EN 300 395-2: "Terrestrial Trunked Radio (TETRA) Speech codec for full-rate traffic channel Part 2: TETRA codec", version 1.3.1 (25 January 2005).

[7] 3GPP TR 26 975: "Performance characterization of the Adaptive Multi-Rate (AMR) speech codec".

[8] 3GPP TR 46.055: "Performance characterization of the GSM Enhanced Full Rate (EFR) speech codec".

[9] (void)

[10] IETF RFC 3550: "RTP: A Transport Protocol for Real-Time Applications".

[11] 3GPP TS 26.346: "Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs".

[12] 3GPP TS 23.468: "Group Communication System Enablers for LTE (GCSE_LTE); Stage 2".

[13] 3GPP TR 26 976: "Performance characterization of the Adaptive Multi-Rate Wideband (AMR-WB) speech codec".

[14] 3GPP TS 22.076: "Noise suppression for the AMR codec; Service description; Stage 1".

[15] 3GPP TS 26.131: "Terminal acoustic characteristics for telephony; Requirements".

[16] NTIA Report 15-520: "Speech Codec Intelligibility Testing in Support of Mission-Critical Voice Applications for LTE", S.D. Voran & A.A. Catellier September 2015.

[17] (void)

[18] (void)

[19] 3GPP TS 36.300: "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2".

[20] 3GPP TR 26.947: "Multimedia Broadcast/Multicast Service (MBMS); Selection and characterisation of application layer Forward Error Correction (FEC)".

[21] (void)

[22] (void)

[23] (void)

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

[25] 3GPP TS 26.442: "Codec for Enhanced Voice Services (EVS); ANSI C code (fixed-point)".

[26] 3GPP TS 26.448: "Codec for Enhanced Voice Services (EVS); Jitter buffer management".

[27] ITU-T Recommendation P.807 (02/2016): "Subjective test methodology for assessing speech intelligibility".

3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

ADP	Associated Delivery Procedures
AS	Application Server
BC	Broadcast
BM-SC	Broadcast-Multicast - Service Centre
GCS	Group Communication Service
MCPTT	Mission Critical Push-To-Talk
MBMS	Multimedia Broadcast/Multicast Service
MBSFN	Multimedia Broadcast Single Frequency Network
MOS	Mean Opinion Score
NTIA	National Telecommunications & Information Administration
TETRA	TErrestrial Trunked Radio
SC-PTM	Single Cell-Point To Multipoint
SWB	Super Wide Band
UC	Unicast

4 Reference Model

Figure 1 shows a reference model of MCPTT support over UC and BC. The GCS AS interacts with UE over GC1 interface for application signalling. The GCS AS determines whether to deliver the audio over UC or BC. GCS AS interacts with BM-SC over MB2 interface to deliver audio to BM-SC. The BM-SC delivers the audio over broadcast channel to the UE via SGi-mb interface. The GCS AS interacts with P-GW over SGi interface to deliver audio to the UE. The **red** line represents the audio delivered over UC channel. The **green** line represents the audio delivered over BC channel.

NOTE: The UE interacts with the BM-SC using HTTP method via SGi interface for MBMS Associated Delivery Procedure. Whether the ADP procedure applies to the MCPTT is TBD.

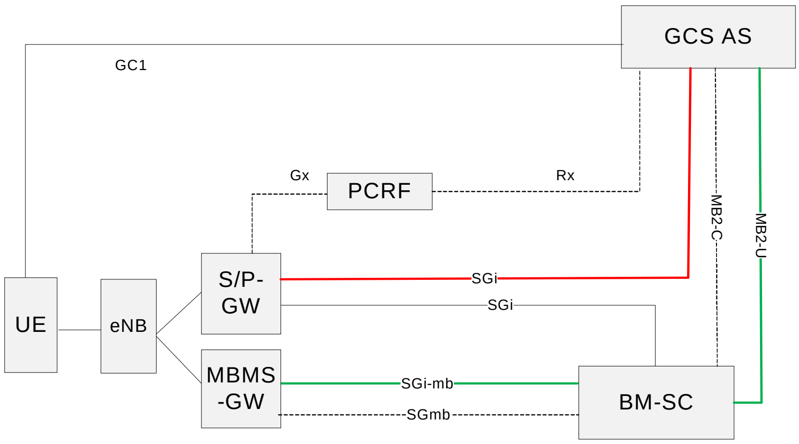


Figure 1: MCPTT support Reference Model

5 Key Issues for Supporting MCPTT

5.1 Key Issue#1: Codec for MCPTT

5.1.1 Review of Codec Alternatives and their Relative Perceptual Performance

5.1.1.1 Overview of the 3GPP Codec Comparison

The EVS Selection and Characterization Phase Test Results provided in the main body and Annex D of TR 26.952 [3] give a detailed assessment of the performance of the EVS Codec in realistic scenarios compared to both AMR and AMR-WB. A summary of this comparison is provided in the next two subclauses.

In the fourth subclause the relative performance of different audio bandwidths coded with AMR, AMR-WB and EVS is provided showing that the SWB modes of EVS outperform the WB and NB Primary modes of EVS, AMR-WB and AMR.

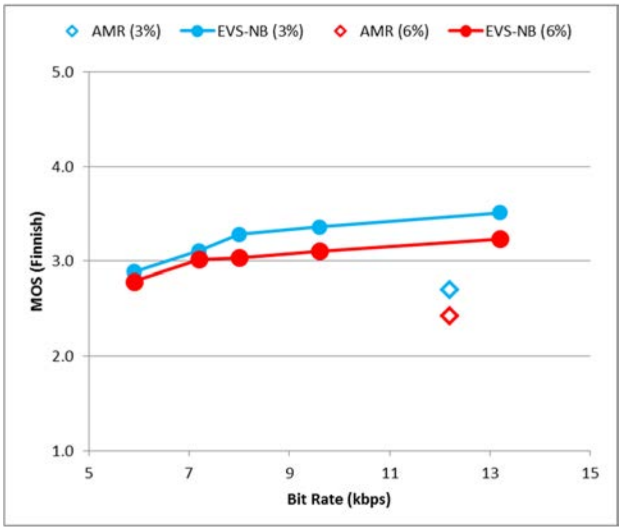
In the fifth subclause, a review of the TETRA codec performance in comparison to the 3GPP Codecs is provided.

This version of the document includes a review of codec alternatives and their relative intelligibility in high noise conditions, e.g., at SNRs in the range of -30 dB to 5 dB. The NTIA report [16] covered six noise types for an intelligibility study that included a range of public safety and civilian environments. Results of intelligibility testing for additional public safety specific high noise background conditions are not included in this document.

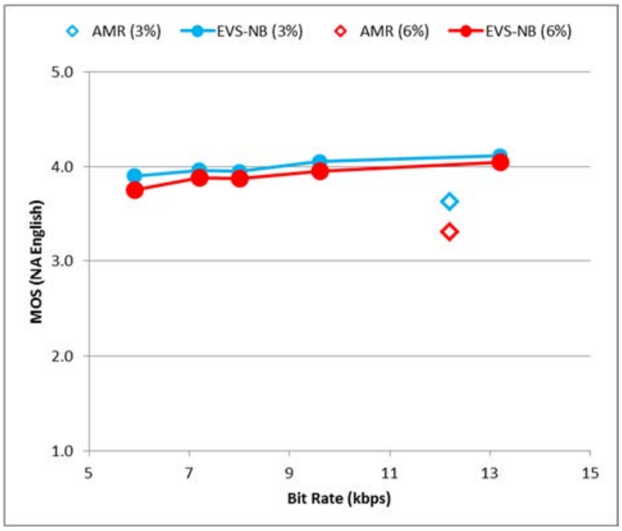
5.1.1.2 Narrowband Comparison vs AMR

For Narrowband (NB) signals, four experiments were conducted in the EVS Selection and four in the EVS Characterization. Taken together, these results provide a complete picture of the performance of EVS with respect to AMR but the highlights are provided in Figures 2 to 6 below.

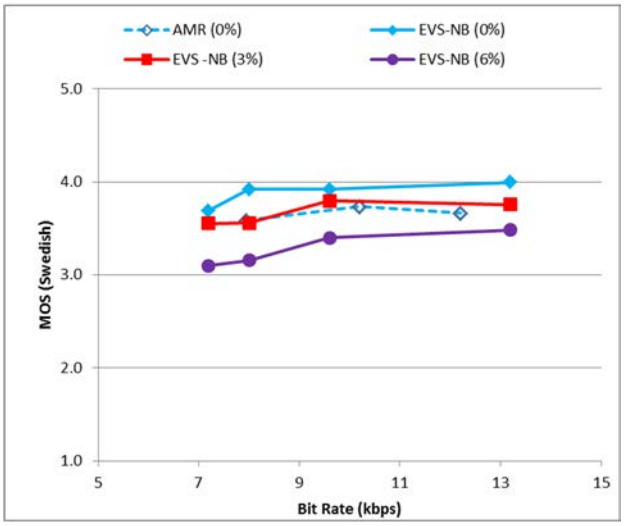
It can be seen that EVS always significantly out-performs AMR in terms of intrinsic audio quality for both speech and Mixed/Music signals. EVS is also significantly more robust to frame erasures; both randomly distributed or according to the Delay and Error profiles from TS 26.114 [4] using the EVS JBM.



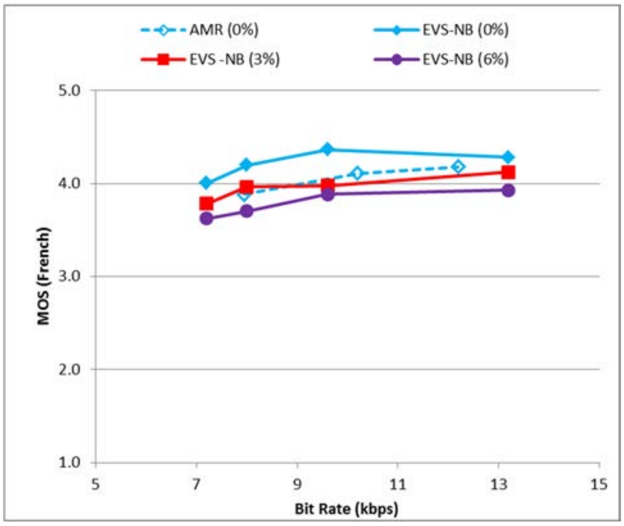
(a)



(b)

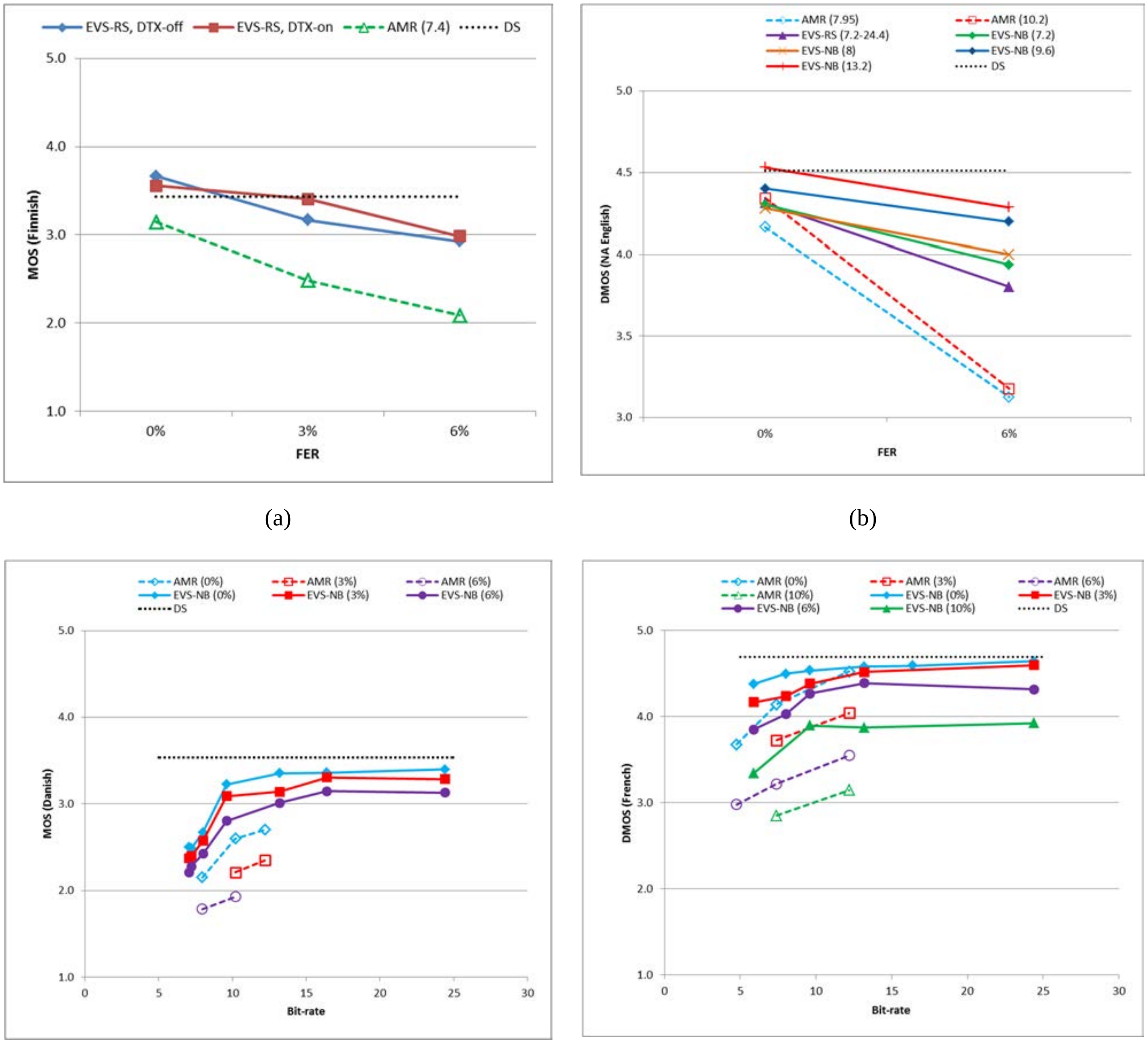


(c)



(d)

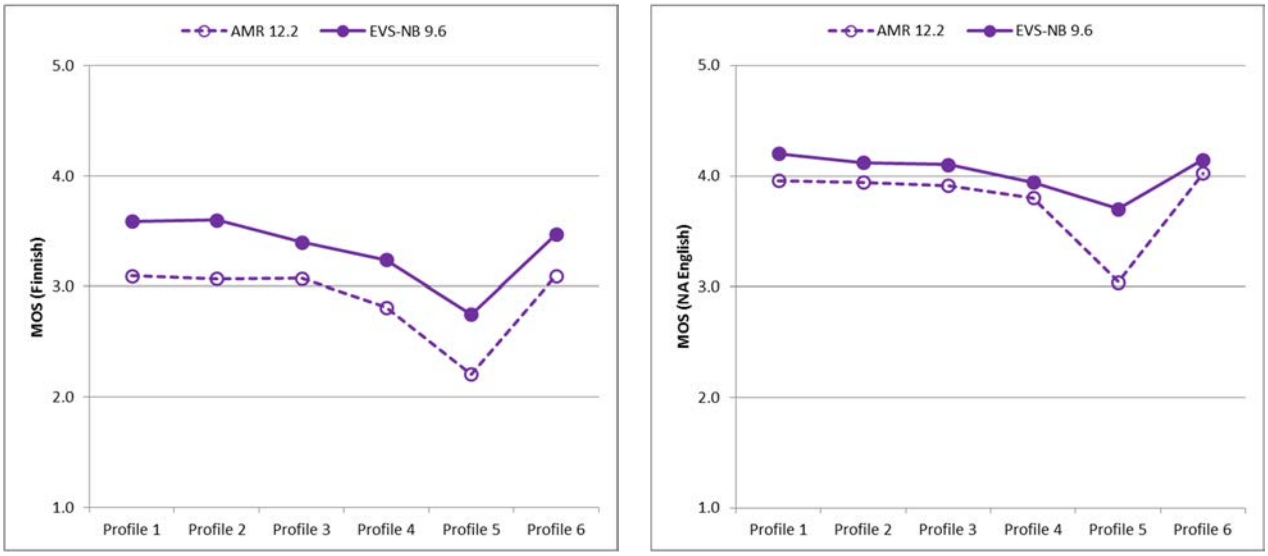
Figure 2: EVS NB vs AMR – Speech - Random Frame Erasures - Selection



(a) (b)

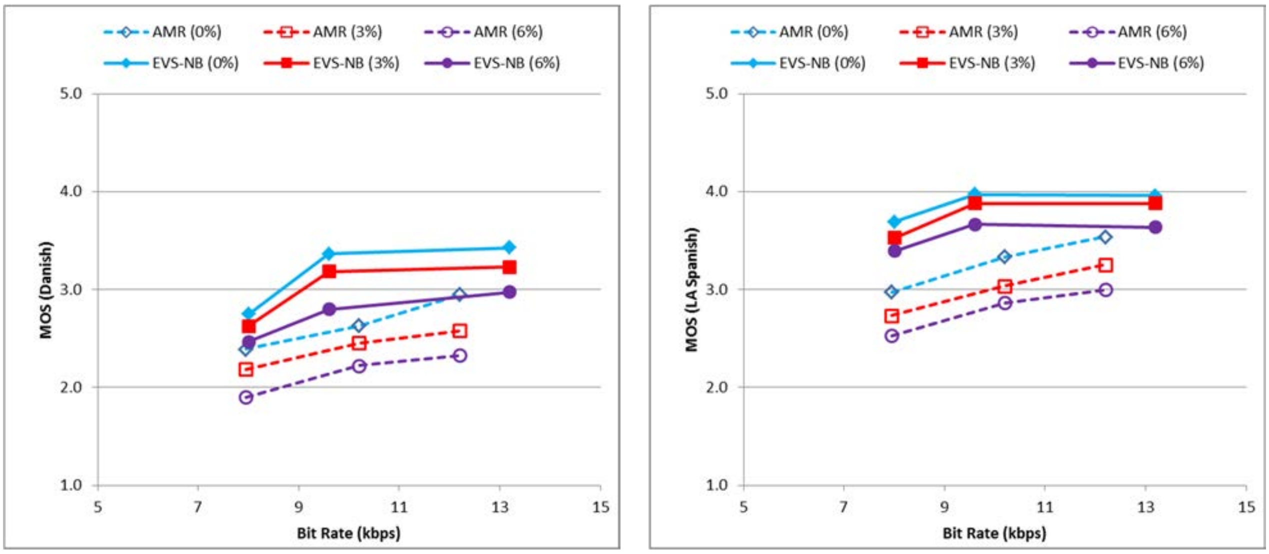
(c) (d)

Figure 3: EVS NB vs AMR – Speech - Random Frame Erasures - Characterization



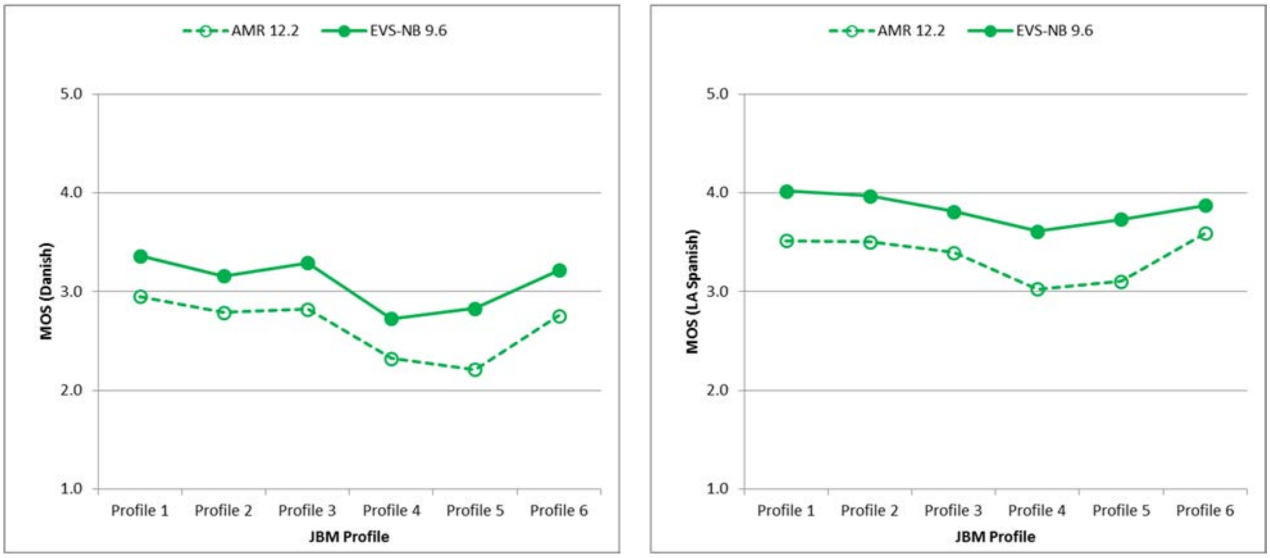
(a) (b)

Figure 4: EVS NB vs AMR – Speech - TS 26.114 Delay & Error Profiles



(a) (b)

Figure 5: EVS NB vs AMR – Music & Mixed Content - Random Frame Erasures



(a) (b)

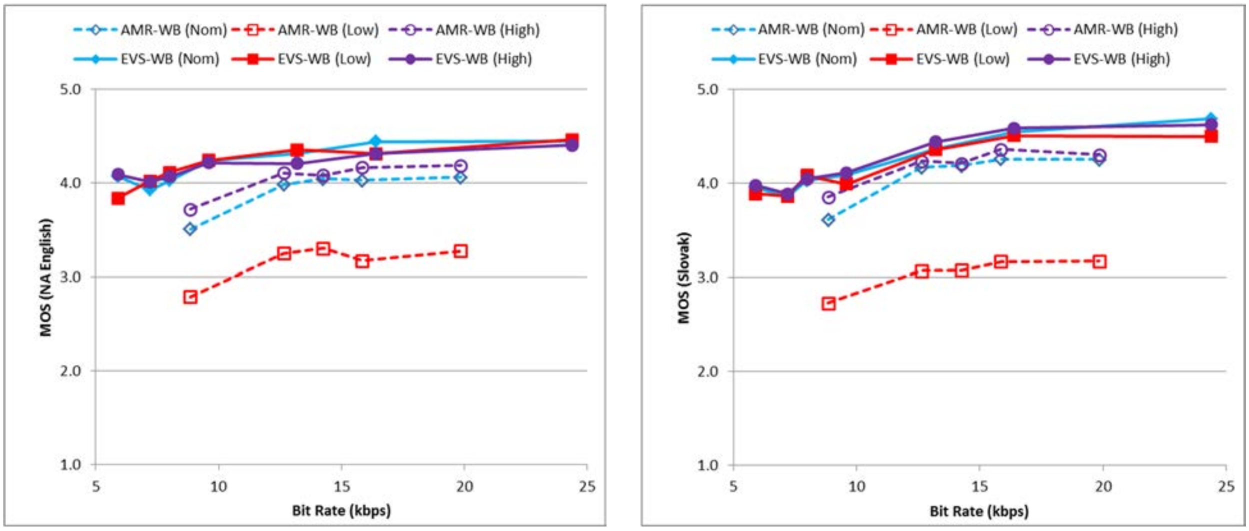
Figure 6: EVS NB vs AMR – Music & Mixed Content - TS 26.114 Delay & Error Profiles

5.1.1.3 Wideband Comparison vs AMR-WB

For Wideband (WB) signals, seven experiments were conducted during the EVS Selection and five experiments during Characterization; focused on determining the performance of the EVS Wideband Primary Modes of operation. Taken together these experiments provide unique information about the performance of EVS with respect to AMR-WB but the highlights are provided below in Figures 7 to 10.

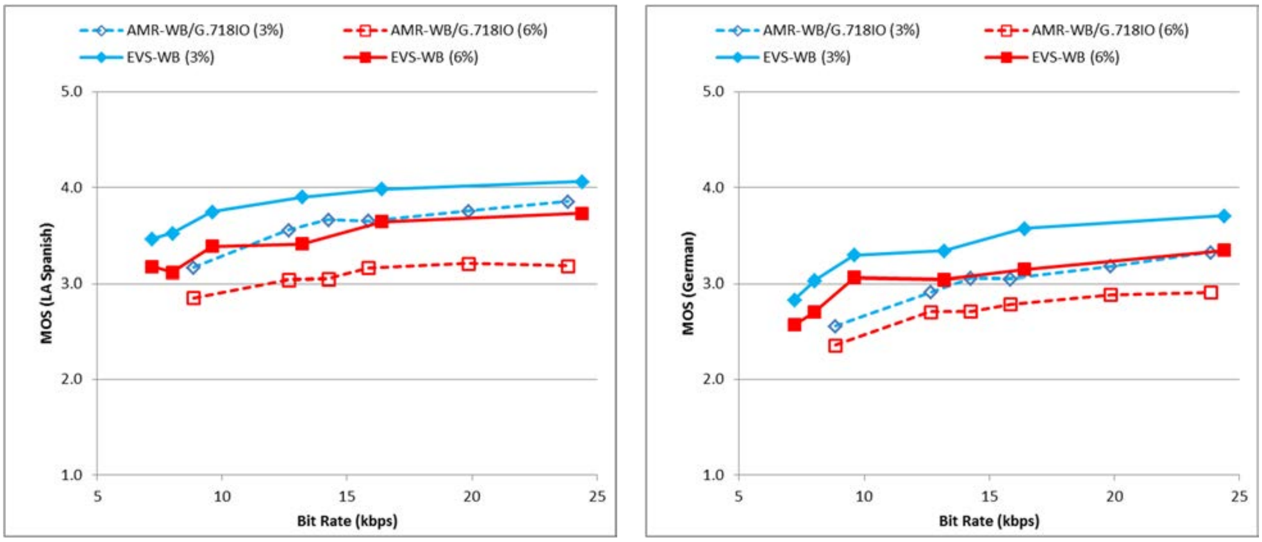
As in the case of AMR and NB, it can be seen that EVS always significantly out-performs AMR-WB or AMR-WB/G.718IO in terms of intrinsic audio quality for both speech and Mixed/Music signals. EVS is also significantly more robust to input level and frame erasures; both randomly distributed or using the EVS JBM in conjunction with the packet delay and error profiles taken from either TS 26.114 or the new profiles defined for LTE.

What is less clear from the frame erasure plots is that AMR-WB, in its basic form, performs significantly less well than these curves would suggest. Work in ITU-T as part of the G.718 exercise led to significant improvements to the packet loss concealment of AMR-WB (G.722.2) and these improvements are shown in Figures 11 & 12 (FER and BFER); taken from the Characterization Report of Recommendation ITU-T G.718 [5]. The enhancements achieved during the development of G.718 formed part of the justification of the EVS work item and thus it can be assumed that EVS will perform even better than suggested by Figures 8, 9 and 10.



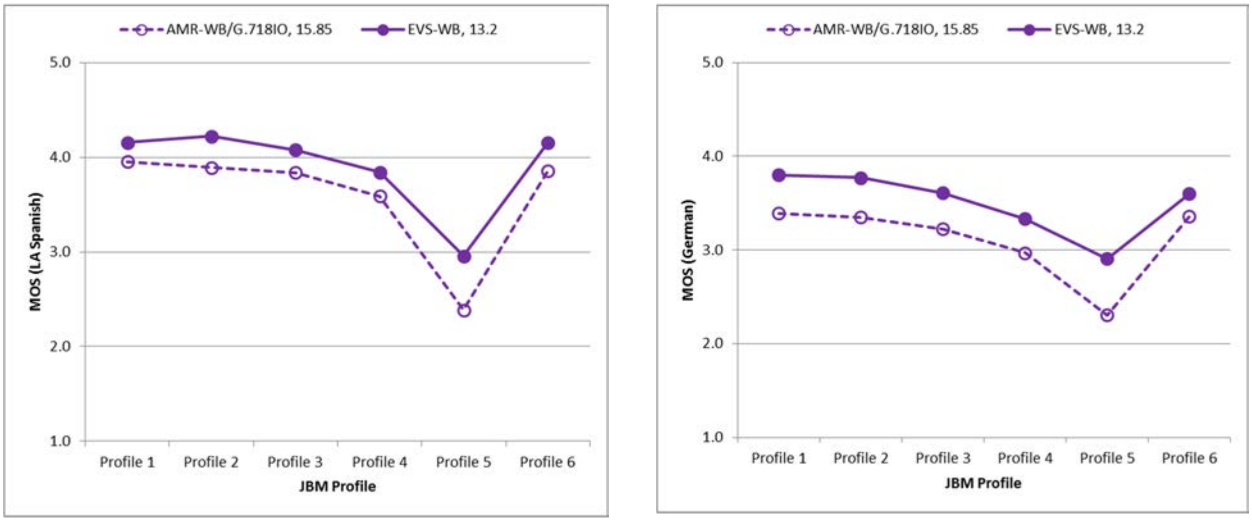
(a) (b)

Figure 7: EVS WB vs AMR-WB – Speech – Clean Channel & Levels



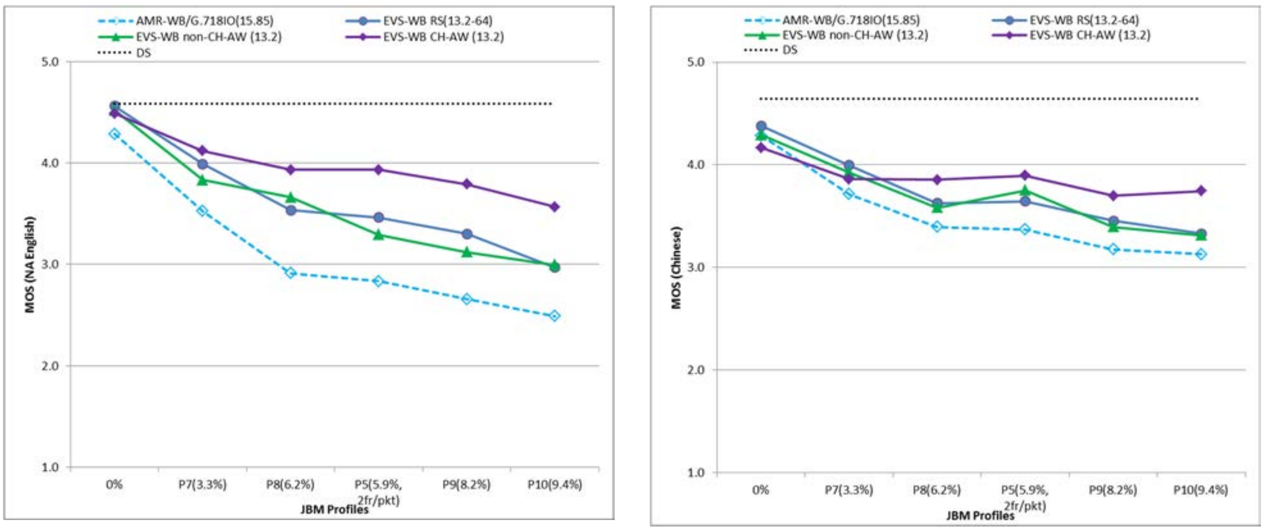
(a) (b)

Figure 8: EVS WB vs AMR-WB – Speech - Random Frame Erasures



(a) (b)

Figure 9: EVS WB vs AMR-WB – Speech - TS 26.114 Delay & Error Profiles



(a) (b)

Figure 10: EVS WB vs AMR-WB – Speech – New EVS JBM Delay & Error Profiles

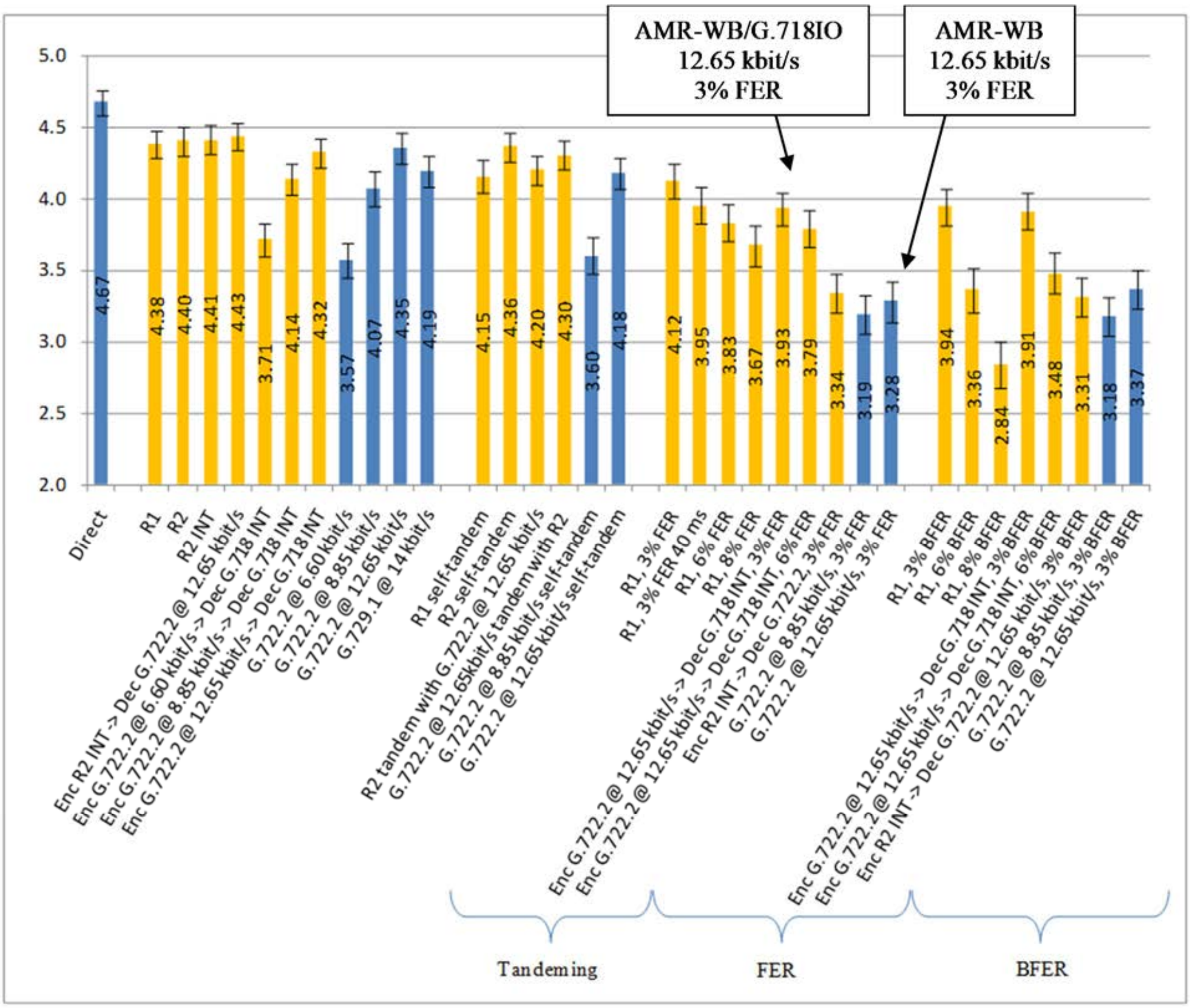


Figure 11: AMR-WB (G.722.2) vs G.718IO – Speech (American English) Figure 27 of [3]

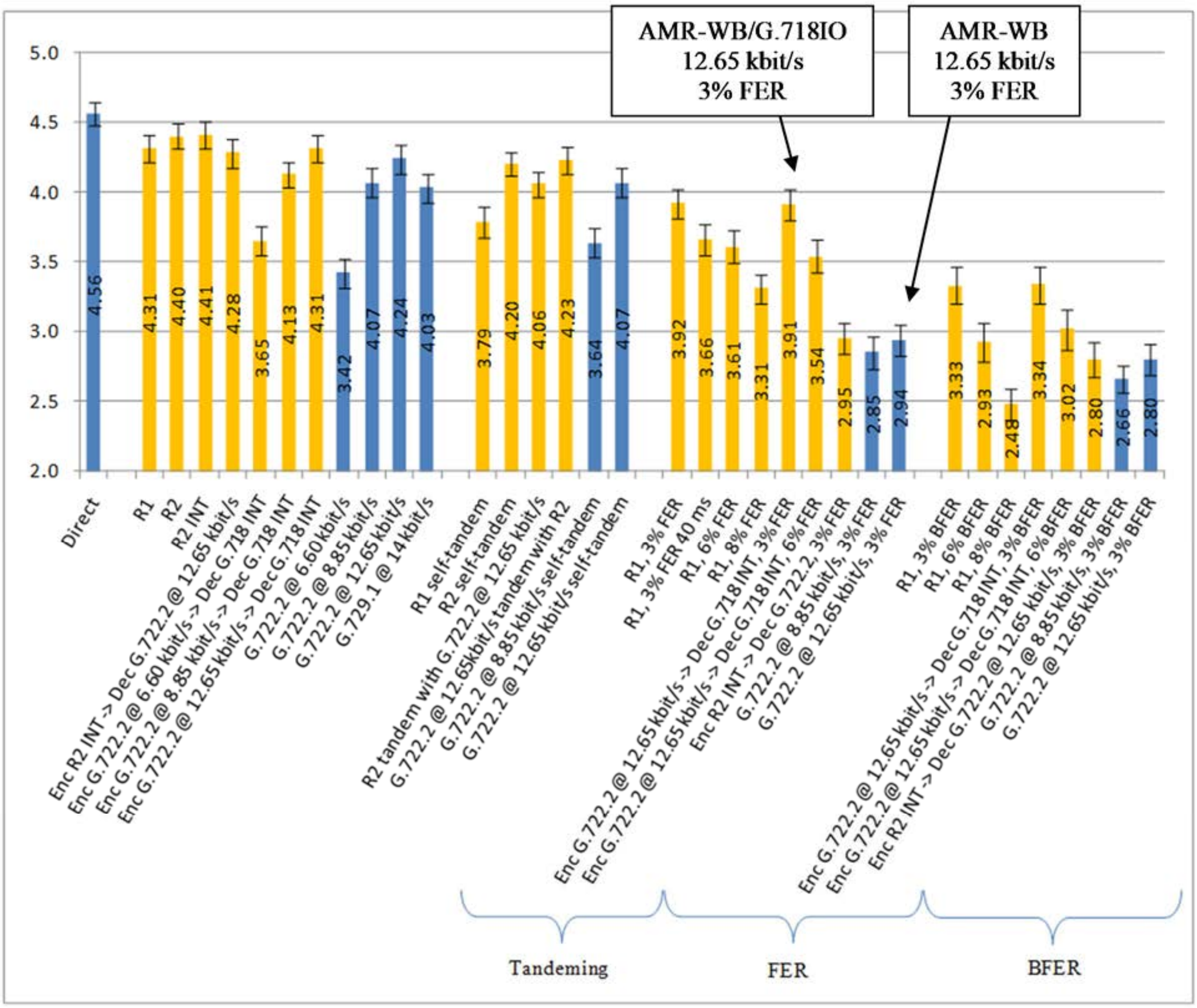


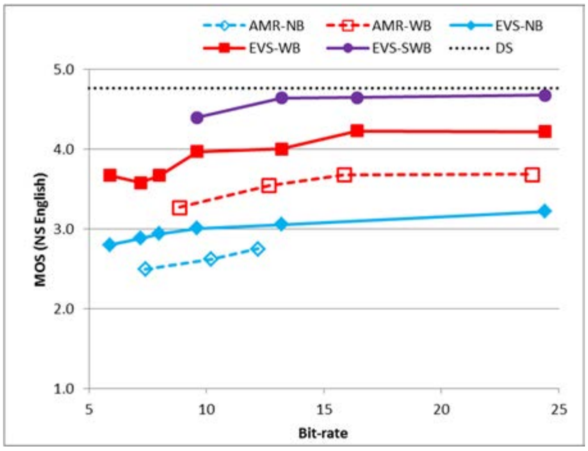
Figure 12: AMR-WB (G.722.2) vs G.718IO – Speech (French) Figure 28 of [3]

5.1.1.4 Super-wideband EVS and Relationships to Other Bandwidths

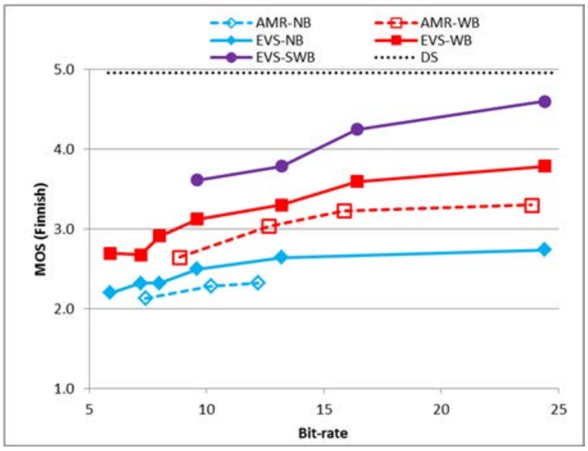
Three mixed bandwidth tests were performed during the EVS Characterization and the results are shown in Figure 13.

It is clear from Figure 13 that the Super-wideband (SWB) modes of EVS outperform the WB modes, which themselves outperform the NB modes. On the whole it is clear that these trends hold across input types and bit rates. The EVS codec can also be seen to scale well with bit rate within each bandwidth and asymptotically approaches the Direct Source (DS) in the case of SWB and progressively lower value in the cases of the reduced bandwidth signals; WB and NB.

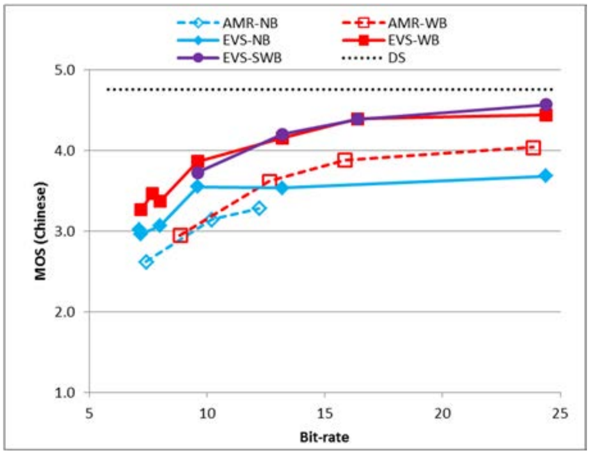
These mixed bandwidth test results also reinforce the performance advantages of EVS compared to AMR and AMR-WB.



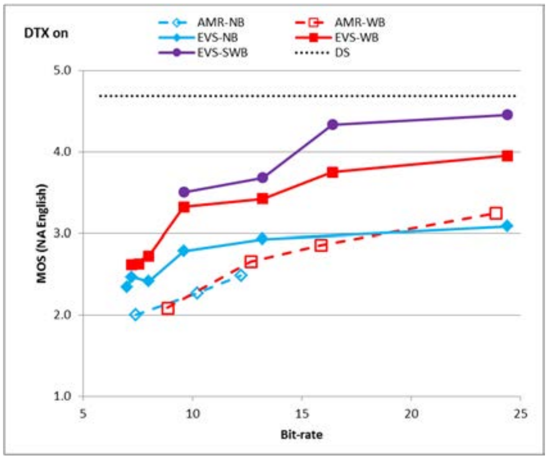
(a) Clean



(b) Car Noise



(c) Music & Mixed Content - Chinese



(d) Music & Mixed Content - US English

Figure 13: EVS vs AMR and AMR-WB – Bandwidth and Bitrate Differences

5.1.1.5 Comparison of the 3GPP Codecs to TETRA

The details of the TETRA codec may be found in ETSI EN 300 395-2 [6]. The bit rate of the TETRA codec is 4,567 kbps and it makes use of the ACELP paradigm and 30ms frames. At the time of its selection in 1993, the TETRA ACELP codec represented the state-of-the-art in low bit rate speech codecs and it was well adapted to its specific application and the TETRA 4:1 TDMA air interface.

According to the Characterization tests conducted during the standardization of the TETRA ACELP codec which are also provided in [6]:

> *"For clean speech at a nominal input level of -22 dB the average Q value obtained for the TETRA codec is 13,0 dB for the linear input condition and 16,5 dB for the IRS input condition. For comparison purposes the corresponding values obtained for the Global System for Mobile communications (GSM) full-rate codec are 17,4 dB and 18,9 dB respectively."*

These differences in dBQ are reproduced below (from [6]) in Tables 1 & 2 for various input signals.

From Tables 1 & 2 it is clear that the TETRA codec is consistently inferior to the original GSM Full Rate Speech Codec of the order of 2.4 - 4.4 dBQ.

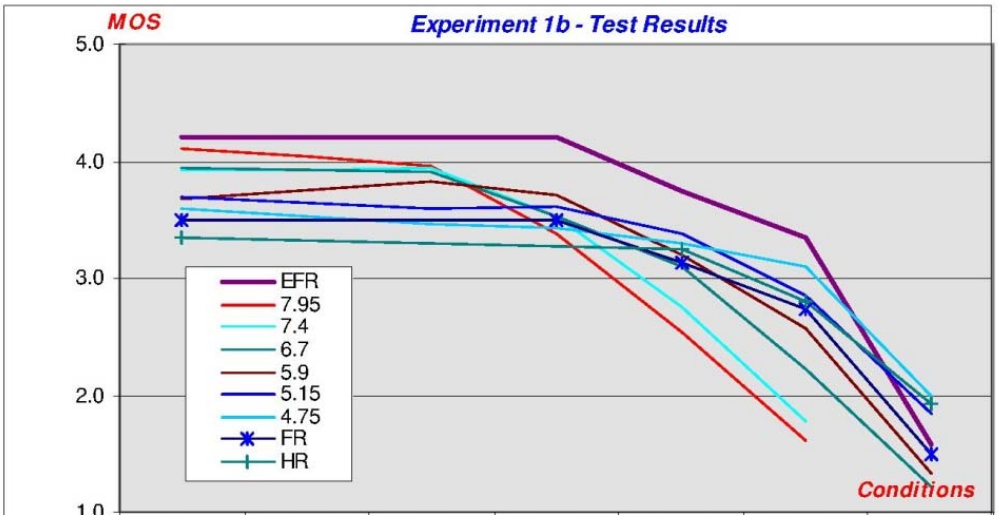
Table 1: TETRA vs GSM TCH-FS for A-Law IRS Input Signals

	TETRA ACELP (Nominal Level) (dBQ)	GSM TCH-FS (Nominal Level) (dBQ)
Quiet	16.5	18.9
Vehicle -10dB	4.1	5.2
Vehicle -20dB	9.5	10.5
Office -10dB	7.2	8.7
Office -20dB	11.4	11.7

Table 2: TETRA vs GSM TCH-FS for FLAT Input Signals

	TETRA ACELP (Nominal Level) (dBQ)	GSM TCH-FS (Nominal Level) (dBQ)
Quiet	13.0	17.4
Vehicle -10dB	6.5	10.0
Vehicle -20dB	9.3	14.7
Office -20dB	9.4	14.6

Since the selection of the TETRA speech coding standard in 1993, firstly SMG, and latterly 3GPP, has developed several generations of codec upgrades for NB speech; firstly the Enhanced Full-Rate Codec (EFR), then the Adaptive Multi-Rate codec (AMR) which included the EFR as the 12.2 kbps mode and most recently the EVS codec. Each of these developments has provided clear and measurable quality improvements over the generation that went before.



	No Errors	C/I=19 dB	C/I=16 dB	C/I=13 dB	C/I=10 dB	C/I= 7 dB	C/I= 4 dB
EFR	4.21			4.21	3.74	3.34	1.58
7.95	4.11	4.04	3.96	3.37	2.53	1.60	
7.4	3.93	3.93	3.95	3.52	2.74	1.78	
6.7	3.94		3.90	3.53	3.10	2.22	1.21
5.9	3.68		3.82	3.72	3.19	2.57	1.33
5.15	3.70		3.60	3.60	3.38	2.85	1.84
4.75	3.59		3.46	3.42	3.30	3.10	2.00
FR	3.50			3.50	3.14	2.74	1.50
HR	3.35				3.24	2.80	1.92

Figure 14: (Figure 5.4 from [7]): AMR Family of Curves for Experiment 1b (Clean Speech in Half Rate)

Examining the performance gains of GSM EFR (= AMR 12.2 - ETSI ETR 305 or GSM 06.55 [8]) over the GSM Full Rate Codec it is clear that EFR significantly outperformed the GSM Full Rate Codec in all tests [8]:

"The EFR codec is better than the actual FR codec for clear speech, for all error conditions (EP1, EP2 andEP3) and for tandeming under error EP1; it is equivalent to G.728 for its intrinsic quality, for background noise conditions and talker dependency."

Considering GSM AMR (ETSI TR 126 975 [7]), from Figure 14 (Figure 5.4 of [7]) it can be seen that all of the AMR coding modes are at least as good as the GSM Full Rate Codec for clean speech [7]. Data is somewhat lacking on noisy speech performance with AMR in [7] but it can be reasonably expected that In noisy speech, the higher bit rates of AMR would exceed the performance of GSM FR due to the increasing similarity with EFR but that the margin would diminish at lower bit rates. However, all of the bit rates of the AMR modes exceed the 4.567 kbps of TETRA.

It can therefore be confidently concluded that the overall perceptual quality of the TETRA Codec will be inferior to that of any mode of the AMR codec.

Such a conclusion is anecdotally supported by the adoption of the AMR 4.75 kbps codec as a codec upgrade to TETRA during the development of the TETRA-2 feature set.

From the EVS Characterization results in TR 26.952 [3] (reproduced in subclause 5.1.1.2) the comparisons between AMR & EVS show an improvement for the EVS codec over all of the coding modes of AMR.

It can therefore be confidently concluded that the perceptual quality of the TETRA Codec is going to be noticeably inferior to any of the EVS NB codec modes. It is also clear from subclause 5.1.1.4 that AMR-WB and the WB, SWB and FB modes of EVS are capable of significantly improving not only the quality, but also the intelligibility, of any MCPTT system when compared to narrowband communication systems such as TETRA and P25. The increased intelligibility of the wider audio bandwidths are also available at bit rates approaching the lower bit rates of AMR with the EVS codec i.e. EVS Wideband VBR (nominally 5.9 kbps) and 7.2 kbps compared to AMR 4.75, 5.9, 6.7 and 7.4 kbps. This feature of the EVS codec simultaneously satisfies the requirements for improved intelligibility and for radio resource efficiency given in subclauses 5.14, 6.15.5 and 6.15.6 of [2].

5.1.1.6 Comparison of Performance over MCPTT Bearers

5.1.1.6.1 "HD-Voice" AMR-WB performance over 3GPP networks

The "HD Voice" reference describes the quality of experience across the listed KPIs (speech quality, speech intelligibility, error resiliency, and call capacity) of AMR-WB in today's commercial VoLTE networks.

For instance AMR-WB at 12.65 kbps operating over a unicast LTE PS channel with 1% FER per mobile link (as specified for QCI=1) resulting in 2% total FER in mobile-to-mobile calls.

To characterize the reference coverage in a VoLTE system using unicast power-controlled channels with HARQ Re-TX and QCI = 1 this document uses the VoLTE field test results illustrated in figures 5.1.1.6.1-1 and 5.1.1.6.1-2 below. Figure 5.1.1.6.1-2 excludes the zero RTP loss rate data to allow the reader to see the non-zero cases more clearly.

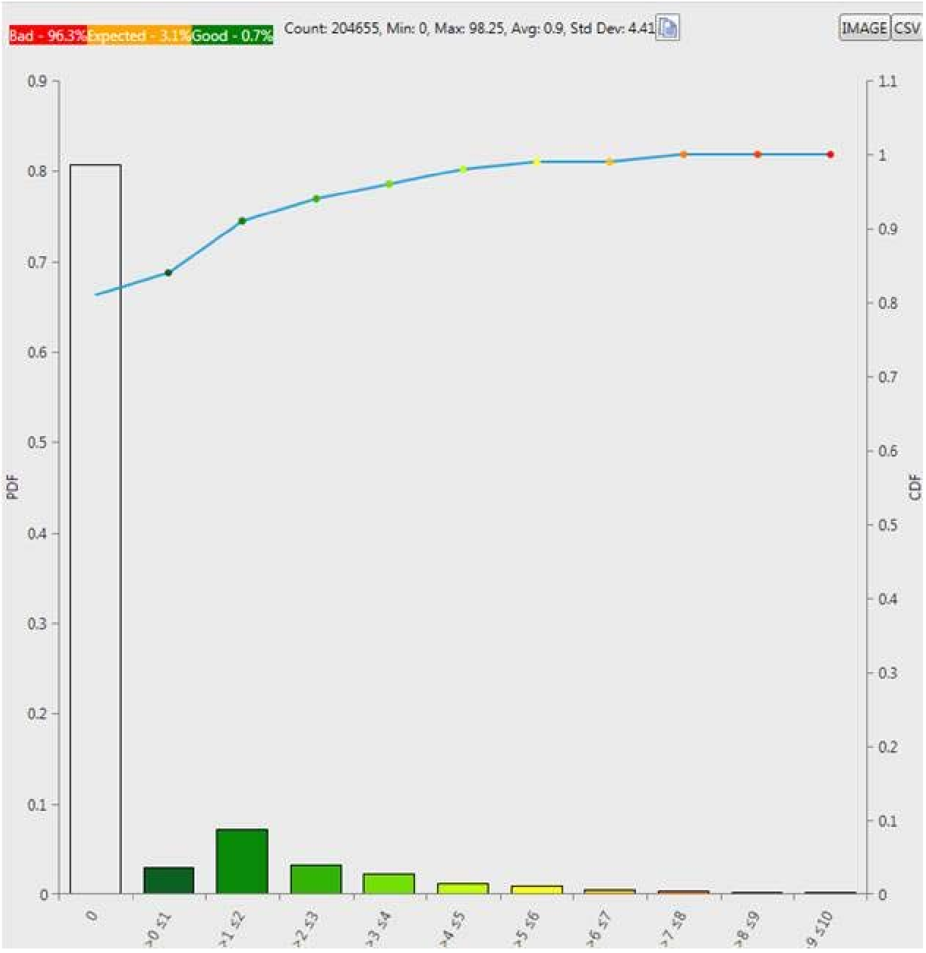


Figure 5.1.1.6.1-1 CDF of end-to-end RTP packet loss rate for VoLTE mobile-to-mobile calls. Zero RTP loss values INCLUDED.

These measurements are based on logs taken over 6000 calls over various commercial LTE networks spanning multiple continents, with each call averaging 34 seconds in duration (actually a mix of many short 30s calls and several hours of long calls). The RTP loss rate is calculated over 1 second windows and includes stationary and mobile UE's in good and bad coverage conditions.

It can that seen that about only about 90% of the cell area has an end-to-end FER <=2%. This is interpreted to mean that the reference "HD Voice" coverage is equivalent to 90% of the cell area. In the remaining 10% the AMR-WB codec speech quality starts to degrade at FERs above 2%.

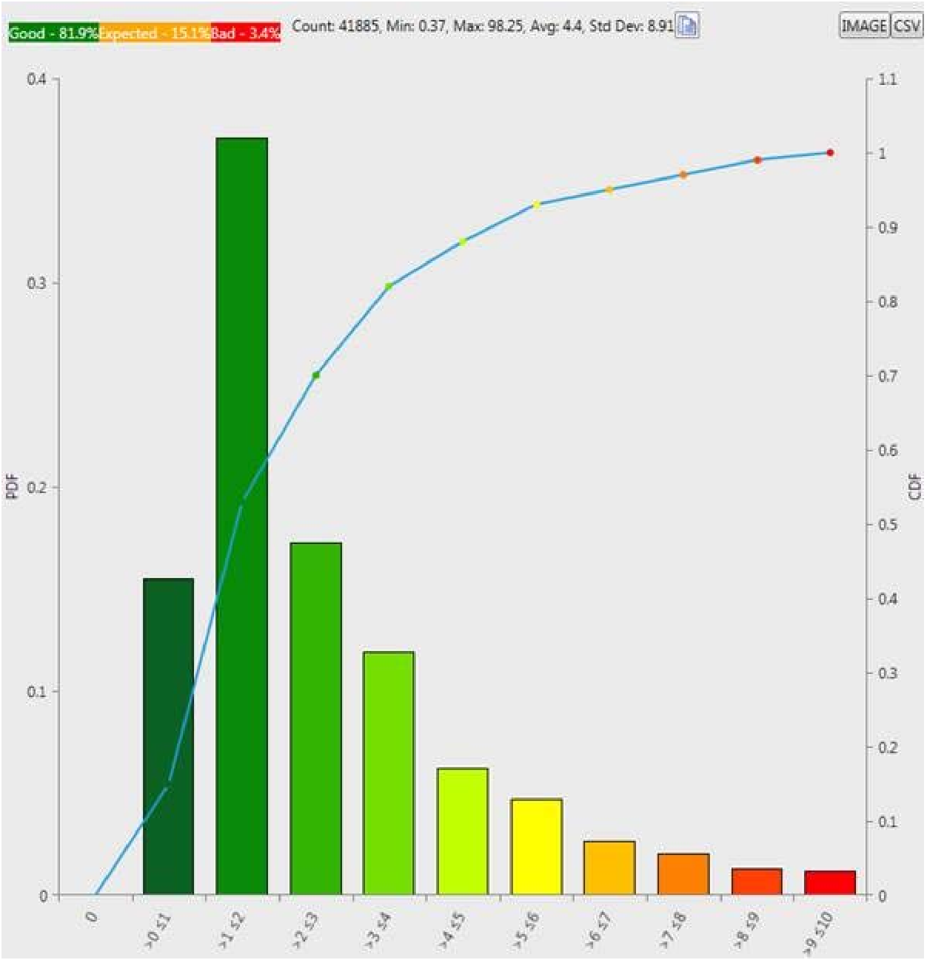


Figure 5.1.1.6.1-2 CDF of end-to-end RTP packet loss rate for VoLTE mobile-to-mobile calls. Zero RTP loss values EXCLUDED.

5.1.1.6.2 MCPTT Bearers

The MCPTT service can be operated over three types of bearers depending on the network topology that is most appropriate among those available. The following clauses describe these bearers and also the channel models used to provide the simulation results in the next clause.

5.1.1.6.2.1 Unicast bearer

MCPTT can be operated over unicast channels in the same way the teleconferencing is performed in today's mobile networks using a central conferencing server for duplicating and distributing media (Figure 5.1.1.6.2.1-1).

Each of the LTE unicast channels is a power-controlled channel that also use retransmission schemes such as HARQ to provide a target BLER or packet loss rate to the VoIP frames transmitted over the channel.

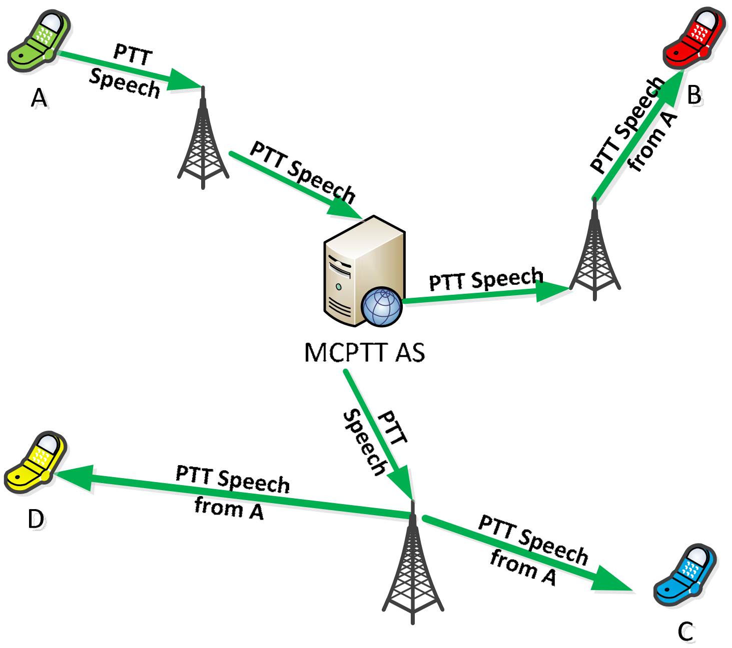


Figure 5.1.1.6.2.1-1 MCPTT topology using unicast bearers

When using AMR-WB in this topology, the coverage, error-resiliency, speech quality, speech intelligibility, and call capacity is equivalent to that of "HD Voice."

When evaluating the performance of EVS and AMR-WB over the unicast bearers, the EVS characterization report [3] used the delay-loss profiles defined in [4] and introduced additional delay-loss profiles.

5.1.1.6.2.2 MBMS bearer

When multiple participants in a group are in a single cell the system can reduce the resources needed to support the users by having them share a common downlink MBMS bearer. This shared channel has the following characteristics:

1) It is not power-controlled. There is no dynamic feedback by which the eNB can decide to dynamically adjust its transmission resources to improve error performance or meet a target error rate.

2) Use of retransmissions is "blind" in that the retransmissions are not sent based on dynamic feedback such as ACK/NACKs. These retransmissions cannot be used to guarantee a certain level of performance or target error rate throughout the cell.

Therefore, error rates on the MBMS bearer can vary considerably throughout the cell, e.g., indoors, basements, elevators, stairwells, or the edge of cell in an SC-PTM topology (see below).

The topology for using an MBMS bearer can be configured in two ways:

1) As a Single-Cell Point-to-Multipoint (SC-PTM) bearer where adjacent cells do not necessarily transmit the same group's content on the same MBMS bearer. In this topology the adjacent cells typically interfere with the MBMS bearer in the serving cell resulting in poorer coverage than the MBSFN topology.

2) As part of a MBSFN, where all the cells are broadcasting the same content on the same MBMS bearers, preventing inter-cell interference and allowing the users to combine these transmissions to improve coverage and reception.

The simulation model used to evaluate the performance of 3GPP speech codecs over MBMS bearers is described in clause A.1

5.1.1.6.2.3 LTE-D bearer

LTE-Direct communication is a broadcast mechanism (no physical layer feedback) that defines two physical channels, control and data, for communication between two (or more) UEs. The resources used for direct communication comprise of control and data resources. For in-network operation, a control resource pool is provided via RRC signalling while for off-network operation, the control resource pool is pre-configured. Further, two modes of resource allocation are supported: Mode 1 (in-network) and Mode 2 (in-network and off-network) as illustrated in Figure 5.1.1.6.2.3-1.

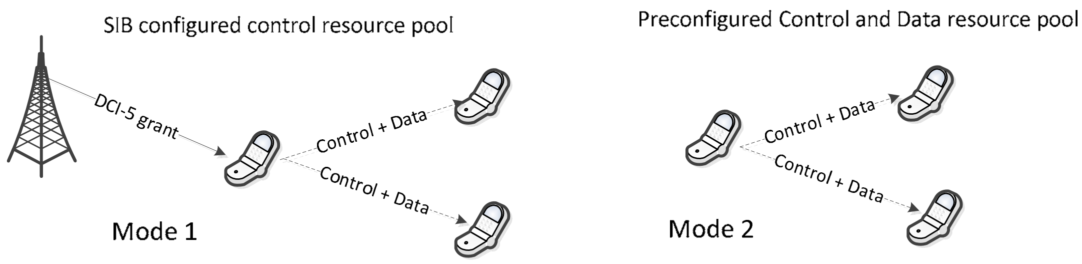


Figure 5.1.1.6.2.3-1 LTE-D operation

Here, the focus is on Mode 2 (off-network) scenario. In Mode 2, the transmitting UEs determine the resources to be used for control and data transmission. UE transmits control to announce resources to be used for subsequent data transmission. Receiving UEs monitor the control resources to determine when to wake-up and listen for data transmission.

The performance of LTE-D based on system and link level simulations for off-network scenario is evaluated. In particular, the mandatory Option 5 hotspot drop has been used for system level simulations.

5.1.1.6.3 AMR-WB and EVS Performance over the MCPTT Bearers

5.1.1.6.3.0 General

Enhanced Voice Services (EVS) is a new speech codec standard (part of 3GPP Release 12) which offers a wide range of new features and improvements for low delay real-time communication. The key advancements fall into three categories namely significantly improved quality for clean/noisy speech and music content, higher compression efficiency and unprecedented error resiliency to packet loss and delay jitter experienced in PS systems. In addition to voice quality and intelligibility aspects, we present methods on how to utilize some of the EVS codec advancements to realize MCPTT system level benefits such as improved coverage and call capacity gains.

5.1.1.6.3.1 EVS Speech Quality

EVS Selection and Characterization Phase Test Results are summarized in the main body and detailed in Annex D of TR 26.952 [3]. In this clause a few test results are highlighted to quantify the improvements of the EVS codec along the three dimensions listed above, i.e., speech quality, compression efficiency, and error resiliency. To further simplify the performance comparison a reference point for benchmarking is established, namely AMR-WB at a bit-rate of 12.65 kbps based on commercial grade HD Voice services available today.

NOTE. The correlation between voice quality and intelligibility is dependent on the test parameters. In general, improved voice quality may result in improved intelligibility. However, it is also possible e.g., in noisy conditions of [-30 dB to 5 dB SNR], that the improvements observed using subjective voice quality testing and the improvements observed using subjective intelligibility testing may not correlate well. In the other end of spectrum, e.g., in clean speech, while the voice quality may have improved significantly, the intelligibility may already have approached a level of saturation.

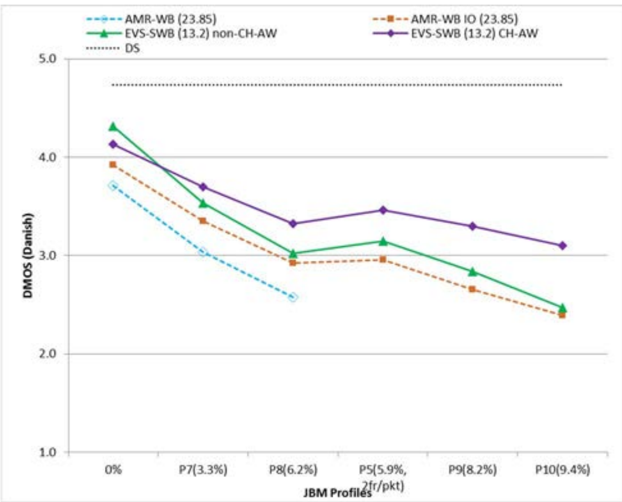
Three mixed bandwidth DCR (Degradation Category Rating) tests were performed as a part of EVS Characterization testing whose results are shown in Figure 13.

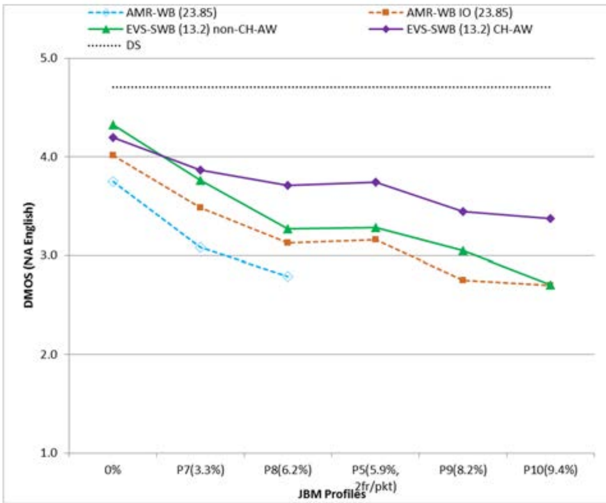
In general, EVS-WB codec offers quality significantly better than AMR-WB at a similar bit-rate and quality equivalent to AMR-WB at a lower bit rate. The EVS-SWB codec performance is significantly better than both AMR-WB and corresponding bit rates of EVS-WB.

For clean speech content (Figure 13a), the lowest bit-rate of EVS-WB namely 5.9 kbps can offer quality significantly better than AMR-WB at 8.85 kbps and equivalent to AMR-WB at 12.65 kbps. The subjective quality of EVS-WB coding starting at 9.6 kbps is significantly better than the AMR-WB coding at its highest bit rate of 23.85 kbps. The super-wideband mode of EVS at 13.2 kbps achieves transparency to the direct source and offers quality significantly better than both 23.85 kbps of AMR-WB and 24.4 kbps of EVS-WB.

For noisy speech (Figure 13b), EVS-WB at 9.6 kbps offers quality on par with AMR-WB at 12.65 kbps. This has also been shown across different languages/noise types and summarized in TR 26.952. However, none of the noisy speech tests included the presence of a front end noise suppression, which is expected to establish the equivalence to AMR-WB 12.65 kbps quality at a bit-rate lower than 9.6 kbps by providing a higher SNR at the input to the coder. EVS-WB at 13.2 kbps offers quality on par with AMR-WB at approximately twice the bit-rate with consistent progression in subjective quality with increasing bit-rates. The subjective quality of EVS-SWB coding at 13.2 kbps is significantly better than that of AMR-WB at 23.85 kbps and EVS-WB at the same bit-rate.

For mixed/music coding (Figure 13c, 13d) under clean channel conditions, both the EVS-WB and SWB codec starting at 13.2 kbps achieves subjective quality that is significantly better than that of AMR-WB at any bit-rate. For North American English music and mixed content (Figure 13d), EVS-SWB coding performs significantly better than EVS-WB at the same bit rate.





(a) (b)

Figure 5.1.1.6.3.1-1: EVS-SWB Channel aware mode clean speech performance under clean and impaired channels (Mixed bandwidth DCR Test), (a) Danish, (b) North American English

Figure 5.1.1.6.3.1-1 shows the EVS-SWB channel aware mode performance at 13.2 kbps under clean channel as well as under five different delay/loss profiles (Profiles 5, 7, 8, 9, and 10) which simulate impaired channel characteristic with varying delay and jitter. Profile 5 is a MTSI delay loss profile from TS 26.114 [4] and profiles 8-10 are VoLTE delay loss profile used for characterization testing of EVS channel aware mode [http://www.3gpp.org/ftp/tsg_sa/WG4_CODEC/EVS_Permanent_Documents/EVS-7c_S4-141392.zip]

Most 3GPP networks are expected to be configured such that the FER is around 1% for each link. While the 2% data point was not tested in this test, comparisons were made between the EVS modes versus AMR-WB at the nearest data points namely, 0% (clean channel) and 3% FER.

In general, the 13.2 kbps EVS-SWB clean speech performance under impaired channel with channel aware mode enabled is significantly better than without channel aware mode which in turn is significantly better than AMR-WB at its highest bit-rate of 23.85 kbps. For both languages, the quality of EVS SWB 13.2 kbps channel aware and non-channel aware modes in clean channel are significantly better than AMR-WB at its highest bit-rate of 23.85 kbps.

For North American English (Figure 5.1.1.6.3.1-1b), EVS SWB 13.2 kbps channel aware mode operating at around 6% FER delivers quality on par with that of the highest bit-rate of AMR-WB (23.85 kbps) under no loss. The 13.2 kbps SWB non-channel aware mode is able to achieve the quality equivalence to AMR-WB 23.85 clean channel when operating at around 3% FER. EVS 13.2 kbps channel aware mode even at 10% FER delivers quality better than AMR-WB 23.85 kbps at 3% FER, while the 13.2 kbps EVS SWB non-channel aware mode can operate at 8% FER but achieve quality equivalence to AMR-WB 23.85 kbps at 3% FER.

For Danish (Figure 5.1.1.6.3.1-1a), EVS SWB 13.2 kbps channel aware mode operating at around 3% FER delivers quality on par with that of the highest bit-rate of AMR-WB (23.85 kbps) under no loss. EVS 13.2 kbps channel aware mode even at 10% FER delivers quality equivalent to AMR-WB 23.85 kbps at 3% FER, while the 13.2 kbps EVS SWB non-channel aware mode can operate at 6% FER to achieve quality equivalence to AMR-WB 23.85 kbps at 3% FER.

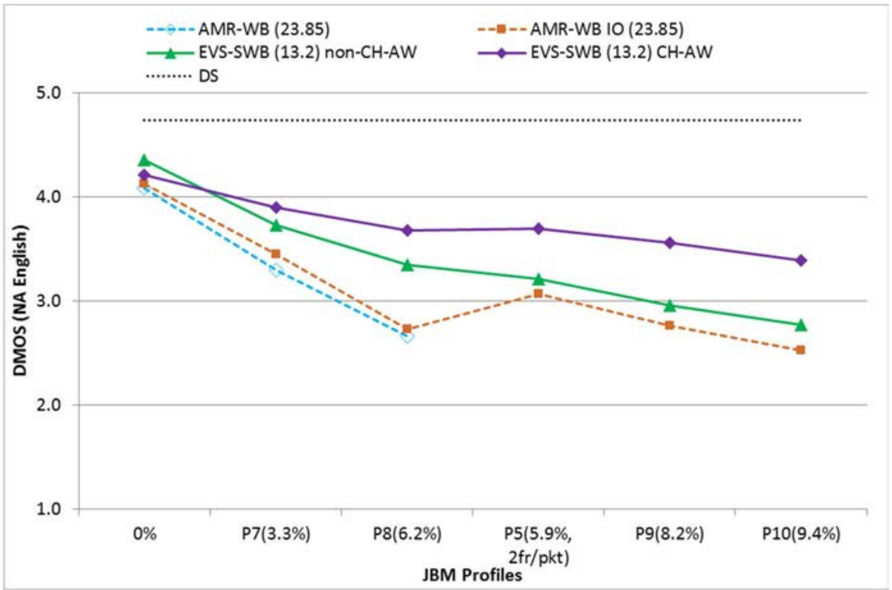


Figure 5.1.1.6.3.1-2: EVS-SWB Channel aware mode noisy speech performance under clean and impaired channels (North American English with car noise @ 15 dB SNR- Mixed bandwidth DCR test)

In general, the 13.2 kbps EVS-SWB noisy speech performance under impaired channel with channel aware mode enabled is significantly better than without channel aware mode which in turn is significantly better than AMR-WB at its highest bit-rate of 23.85 kbps.

For North American English with car noise at 15 dB SNR, EVS SWB 13.2 kbps channel aware mode operating at 10% FER and the EVS SWB non-channel aware mode operating at 6% FER can achieve quality equivalence to AMR-WB 23.85 kbps at 3% FER.

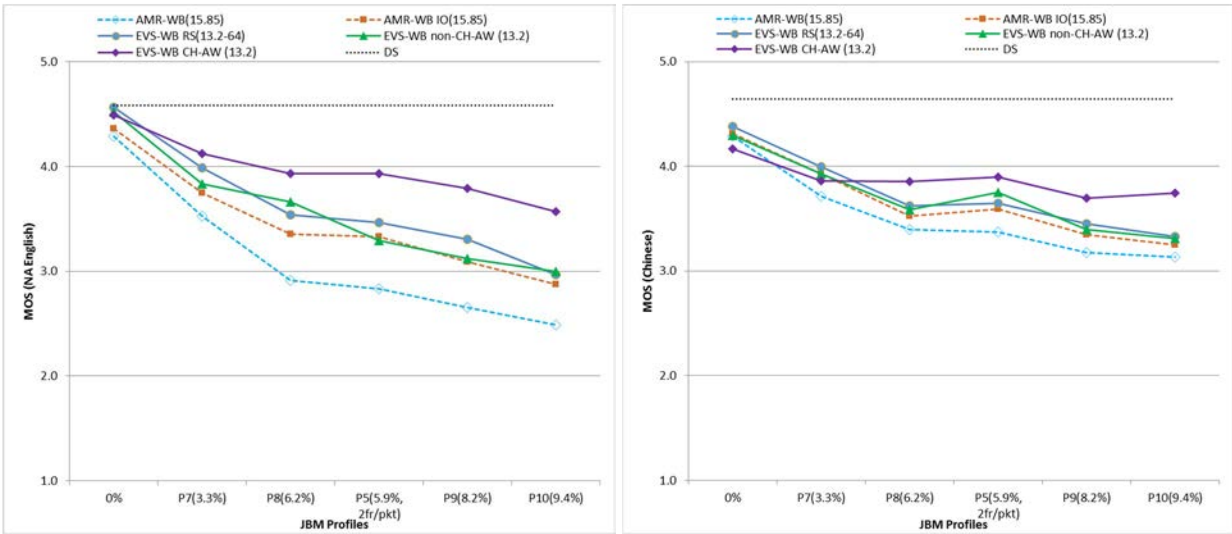


Figure 5.1.1.6.3.1-3: EVS-WB Channel aware mode clean speech performance under clean and impaired channels (Single bandwidth ACR test –Experiment W1 from EVS Characterization Testing) (a) North American English, (b) Mandarin

For North American English (Figure 5.1.1.6.3.1-3a), the 13.2 kbps EVS-WB clean speech performance under impaired channel with channel aware mode enabled is significantly better than without channel aware mode which in turn is significantly better than AMR-WB at 15.85 kbps. The quality of EVS WB 13.2 kbps channel aware and non-channel aware modes in clean channel are significantly better than AMR-WB at 15.85 kbps.

Specifically, it can be seen that for both languages tested, the EVS 13.2 kbps channel aware mode operating at 10% FER can deliver quality on par with AMR-WB at 15.85 kbps at 3% FER. In addition, the 13.2 kbps non-channel aware mode can operate at 6% FER to achieve equivalence to AMR-WB 15.85 kbps at 3% FER.

Since AMR-WB 12.65 kbps at 2% FER ("HD Voice" center of cell speech quality) quality benchmark was not included in the tests summarizes in Figures 5.1.1.6.3.1-3a and 5.1.1.6.3.1-3b, the source used P.OLQA to determine the frame error rate at which the 13.2 kbps EVS WB channel aware and non-channel aware modes is equal to the "HD Voice" reference speech quality.

The source presents below the result of a study to characterize the correlation between subjective MOS and P.OLQA [10] for the North American English Absolute Category Rating (ACR) MOS test results shown in Figure 5.1.1.6.3.1-3a. A North American English database comprising of 3 male and 3 female talkers with 5 sentence pairs per talker was used for computing the P.OLQA scores. The sentence pair wise scores for each condition was averaged to obtain a single P.OLQA score for the condition.

NOTE. As per [12], POLQA has been tested to work with background noise levels. A fixed lower SNR limit at which POLQA can be applied was not reported in [12] and it was noted that it is highly signal-dependent. POLQA analysis in this Clause is done on clean speech in clean and noisy channel conditions. POLQA analysis for extreme noisy conditions [10 to -5 dB] is not performed.

To study the correlation between the POLQA and subjective MOS, the relationship between the two sre plotted in Figure 5.1.1.6.3.1-4. The data points in this plot include the subjective MOS scores from Figure 5.1.1.6.3.1-3a and POLQA scores computed as described above. The data points used for this plot include all conditions (clean channel and all delay loss profiles tested) of AMR-WB 15.85 kbps, EVS-WB 13.2 kbps non-channel aware and channel aware modes (a plot showing the correlation with the channel and non-channel aware modes separated is provided in clause A.2 of the Annex). The "Linear (AMR-WB)" line represents the linear regression between AMR-WB MOS and the corresponding POLQA scores. Therefore, assuming that the "Linear (AMR-WB)" represents the true relationship between POLQA and subjective MOS, then it is seen that POLQA underestimates the subjective quality of EVS WB 13.2 kbps modes.

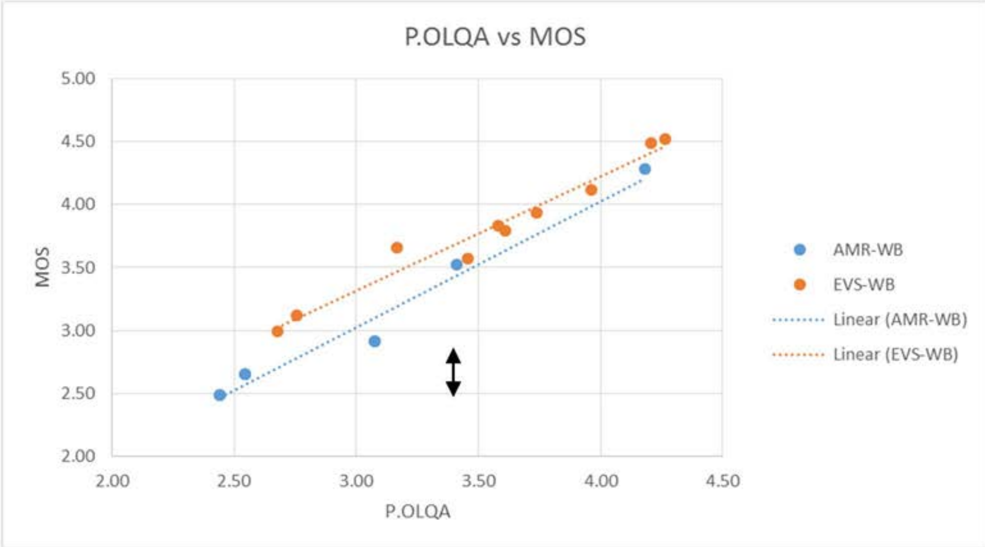


Figure 5.1.1.6.3.1-4: Correlation between Subjective MOS and P.OLQA

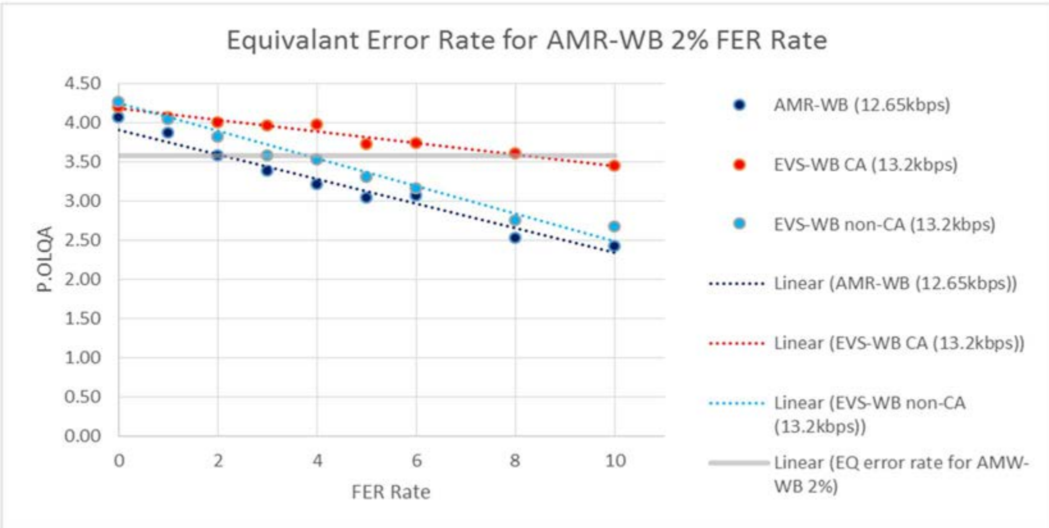


Figure 5.1.1.6.3.1-5: Use of P.OLQA to determine frame error rate for EVS WB 13.2 kbps channel aware and non-channel aware modes at same speech quality as AMR-WB 2% FER

As illustrated in Figure 5.1.1.6.3.1-5, EVS WB 13.2 kbps channel aware mode operating under 8% FER results in the same P.OLQA score as AMR-WB 12.65 kbps at 2% FER. The corresponding data point for the non-channel aware mode is 4% FER. The 3, 6, 8 and 10% FER data points (anchors) shown on the X-axis in Figure 4.1.6 were simulated via delay loss profiles 7, 8, 9 and 10 respectively (used in ACR MOS results shown in Figure 4.1.4a). Delay loss profiles to simulate the other FER data points were created as a random subset of the nearest highest FER anchor. For example the 1 and 2% FER profiles were created as a subset of delay loss profile 7 (3% FER).

Taking into account the under estimation observed in Figure 5.1.1.6.3.1-4 and results from the subjective MOS tests shown in Figure 5.1.1.6.3.1-3a and 5.1.1.6.3.1-3b, it can be seen that the above determination of frame error rate for EVS-WB 13.2 kbps channel aware and non-channel aware modes to achieve the same speech quality as AMR-WB 12.65 kbps at 2% FER, is conservative. The respective SWB modes will provide significantly improved voice quality as compared to the 13.2 kbps WB modes which will result in even higher frame error rates than what is determined above for the WB modes.

5.1.1.6.3.2 Unicast bearer

Speech quality: Power controlled LTE unicast channels are typically configured to operate at a target BLER of 1% per link. The results presented in clause 5.1.1.6.3.1 show that EVS WB at 13.2 kbps (channel aware or non-channel aware) can offer significant quality improvement over "HD Voice" quality. Furthermore the EVS-SWB mode operating at 13.2 kbps (channel aware and non channel aware) offers significantly better audio quality than EVS-WB at the same bit-rate. This applies to a wide range of input signals which include clean speech, speech with background noise and mixed/music content.

Speech intelligibility: EVS offers significant voice quality improvement over AMR-WB (HD voice). The improved robustness to background noise and resiliency to errors are particularly relevant to MCPTT service. In clean channel conditions, as observed in the NTIA report [16] for the very low SNR conditions in the range of [10 to -5 dB], the intelligibility performance of EVS FB codec is equivalent or better than that of AMR-WB (HD voice). Intelligibility in channel errors is not evaluated and is for further study.

Error resiliency and Coverage: Although retransmission schemes such as HARQ maybe used for tight control of the target BLER, due to the power limited uplink, the cell edge or deep indoors may still experience higher BLER (>1%). As seen in clause 5.1.1.6.3.1, under these conditions the EVS WB and SWB channel aware mode will offer significantly better speech quality than AMR-WB at 12.65 kbps due to improved error resiliency. It can be appreciated that the EVS WB channel aware mode at 13.2 kbps can tolerate up to 8% FER and still deliver the same speech quality as AMR-WB 12.65 kbps operating at 2% FER which is center of the cell "HD Voice" speech quality. The ability to sustain the link while tolerating higher path loss results in improved link budget/coverage. The EVS SWB channel aware mode at 13.2 kbps can tolerate even higher FER (up to 10%) for further extending coverage while maintain HD Voice center of cell speech quality. The 13.2 kbps EVS WB and SWB non-channel aware modes can also operate at higher FER(4% for WB non-channel aware mode as shown in clause 4.1) and deliver "HD Voice" center of cell speech quality thereby resulting in improved coverage, albeit lower than that of the channel aware mode.

Call capacity: The 13.2 kbps EVS modes utilize the same transport block size as AMR-WB 12.65 kbps. This results in the same cell site voice capacity as AMR-WB 12.65 kbps. If coverage is kept a constant, the improved error resiliency can be utilized for capacity gains by increasing the packet loss rate by not transmitting the packet at the UE itself. The power at which each packet is transmitted would not be lowered but this mechanism can result in power savings at the UE due to reduced ON time or reduced number of transmissions (analogous to DTX or blanking but for active speech). Incorporating this into the scheduler can reduce the average number of resource blocks required thereby freeing up resources either to add more users or for best effort traffic. The capacity gains are directly proportional to the maximum FER rate at which the EVS mode can still maintain HD voice centre of cell voice quality. Correlating to speech quality discussed in clause 5.1.1.6.3.1, at 13.2 kbps the EVS SWB channel aware mode would offer the highest capacity gains followed by EVS WB channel aware, EVS SWB and WB non-channel aware modes. For example, reducing ON time by 10% (i.e. blanking 10% of the active frame vocoder packets) when operating in EVS SWB channel aware mode can result in 18% capacity gains measured in terms of number of additional users per cell site.

The lower bit-rates of EVS namely the 5.9 VBR and 7.2 kbps WB modes can offer significant cell site voice capacity improvements due to utilizing smaller transport block sizes/resource blocks. It has been demonstrated that the EVS VBR 5.9 mode can achieve 41% more capacity than AMR 12.2, which has the same capacity as AMR-WB 12.65. EVS 7.2 kbps can offer 35% more capacity than AMR 12.65 kbps. Correlating to speech quality results discussed in clause 5.1.1.6.3.1, it can be seen that this significant capacity gain improvement is achieved while maintaining "HD Voice" speech quality.

5.1.1.6.3.3 MBMS bearer

The simulation results in this clause demonstrate EVS and AMR-WB performance on the downlink MBMS bearer and compare this to the reference. The error conditions and scheduler assumptions considered in the MBMS downlink bearer in these simulations, including the definition of Case 1 and Case2, are detailed in clause A.1. The results provided are the output of simulations conducted by one source company using the models described in clause A.1.

Figures 5.1.1.6.3.3-1 (Superwideband) and 5.1.1.6.3.3-2 (Wideband) show speech quality comparison of EVS vs AMR-WB for Case 1. Simulated error conditions in the downlink MBMS bearer channel is given in Table A.1.3-1. In all MBMS bearer scenarios considered, EVS significantly outperforms AMR-WB in terms of voice quality.

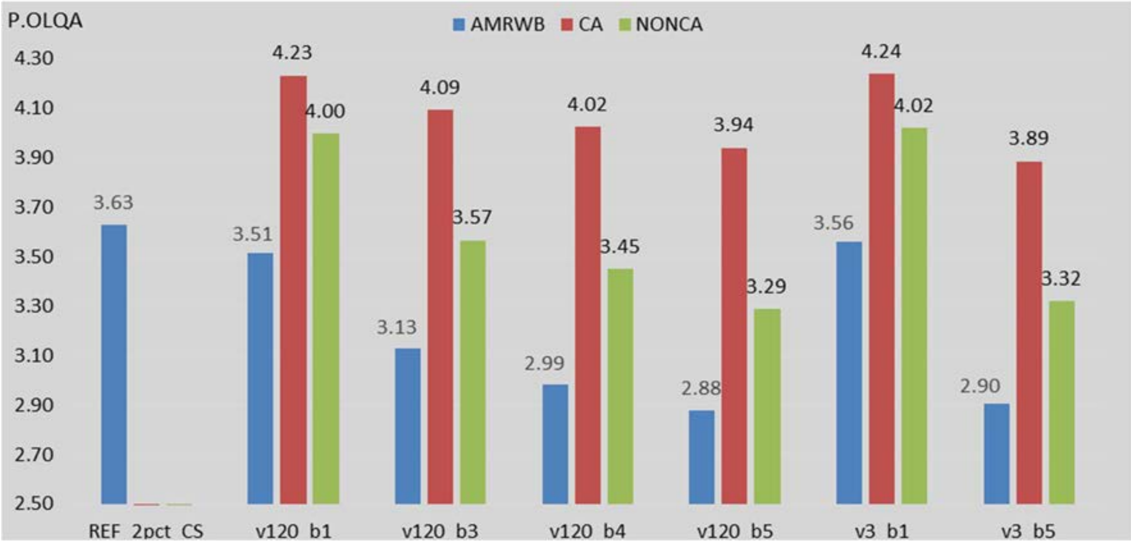


Figure 5.1.1.6.3.3-1: P.OLQA scores for Case 1 – Superwideband Speech

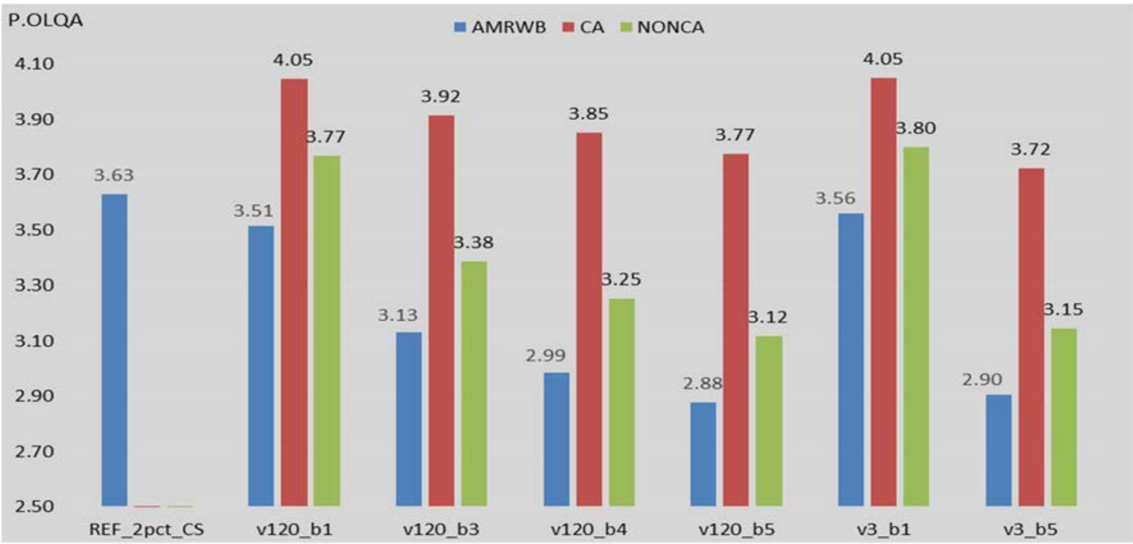


Figure 5.1.1.6.3.3-2: P.OLQA scores for Case 1 – Wideband Speech

Figures 5.1.1.6.3.2-3 and 5.1.1.6.3.2-4 show speech quality for Case 2 described in clause A.1.3. Total uplink and network error levels of 1% FER are considered with various MBMS bearer conditions in the downlink with 40ms scheduling at eNB, and de-jitter buffering at the receiving MCPTT UE.

AMR-WB is unable to meet the reference error resilience, speech quality, and speech intelligibility in all scenarios except for the 1% FER cases. EVS meets or exceeds these reference KPIs in most cases. In addition, both EVS wideband and superwideband modes offer notable improvement in voice quality over AMR-WB under various error conditions considered here.

In certain cases (e.g. v120_b5_4 profile in Figure 5.1.1.6.3.2-3) the improvement in P.OLQA scores by EVS over AMR-WB is 0.65, while AMR-WB yield low P.OLQA scores around 2.84. This may translate into scenarios where an end user may not understand certain portions of speech with AMR-WB under certain FER conditions in the MBMS coverage area, while EVS would provide clear and meaningful speech.

In most test cases considered under Case 1 and Case 2 above, AMR-WB is unable to meet the reference "HD Voice" quality as it suffers from a significant reduction in voice quality during channel errors. From a coverage perspective, this would be experienced as significant degradation in voice quality in areas exhibiting above channel characteristics with errors. Above results also demonstrate how the EVS can compensate for errors significantly better than AMR-WB. Note that the improvement of speech quality over AMRWB offered by EVS Channel Aware mode increases, when more errors are present in the downlink MBMS bearer channel.

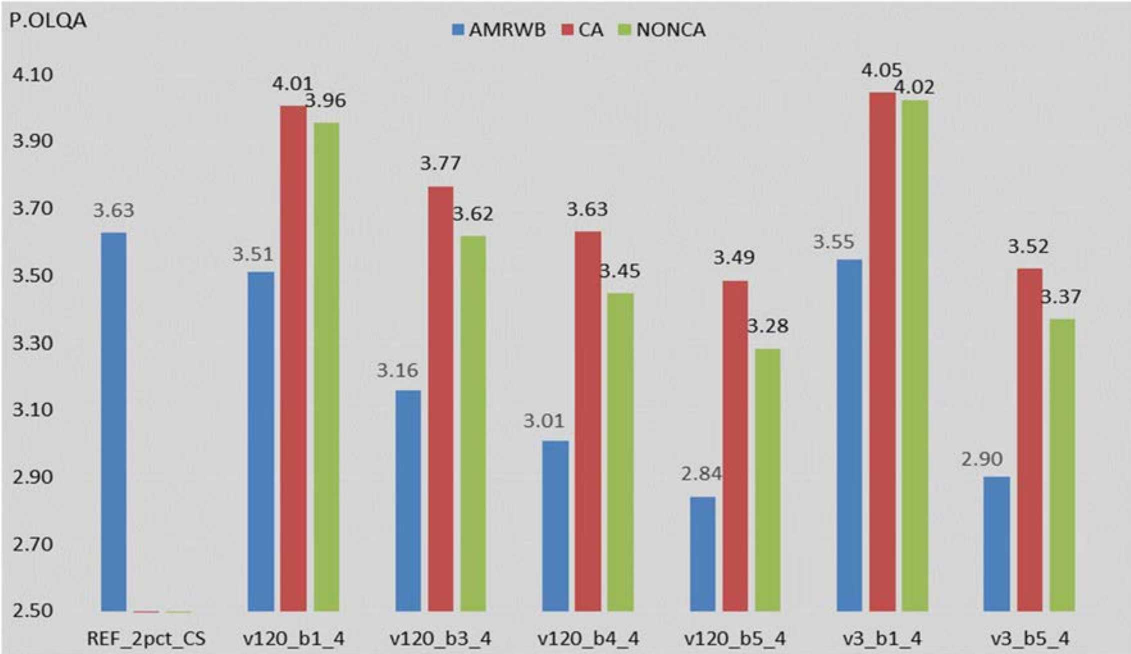


Figure 5.1.1.6.3.3-3: P.OLQA scores for Case 2 – Superwideband Speech

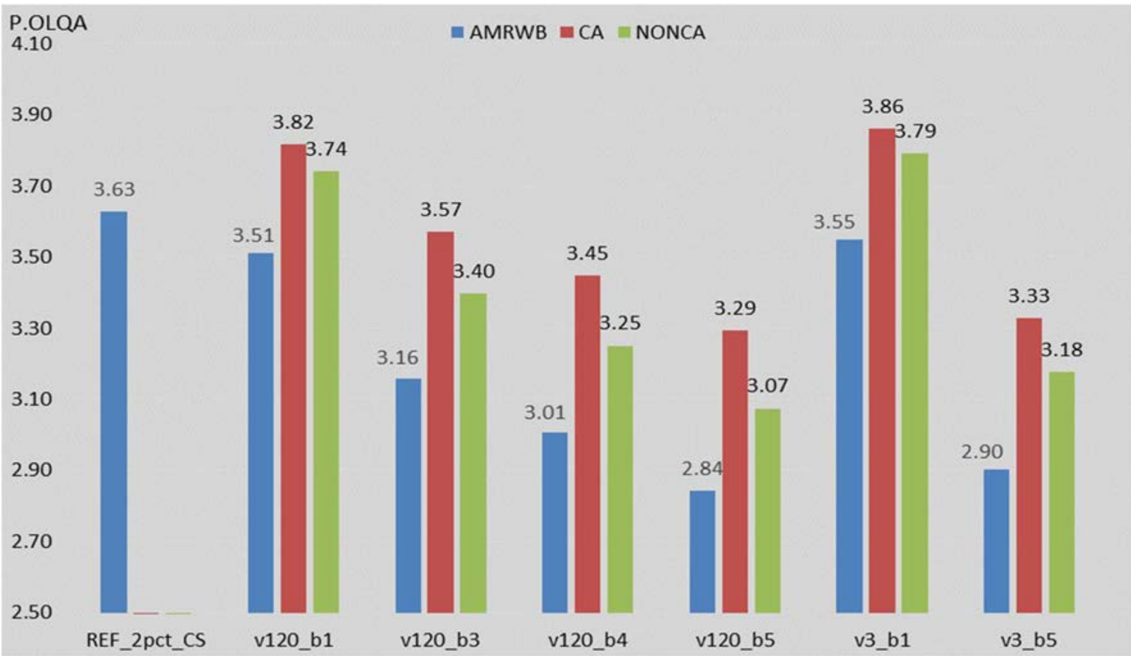


Figure 5.1.1.6.3.3-4: P.OLQA scores for case 2 – Wideband Speech

Error resiliency, Speech Quality, and Speech Intelligibility:

The results in this clause show that EVS can withstand MBMS channel errors much better than AMR-WB. With increasing error rates in the MBMS downlink, AMR-WB speech quality degrades at a faster pace, while EVS is able to maintain a reasonable level of speech quality.

In all of the MBMS bearer scenarios, with the exception of FERs >= 3% in Case 2 wideband, the EVS channel-aware mode is able to meet or exceed the performance of the reference error resiliency, speech quality, and speech intelligibility.

In most of the MBMS bearer scenarios (FER<4% for SWB and FER<3% for WB), EVS non-channel aware mode is able to meet or exceed the performance of the reference.

In all the cases where EVS non-channel aware mode cannot meet the reference, it significantly outperforms AMR-WB. AMR-WB fails to meet the reference in all scenarios except for the 1% FER cases.

Coverage:

Figures 5.1.1.6.3.3-1, 5.1.1.6.3.3-2, 5.1.1.6.3.3-3, and 5.1.1.6.3.3-4 show that EVS channel aware and non-channel aware modes provide significantly better and better coverage than AMR-WB, respectively, due to their ability to handle higher FERs with less degradation in voice quality.

The simulations discussed in clause A.1.1 demonstrated that less than 95% of the cell coverage area can support FER <=1% on the MBMS bearer downlink in SC-PTM when two rings of adjacent cells are only transmitting at 50% interference. Unfortunately, it was not possible to generate more simulation results showing exactly how much cell coverage area provides <=1% FER and how much provides <=5% FER using SC-PTM.

However, it is expected that relaxing the FER target to 5% in the simulations will improve the coverage enough to cover at least 90% of the cell and thus demonstrate that EVS channel aware mode operated over the SC-PTM in the simulated conditions can at least meet the reference coverage.

A comparison of MBMS MBSFN coverage to the reference was not provided as the source did not have simulations that could be used for this analysis.

Call Capacity:

It is not possible to make a reasonable comparison of the capacity of an MBMS bearer with the capacity of an HD Voice unicast bearer system. For example, the number of users supported by the MBMS bearer could be very high when there's a very dense concentration of listeners near the center of the cell.

Nevertheless, the following observations and statements can be made regarding MBMS capacity.

When operating EVS at 13.2 kbps it consumes the same MBMS bearer resources as AMR-WB 12.65kbps. However, due to the higher error-resiliency of EVS, the number of EVS users that can receive the MBMS speech traffic intelligibly is greater than that of AMR-WB users. Therefore EVS provides better call capacity than AMR-WB.

The EVS VBR mode with a lower average rate of 5.9kbps, which provides the same voice quality as AMR-WB 12.65, can also provide an improvement in capacity by using less of the MBMS bearer resources. This would allow the MBMS bearer to carry concurrently more speech streams of EVS VBR than AMR-WB.

5.1.1.6.3.4 LTE-D bearer

In order to study and quantify potential coverage (range extension) and power savings benefits of EVS as compared to AMR-WB, the following parameters of RAN1 VoIP traffic model for D2D group broadcast were adapted to reflect EVS codec modes and characteristics.

- Number of Bytes every 20 ms
- ON time – Markov chain modelling for ON-OFF to transition.
- Outage criteria or BLER target
- All other parameters were kept the same,

Gains to KPIs such as link budget/coverage and fraction of successful links can be realized by either relaxation of target BLER requirements (possible due to the EVS codec's improved error resiliency) or via lower bit-rates (utilizing the EVS codec's improved efficiency). The results provided are the output of simulations conducted by one source company.

The EVS bit-rates and operating FER rates (BLER target) summarized in Table 5.1.1.6.3.3-1 were chosen for RAN1 LTE-D link budget and system simulations. It is important to note each of these EVS configurations provide equivalent speech quality to AMR-WB at 12.65 kbps under 2% FER ("HD Voice center of cell experience), i.e. speech quality has been equalized to this benchmark for purpose of quantifying KPI improvements offered by the EVS modes. In-depth details on audio quality are provided in clause 5.1.1.6.3.1.

Table 5.1.1.6.3.4-1: EVS bit-rates and operating FER rates used for RAN1 LTE-D simulations

Codec and mode	On time	Packet size	BLER target
AMR 12.2 kbps/AMR-WB 12.65 kbps	75%	44 Bytes	2%
EVS 13.2 kbps channel aware mode (Option 1)	72.5%	44 Bytes	8%
EVS 13.2 kbps channel aware mode (Option 2)	66.5%	44 Bytes	2%
EVS 13.2 kbps non channel aware mode (Option 1)	72.5%	44 Bytes	4%
EVS 13.2 kbps non channel aware mode (Option 2)	70.5%	44 Bytes	2%
EVS WB VBR 5.9 or 7.2 or 8 kbps	72.5%	31 Bytes	2%
NOTE 1: Option 1 is RX side BLER relaxation. Option 2 is TX side relaxation where 6% of the packets are dropped at the transmitter while keeping the same RX BLER target of 2% (net FER is 8%) for the channel aware mode. For non-channel aware mode, 2% of packets are dropped at the transmitter and RX BLER target is kept at 2% (net FER is 4%). NOTE 2: The EVS-VBR (variable bit-rate) mode combines bit-rates of 2.8 kbps, 7.2 kbps and 8 kbps to achieve an average bit rate of 5.9 kbps over active speech. For purpose of this simulation, 2.8 kbps and 7.2 kbps packets were zero padded and sent at the same payload size as 8 kbps, i.e. 31 Bytes.			

The input speech database was chosen such that the resulting Voice Activity Factor on encoding using the AMR-12.2 kbps or AMR-WB 12.65 kbps codec (used in current VoIP traffic model) is 75%. Note that EVS Voice Activity Factor is lower by 2.5 % as compared to AMR-12.2/AMR-WB 12.65 kbps.

RAN1 LTE-D simulation assumptions

(a) Link level assumptions

Parameter	Value
Channel Model	EPA
Doppler	10 Hz
Number of PRBs	2
Number of transmissions/packet	4

(b) System level assumptions

Parameter	Value
Drop/channel model	Option 5 Hotspot
TX Power	31 dBm
Resource pool	Dedicated 10 MHz
Resource allocation	Mode 2

(c) RAN1 power consumption model

- Sleep power = 0.01 unit per sub-frame
- RX Power = 1 unit per sub-frame
- TX power
    - 20 unit per sub-frame for 31 dBm
    - 1 unit per sub-frame for 0 dBm and below
    - Linearly scaled with transmit power between 1mW and 10^3.1mW

Link level comparison and coverage gains offered by EVS modes

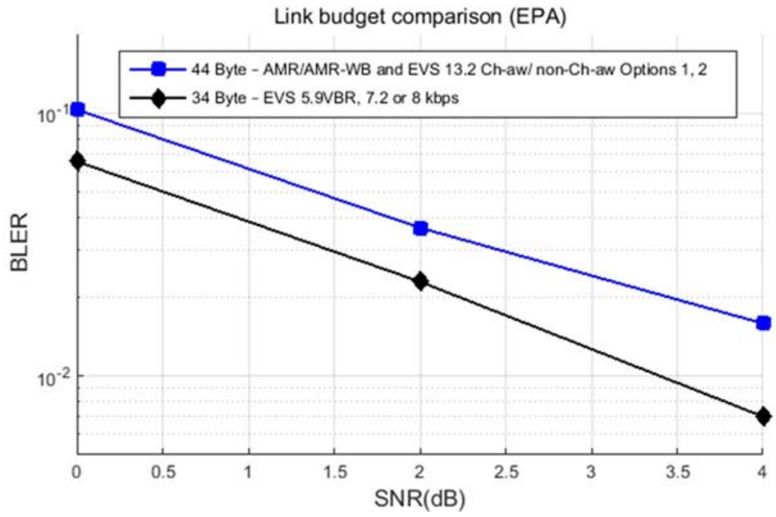


Figure 5.1.1.6.3.4-1: Link budget comparison - EPA

Table 5.1.1.6.3.4-2: Link budget gains and coverage/range extension offered by EVS modes

Model	BLER Target	Gain	Distance Gain	Area Gain
AMR 12.2 kbps/AMR-WB 12.65 kbps	2%	-	-	-
EVS – 13.2 kbps channel aware mode (Option 1)	8%	2.8 dB	17%	38%
EVS – 13.2 kbps channel aware mode (Option 2)	2%	0 dB	-	-
EVS 13.2 kbps non channel aware mode (Option 1)	4%	1.9 dB	11%	24%
EVS – 13.2 kbps non channel aware mode (Option 2)	2%	0 dB	-	-
EVS WB VBR 5.9 or 7.2 or 8 kbps	2%	1.4 dB	8%	17%

The amount of gain in coverage will depend on the scenario. An exponent of 4 in path loss (used for O2O LOS and NLOS) was assumed in translation of link budget gains to coverage gain (distance and area).

Power Consumption comparison of EVS modes to AMR/AMR-WB

Table 5.1.1.6.3.4-3 summarizes the transmit (TX) and receive (RX) side power gains offered by the EVS modes. For the 13.2 kbps channel aware mode Option 2, it is assumed that the RX on time is not reduced as the receiver would keep decoding in the absence of information regarding number of packets being sent (as per Rel-12).

Table 5.1.1.6.3.4-3: Power consumption gains offered by EVS modes

Model	TX Power	TX Power Gain	RX Power	RX Power Gain
AMR 12.2 kbps/AMR-WB 12.65 kbps	3.01 units	-	0.16 units	-
EVS – 13.2 kbps channel aware and non-channel aware modes (Option 1)	2.91 units	3.3%	0.155 units	3.1%
EVS – 13.2 kbps channel aware mode (Option 2)	2.67 units	11%	0.155 units	3.1%
EVS 13.2 kbps non channel aware mode (Option 2)	2.83 units	6%	0.155 units	3.1%
EVS WB VBR 5.9 or 7.2 or 8 kbps	2.91 units	3.3%	0.155 units	3.1%

Example computation:

AMR TX Power = 20 (TX Power) * 0.75 (On time) * 4/20 (Transmissions frequency) + 0.01 (sleep power)

AMR RX Power = 1 (RX Power) * 0.75 (On time) * 4/20 (Reception frequency) + 0.01 (sleep power)

Comparison of Call Capacity measured in terms of fraction of successful links (system level simulation)

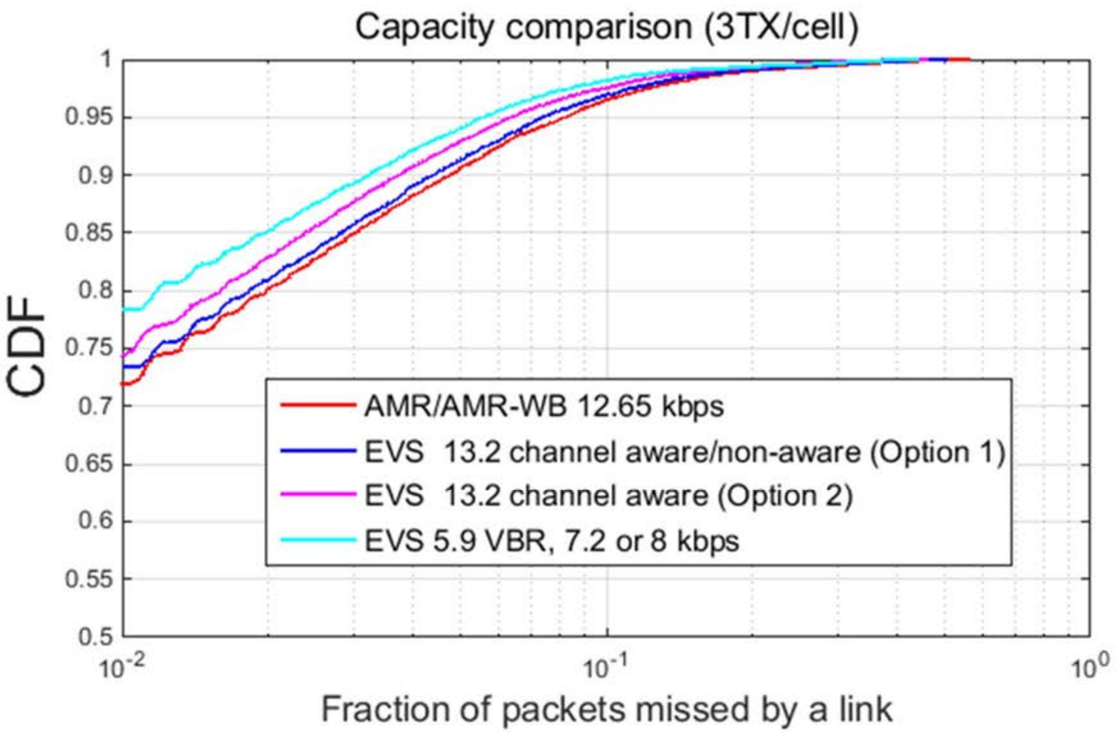


Figure 5.1.1.6.3.4-2: System capacity comparison (3TX/cell)

The capacity comparison plot shown above presents the results of system level simulation keeping the same number of TX (3) per cell for both AMR/AMR-WB and the different EVS modes. Capacity is measured in terms of the fraction of successful links, i.e. the fraction of links where fraction of packets missed by a link <=Target BLER.

The Table 5.1.1.6.3.4-4 below summarizes the gain measured in terms of fractional of successful links.

Table 5.1.1.6.3.4-4: Increase in fraction of successful links offered by EVS modes (3TX/cell)

Model	BLER Target	Fraction of successful links	Gain
AMR 12.2 kbps/AMR-WB 12.65 kbps	2%	80%	-
EVS – 13.2 kbps channel aware mode (Option 1)	8%	96%	20%
EVS – 13.2 kbps channel aware mode (Option 2)	2%	83%	3.7%
EVS 13.2 kbps non channel aware mode (Option 1)	4%	89%	11%
EVS WB VBR 5.9 or 7.2 or 8 kbps	2%	85%	6.2%

The system level simulation to study capacity gains of the EVS modes over AMR/AMR-WB was repeated after increasing the number of TX from 3 to 4 per cell. As shown in Table 5.1.1.6.3.4-5, this results in a lower fraction of successful links (RX side) when compared to the corresponding entry in Table 5.1.1.6.3.4-4, but results in higher gains when TX and RX are combined.

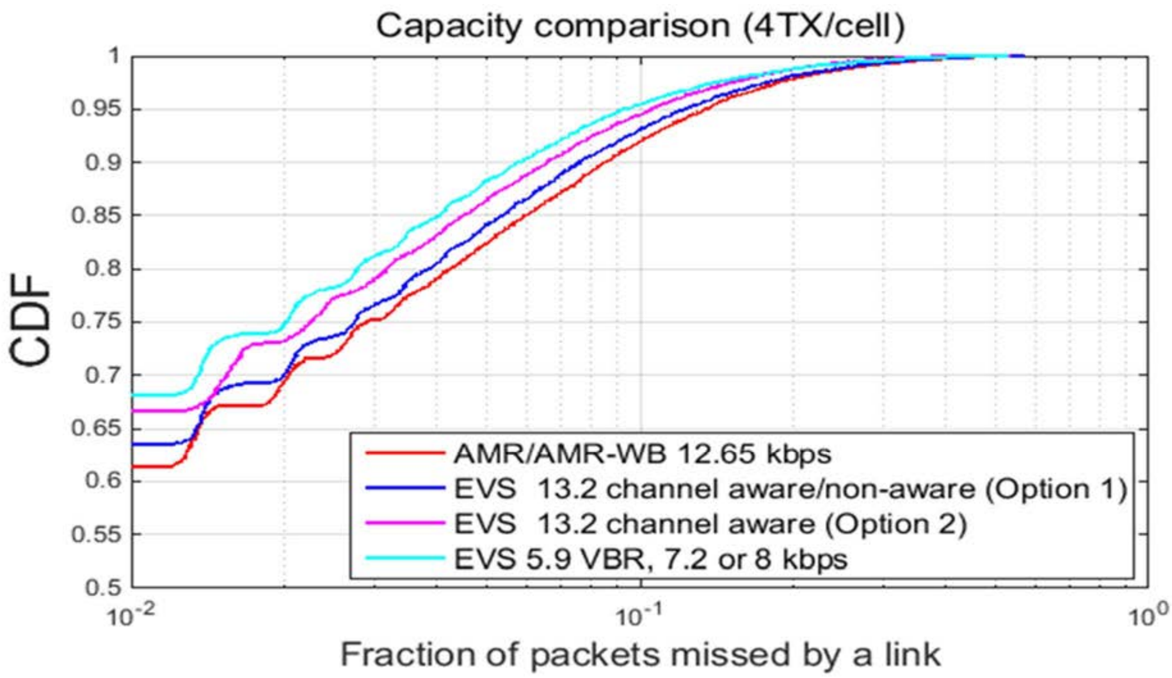


Figure 5.1.1.6.3.4-3: System capacity comparison (4TX/cell)

Table 5.1.1.6.3.4-5: Increase in fraction of successful links offered by EVS modes (4TX/cell).

Model	BLER Target	Fraction of successful links	Gain over AMR/AMR-WB 4TX/cell	Gain over AMR/AMR-WB 3TX/cell (TX and RX combined)
AMR 12.2 kbps/AMR-WB 12.65 kbps	2%	70%		16%
EVS – 13.2 kbps channel aware mode (Option 1)	8%	91%	30%	51.6%
EVS – 13.2 kbps channel aware mode (Option 2)	2%	74%	5.7%	23.3%
EVS 13.2 kbps non channel aware mode (Option 1)	4%	81%	15.7%	35%
EVS WB VBR 5.9 or 7.2 or 8 kbps	2%	75%	7%	25%

Error Resiliency and Coverage:

EVS 13.2 kbps channel aware mode (Option 1) offers the highest gain (over AMR 12.2kbps/AMR-WB 12.65 kbps) in terms of link budget (2.8 dB) and coverage (17% gain in distance and 38% gain in area) followed by the 13.2 kbps non channel aware mode (Option 1) and 5.9 VBR/7.2/8 kbps low bit-rate WB modes.

As seen in Figure 5.1.1.6.3.4-2, for AMR-WB 12.65 kbps case (red curve), only 80% of links have BLER <= 2% i.e. are at HD-Voice center of cell speech quality. The 20% of links where BLER >2% may be classified as regions of bad coverage which will experience significantly lower speech quality than the center of the cell. With EVS 13.2 kbps channel aware mode (Option 1) or non-channel aware mode (Option 1), the fraction of links which will experience "HD Voice" center of cell speech quality, increases to 96% or 89% respectively which roughly speaking brings EVS over LTE-D coverage closer to that of the reference "HD-Voice" coverage. The respective SWB modes will further increase the fraction of links which experience "HD Voice" center of cell speech quality.

Speech Quality and Intelligibility:

EVS offers significantly improved voice quality over "HD Voice" center of cell experience for 80% of links which experience BLER <=2% (Figure 5.1.1.6.3.4-2). The conclusions drawn in clause 5.1.1.6.3.2 (Speech Quality and Intelligibility on LTE unicast bearer) also apply to this scenario. Techniques described in this paper such as increasing the target BLER or lowering of bit-rate can be applied to AMR-WB to realize range/coverage extension. However this will result in significantly degraded speech quality and intelligibility failing the "HD Voice" quality requirement.

Power Consumption: The EVS modes offer power consumption gains over AMR/AMR-WB due reduced ON time. The 13.2 kbps EVS channel aware mode (Option 2) can achieve 8.5% reduction in the ON time when compared to AMR/AMR-WB, which translates to 11% TX Power gain followed by the non-channel aware mode (Option 2), channel aware/non-channel aware modes (Option 1) and 5.9 VBR/7.2/8 kbps low bit-rate WB modes.

Fraction of successful links (Call Capacity):

It is not possible to make a reasonable comparison of the capacity of an LTE-D bearer with the capacity of an HD Voice unicast bearer system. For example, the capacity of LTE-D could be very high when there's a very dense concentration of listeners near the talker.

Nevertheless, the following observations and statements can be made regarding LTE-D capacity.

The EVS 13.2 kbps channel aware mode (Option 1) offers the highest gain in fraction of successful links followed by the non-channel aware mode (Option1) and channel aware mode Option 2/5.9 VBR/7.2/8 kbps low bit-rate WB modes. The ability of EVS, especially in channel-aware mode, to support another transmitter without causing as severe interference to EVS listeners as to AMR-WB listeners in the cell is very relevant in that it allows the cell to support another MCPTT group. This is essentially increasing capacity for the cell to support MCPTT groups by 33% (from 3 to

4) which could be critical in disaster situations where multiple public safety groups (e.g., police, fire-fighters, medical, hazardous materials teams) may be operating concurrently in a particular cell.

The target BLER operating points in Table 5.1.1.6.3.3-1 for the 13.2 kbps EVS channel aware and non-channel aware modes were determined for the WB modes and were based on speech quality results presented in clause 5.1.1.6.3.1. The 13.2 kbps SWB channel aware and non-channel aware modes will provide significantly better quality than the corresponding WB modes. Alternatively, the improvement in voice quality by going to SWB, can be used to further relax the target BLER requirement or reduce the ON time which will translate to even further gains to the corresponding KPIs (coverage, fraction of successful links or power consumption)

5.1.1.6.4 Conclusions

Coverage:

- EVS exceeds the reference coverage for unicast and LTE-D bearers.
- EVS exceeds the coverage of AMR-WB over MBMS SC-PTM bearers. Furthermore, it is expected that EVS should meet, if not exceed, the reference coverage for MBMS SC-PTM bearers.
- AMR-WB meets the reference coverage for unicast bearers
- AMR-WB does not meet the reference coverage for LTE-D bearers.
- Based on the analysis in Clause 5.1.1.6.3.3, AMR-WB does not appear to meet the reference coverage for MBMS SC-PTM bearers.

The simulations shows the coverage improvement benefits using EVS Codec for the unicast and MBMS bearers. Coverage may be improved by deploying more infrastructure, including emergency deployments (e.g. cells on wheels). In this case, the EVS channel aware mode will still provide improved coverage benefits relative to AMR-WB, however, the relative benefit in coverage may need to be re-evaluated.

Speech Quality/Error-resiliency:

- When the MCPTT bearer channel conditions are similar to that of the reference (end-to-end FER <=2%), EVS provides better speech quality than the reference. When the channel conditions get worse, EVS can maintain the same speech quality as the reference for all the MCPTT bearers due to its improved error resiliency.
- AMR-WB is able to provide the reference error-resiliency/speech quality over the unicast bearer.
- AMR-WB is unable to provide the reference error-resiliency/speech quality for both the MBMS SC-PTM and LTE-D bearers.

In this report, the subjective MOS evaluations from 3GPP TS 26.952 contained both clean speech and noise while the POLQA evaluations are based on a clean speech database.

Speech Intelligibility:

- EVS offers significant voice quality improvements over the reference. The improved robustness to background noise and resiliency to errors are particularly relevant to the MCPTT service and in general is expected to result in equal or better speech intelligibly compared to the reference.
- AMR-WB meets the reference speech intelligibility for unicast bearers.
- AMR-WB does not meet the reference speech intelligibility for LTE-D bearers.
- Based on the analysis in Clause 5.1.1.6.3.3, AMR-WB does not appear to meet the reference speech intelligibility for MBMS SC-PTM bearers.

Call Capacity:

- EVS exceeds the reference capacity for unicast bearers.
- AMR-WB meets the reference capacity for unicast bearers.
- For LTE-D and MBMS (SC-PTM and MBSFN) bearers, it is not possible to make a relevant comparison to the reference capacity. However EVS provides better capacity than AMR-WB for all of these MCPTT bearers.

In summary:

- In summary, in the cases where the KPIs and MCPTT bearers allow a comparison to the reference "HD Voice" experience, EVS meets or exceeds (or is expected to at least meet in the case of MBMS SC-PTM coverage) the performance of the reference for all the bearers. For the LTE-D bearers, AMR-WB does not meet the reference requirements, while for the MBMS SC-PTM bearers, AMR-WB does not appear to meet the reference requirements.

5.1.1.7 Listening effort evaluation of AMR-WB and EVS under impaired channels

5.1.1.7.1 Test setup

For the evaluation of the listening effort, a P.800 listening effort test was conducted at Fraunhofer IIS in German language. The P.800 [24] procedure was simply chosen because listening effort evaluation is a key application part of the standard. The scale defined in [24] is given in the following:

5 Complete relaxation possible; no effort required.
4 Attention necessary; no appreciable effort required.
3 Moderate effort required.
2 Considerable effort required.
1 No meaning understood with any feasible effort.

The speech material consisted of six talkers (3 male, 3 female), 4 samples of each talker, where each sample contains two sentences of the Berlin sentence corpus. The clean speech material was combined with the two noise types (coffee and siren), kindly provided by the NTIA over the 3GPP reflector, according to the EVS characterization processing. For concatenation, a long noise sequence was assembled out of the provided short samples. As SNR levels 10dB for coffee noise and 5dB for the siren noise were chosen. Each noise type was evaluated in a separate experiment.

26 conditions were evaluated consisting of 6 codec modes times 4 channel conditions, plus 2 direct condtions. The codecs AMR-WB, EVS-WB, and EVS-SWB have been included operating at the bit rates 13.2kbps (12.85kbps for AMR-WB) and 24.4kbps (23.85kbps for AMR-WB), all using DTX. The 13.2 kbps EVS modes were operated in channel-aware (CA) mode (CA parameters: p=HI, o=3). All conditions have been evaluated using clean channel and error profiles for 5%, 10%, 20% frame loss rate, using the informative EVS-JBM for MTSI [4], specified in [26].

The delay and error profiles have been derived with the tool provided by Qualcomm over the 3GPP SA4 reflector. For the EVS decoder, the JBM correction for low jitter but high loss in [25] was active.

The listening test was conducted in a listening room according to the requirements of P.800 using Sennheiser HD280 pro as the instrument for diotic listening. The randomization of the samples was done according to P.800 using the balanced block design method.

5.1.1.7.2 Test results

The following plots show the mean scores and 95% confidence intervals for the experiment with coffee background noise, where 25 naive subjects participated. Unfortunately, not enough subjects did follow the lab invitation for the siren noise experiment and it was not possible to compensate the missing subjects. As a consequence, the it was decided to omit this experiment. It should be noted that the number of listeners is rather small for a codec selection process, however at least the number of 24 listeners has been reached which is the usual minimum requirement for listening tests in a codec qualification process. It should also be noted that this experiment can only be considered a snapshot of listening effort in a single language. Those results do therefore not claim to be exhaustive and should only be used to evaluate trends.

Given that 25 listeners participated in the test and there were six trials per condition, a total number of 3900 trials were reported, meaning 150 trials per condition.

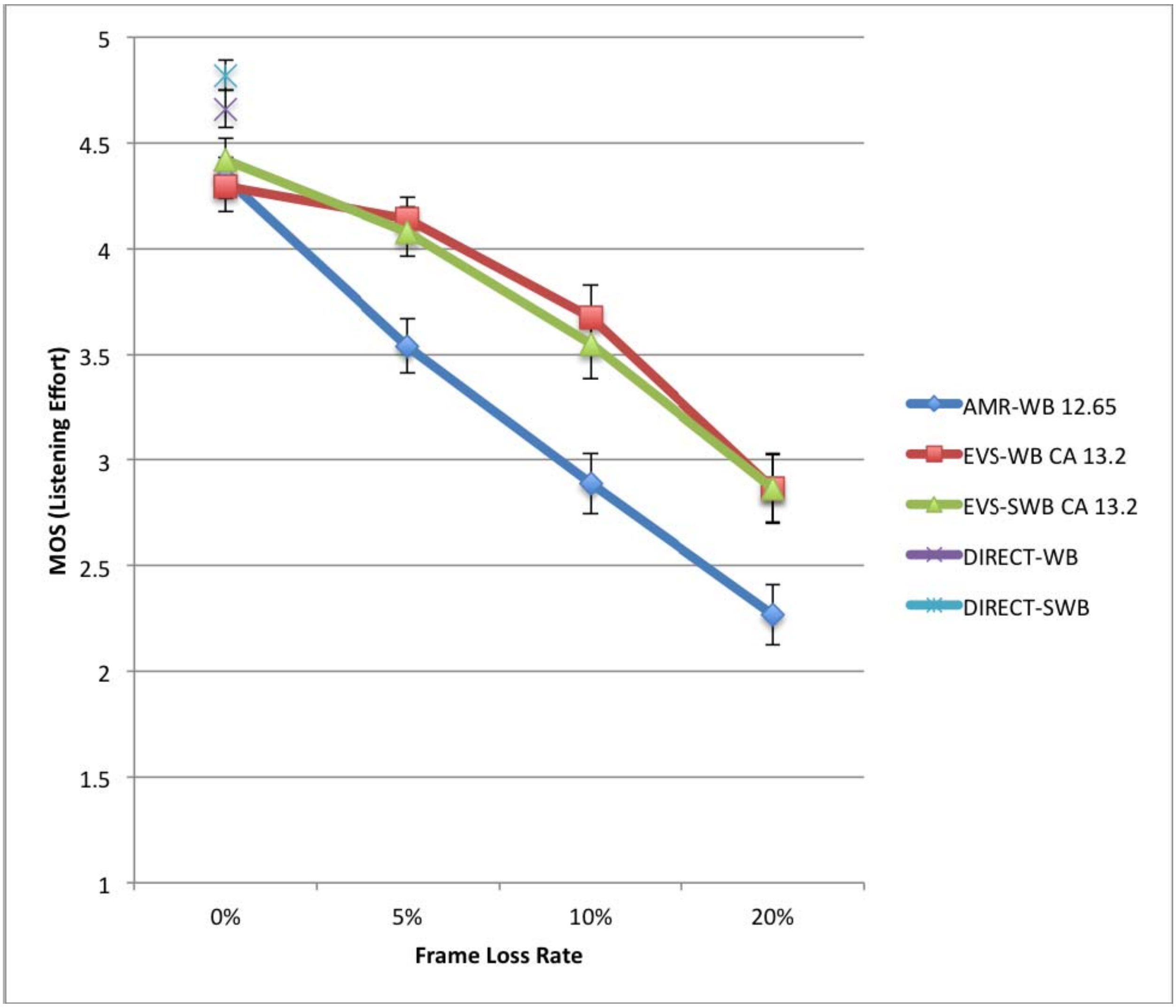


Figure 5.1.1.7.2-1: Plot of listening effort in 10dB SNR coffee background noise for 13.2 kbps gross bit rate incl. AMR-WB, EVS-WB in channel aware mode and EVS-SWB in channel aware mode

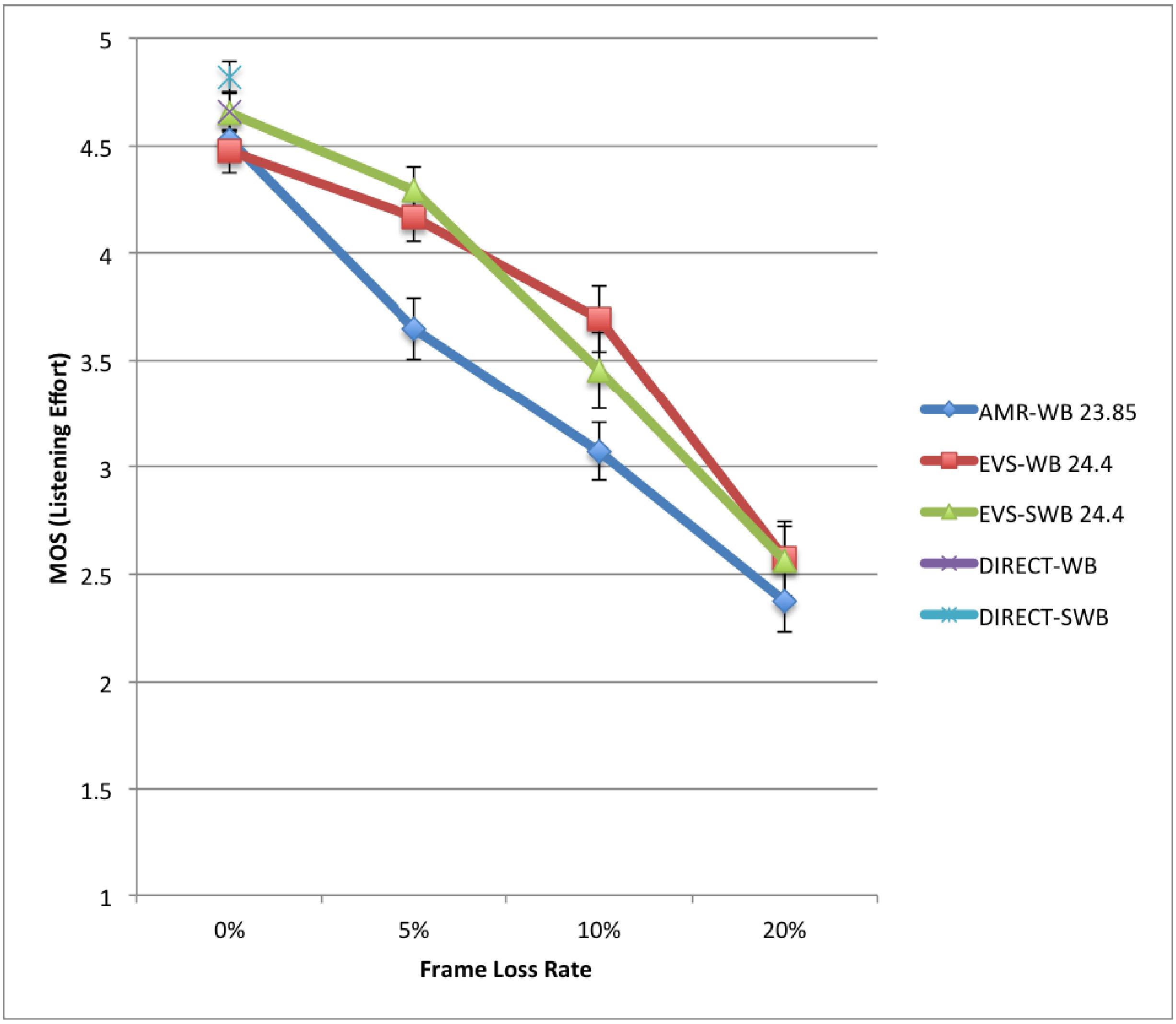


Figure 5.1.1.7.2-2: Plot of listening effort in 10dB SNR coffee background noise for 24.4 kbps gross bit rate incl. AMR-WB, EVS-WB and EVS-SWB

The following points can be observed:

- All error-free conditions show an average score above 4 on the listening effort scale. These operation points seem to guarantee sufficiently low listening effort and are not considered as critical.
- As the experiment design is focusing on high packet loss conditions, conclusions on other influencing factors such as audio bandwidth or bit rate can not be drawn
- For packet loss rates of 5% and 10%, EVS always shows a significantly higher score compared to corresponding AMR-WB conditions
- EVS with 10% packet loss rate performs similar to AMR-WB at 5% packet loss rate
- Using the channel aware mode, EVS@13.2kbps with 20% packet loss rate performs similar to AMR-WB@12.65 with 10% packet loss rate.

5.1.2 Review of the Codec Alternatives and their Relative Speech Intelligibility in Clean and Low SNRs

The high noise intelligibility results in NTIA Report 15-520 [16] provide valuable information for the selection of the codec for MCPTT. For simplicity, operating conditions of codecs are not repeated in the following paragraphs but they can be found in [16].

The report provides intelligibility scores, as measured with the Modified Rhyme Test (MRT) methodology, across a range of noise conditions for different codecs, bit rates and bandwidths but for SA4 only 3 codecs are really of direct interest; namely AMR (Sometimes referred to as AMR NB in [16]), AMR-WB and EVS although the benchmarks of Analogue FM and the P25 codec are of interest. These intelligibility results were conducted without radio channel impairments.

In clean channel conditions, in two cases out of 12 EVS-NB performed statistically significantly better than AMR but in the other 10 cases there was no statistically significant difference.

In one case out of 18, EVS-WB performed statistically worse than AMR-WB, but in this case EVS was operating at its lowest bit rate with DTX/CNG enabled and at a lower bit rate than AMR-WB. In the other 17 cases there was no statistically significant difference.

In three cases out of 24, EVS-FB performed statistically significantly better than AMR-WB or EVS-WB, in one case EVS-FB performed statistically worse than AMR-WB (Alarm noise -30dB SNR) and in the other 20 cases there was no statistically significant difference.

- From the NTIA report [16] using a WB codec results in higher intelligibility than using a NB codec in noisy conditions.
- In the NTIA report [16] for SNRs within the range 10dB to -5dB, FB is always equivalent to or better than WB and NB in noisy conditions from an intelligibility point of view.
- The NTIA report [16] shows that the intelligibility increases (up to a saturation level) with coded bitrate within confidence limits for a given audio bandwidth.

5.1.2.1 Speech Intelligibility

This Clause presents results from a subjective intelligibility test, P.INTELL (as specified in ITU.T P.807 [27]), conducted by an independent laboratory Dynastat, Inc.

P.INTELL - Method and Procedure

The P.INTELL test is designed to evaluate the Speech Intelligibility of eight test conditions. The experimental design of P.INTELL is based on the partially-balanced/randomized-blocks experimental design that has been used in most codec standardization testing efforts in the past decade for Speech Quality tests, i.e., tests described in ITU-T Rec. P.800 [24]. The partially-balanced/randomized-blocks experimental design is described and recommended in the ITU-T Handbook of Subjective Testing Practical Procedures.

The test parameters for the P.INTELL test included:

- Eight test conditions
- Four talkers - two males and two females
- Four samples per talker, where "sample" is described in the next section
- 32 subjects - four panels of eight naive subjects, each panel with an independent randomization of the speech materials

Structure of the P.INTELL Test

The P.INTELL source speech database includes 96 items, where each item is a pair of single-syllable English words. For half of the items, the words differ only in the initial consonant, i.e., rhyming word pairs. For the other half, the words differ only in the final consonant, i.e., alliterative word pairs. The critical consonants in both the rhyming and alliterative test-items differ only in a single Distinctive Feature, either *Voicing*, *Nasality*, *Sustention*, *Sibilation*, *Graveness* or *Compactness*. In P.INTELL a "sample" results from a single subject selecting one of two words for each of the 96 presentation pairs. Table 5.1.2.1-0 shows the P.INTELL test items with the six distinctive features, four samples, and four vowel/consonant start-stop characteristic.

Table 5.1.2.1-0. P.INTELL test items

Distinctive Feature	Sample-1				Sample-2				Sample-3				Sample-4			
	VWL	CONS	Present	Absent	VWL	CONS	Present	Absent	VWL	CONS	Present	Absent	VWL	CONS	Present	Absent
Voicing	ɑ	b/p (i)	BOND	POND	o	v/f (i)	VOLE	FOAL	æ	g/k (i)	GAFF	CALF	ɛ	d/t (i)	DENSE	TENSE
	u	z/s (i)	ZOO	SUE	ɔ	d/t (i)	DAUNT	TAUNT	ɪ	dʒ/tʃ (i)	GIN	CHIN	i	b/p (i)	BEAN	PEEN
	ɛ	v/f (f)	REV	REF	ɪ	dʒ/tʃ (f)	RIDGE	RICH	o	g/k (f)	BROGUE	BROKE	ɑ	dʒ/tʃ (f)	HODGE	HOTCH
	i	v/f (f)	SHEAVE	SHEAF	æ	g/k (f)	BAG	BACK	ɔ	z/s (f)	LAWS	LOSS	u	b/p (f)	LUBE	LOOP
Nasality	ɛ	n/d (i)	NECK	DECK	ɑ	n/d (i)	KNOCK	DOCK	ɔ	m/b (i)	MOSS	BOSS	æ	m/b (i)	MAD	BAD
	i	m/b (i)	MEAT	BEAT	u	m/b (i)	MOOT	BOOT	u	n/d (i)	NOTE	DOTE	ɪ	n/d (i)	NIP	DIP
	ɑ	m/b (f)	BOMB	BOB	ɛ	m/b (f)	GEM	JEB	æ	n/d (f)	FAN	FAD	ɔ	n/d (f)	BRAWN	BROAD
	u	n/d (f)	NOON	NUDE	i	n/d (f)	SCREEN	SCREED	ɪ	m/b (f)	RIM	RIB	u	n/d (f)	MOAN	MODE
Sustention	æ	ð/d (i)	THAN	DAN	ɛ	f/p (i)	FENCE	PENCE	ɑ	v/b (i)	VOX	BOX	ɔ	θ/t (i)	THONG	TONG
	ɪ	θ/t (i)	THICK	TICK	i	ʃ/tʃ (i)	SHEET	CHEAT	u	ʃ/tʃ (i)	SHOES	CHOOSE	u	ð/d (i)	THOUGH	DOUGH
	ɔ	f/p (f)	GOFF	GAWP	ɑ	v/b (f)	SLAV	SLOB	i	ð/d (f)	SEETHE	SEED	æ	f/p (f)	CALF	CAP
	u	ð/d (f)	LOATHE	LOAD	u	f/p (f)	GOOF	GOOP	ɛ	ʃ/tʃ (f)	FLESH	FLETCH	ɪ	v/b (f)	LIVE	LIB
Sibilation	ɔ	dʒ/g (i)	JAWS	GAUZE	æ	s/θ (i)	SANK	THANK	ɛ	tʃ/k (i)	CHAIR	CARE	ɑ	tʃ/k (i)	CHOP	COP
	o	dʒ/g (i)	JOE	GO	ɪ	dʒ/g (i)	JILT	GUILT	i	z/ð (i)	ZEE	THEE	u	tʃ/k (i)	CHEW	COO
	æ	tʃ/k (f)	PATCH	PACK	ɔ	s/θ (f)	ROSS	WROTH	ɑ	tʃ/k (f)	NOTCH	KNOCK	ɛ	dʒ/g (f)	EDGE	EGG
	ɪ	s/θ (f)	MISS	MYTH	u	s/θ (f)	GROSS	GROWTH	u	z/ð (f)	SUES	SOOTHE	i	z/ð (f)	BREEZE	BREATHE
Graveness	i	p/t (i)	PEAK	TEAK	ɪ	f/θ (i)	FIN	THIN	u	f/θ (i)	FORE	THOR	u	p/t (i)	POOL	TOOL
	ɛ	m/n (i)	MET	NET	æ	b/d (i)	BANK	DANK	ɔ	b/d (i)	BONG	DONG	ɒ	w/r (i)	WAD	ROD
	u	m/n (f)	LOOM	LOON	u	b/d (f)	STROBE	STRODE	ɪ	m/n (f)	SHIM	SHIN	i	f/θ (f)	REEF	WREATH
	ɑ	p/t (f)	HOP	HOT	ɔ	f/θ (f)	TROUGH	TROTH	æ	p/t (f)	RAP	RAT	ɛ	b/d (f)	WEB	WED
Compactness	ɪ	h/f (i)	HIT	FIT	u	ʃ/s (i)	SHOW	SO	u	j/r (i)	YOU	RUE	ɛ	k/p (i)	KEG	PEG
	æ	g/b (i)	GAT	BAT	ɔ	k/t (i)	CAUGHT	TAUGHT	ɒ	g/d (i)	GOT	DOT	i	j/w (i)	YIELD	WIELD
	u	k/t (f)	OAK	OAT	ɪ	g/b (f)	BIG	BIB	ɛ	g/d (f)	BEG	BED	ɑ	g/b (f)	NOG	KNOB
	ɔ	g/d (f)	FLOG	FLAWED	æ	ʃ/s (f)	CLASH	CLASS	i	k/p (f)	SEEK	SEEP	u	k/p (f)	DUKE	DUPE

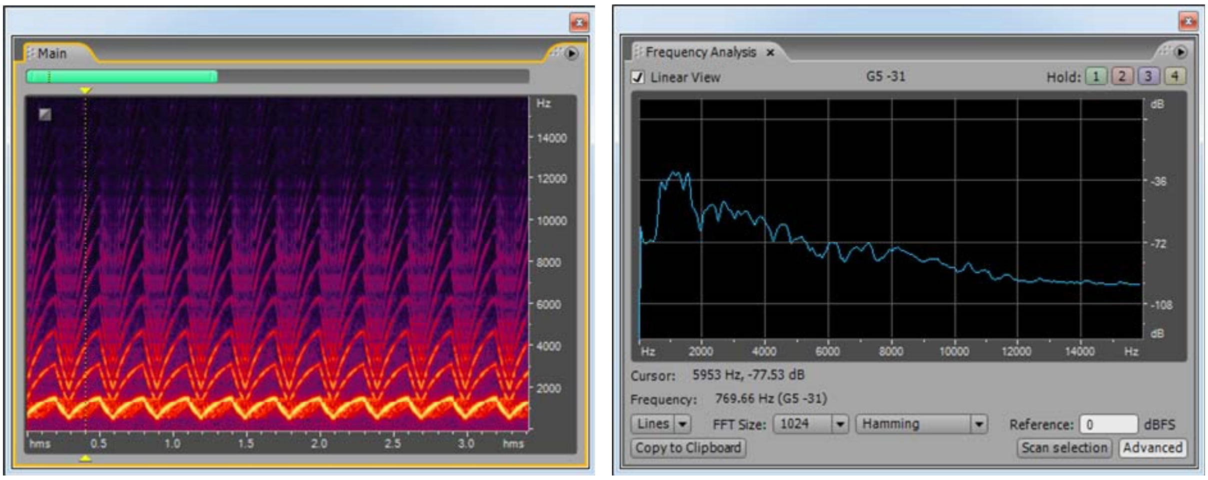
Test Conditions

Table 5.1.2.1-1 shows the test conditions list to evaluate the speech intelligibility of the two codecs AMR-WB and the EVS-SWB CA. The AMR-WB at 12.65 kbps and EVS-SWB CA at 13.2 kbps are tested at the frame erasure rates of 2%, 8%, and 20%. The test also included the original un-coded signal.

Table 5.1.2.1-1. Test conditions list

	Condition	Bit Rate (kb/s)	CA FEC-Offset	DTX	Uplink	Downlink	Noise	SNR
C1	Un-coded with noise	-	-	-	-	0%	Siren	5dB
C2	AMRWB 2%FER	12.65	-	on	MTSI 2	2%	Siren	5dB
C3	AMRWB 8%FER	12.65	-	on	MTSI 2	8%	Siren	5dB
C4	AMRWB 20%FER	12.65	-	on	MTSI 2	20%	Siren	5dB
C5	EVS-SWB-CA 2% FER	13.2	3	on	MTSI 2	2%	Siren	5dB
C6	EVS-SWB-CA 8% FER	13.2	3	on	MTSI 2	8%	Siren	5dB
C7	EVS-SWB-CA 20% FER	13.2	3	on	MTSI 2	20%	Siren	5dB

The intelligibility test was conducted in high noise and in impaired channels. A background noise that mimics that of US police car Siren is mixed with the input source at 5 dB. The Siren noise characteristic is as shown in Figure 5.1.2.1-1 below (which has most of the noise energy around 500-1500 Hz). The following procedure is used for noise mixing (similar to that of the steps used in 3GPP tests): The input source is normalized to -26 dBov, and noise is scaled such that the Siren noise loudness is at -31 dBov. The normalized input source and the scaled noise is mixed at SNR of 5 dB.



(a) (b)

Figure 5.1.2.1-1. An example Siren noise characteristic (a) spectrogram, (b) LT frequency response

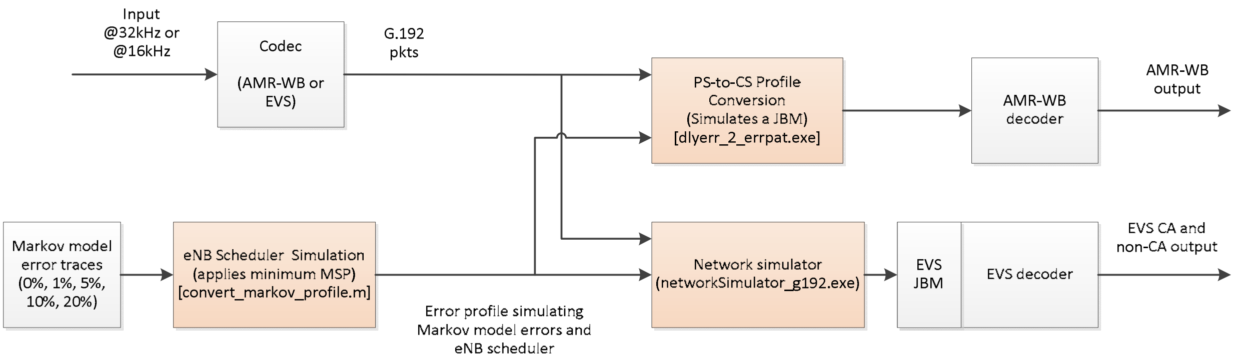


Figure 5.1.2.1-2. Steps to simulate impaired channel conditions

Figure 5.1.2.1-2 shows the steps to simulate the impaired channel conditions that combines the Markov model based error traces with the VoLTE uplink jitter and the MBMS/LTE-D downlink scheduler jitter. The steps are further elaborated below.

Step 1: Apply Markov channel models (Table A.1.2-3 in Annex A.1.2) to produce the error traces.

Step 2: Simulate the eNB scheduling procedures described in Case1 in clause A.1.3. This models the minimum amount of jitter that would be introduced in both the MBMS and LTE-D bearers and is caused by the MSP for MBMS (see clause 5.3.6) and a similar minimum scheduling period for LTE-D bearers. To simulate the VoLTE uplink jitter, the MTSI profile 2 is used.

Step 3: Convert the Markov model error traces to error profile.

Step 4: For EVS, use the *.dat file and the following command lines to generate the decoder output.

For EVS SWB CA @13.2:

```
/* encoder */
evs_cod -dtx -rf 3 13200 32 input.pcm temp.pkt
/* Network simulator */
networkSimulator_g192.exe v3_b10.dat  temp.pkt  temp_voip.pkt data.trace 1 0
/* decoder */
evs_dec -voip -tracefile temp 32 temp_voip.pkt   output.pcm
```

Step 5: For AMR-WB, the procedure is the same through the eNB scheduler simulation. Then, since no JBM is available for AMR-WB, we used the 3GPP utility function (dlyerr_2_errpat.exe) made available by SA4 that simulates jitter-buffer handling, resulting in a circuit-switched packet file that can be passed to the normal AMR-WB decoder.

For AMR-WB @12.65:

```
/* encoder */
amrwb_enc -dtx –itu 2 input.pcm temp.pkt
/* Delay to error pattern */
dlyerr_2_errpat.exe -L 22000 -d 200 -f 1 -w -s 0 -i v3_b10.dat -o   v3_b10.dat
eid-xor.exe -vbr -fer   temp.pkt  v3_b10.dat tempfer.pkt
/* decoder */
amrwb_dec.exe -itu tempfer.pkt output.pcm
```

Low jitter scenario – 3GPP EVS JBM Behavior

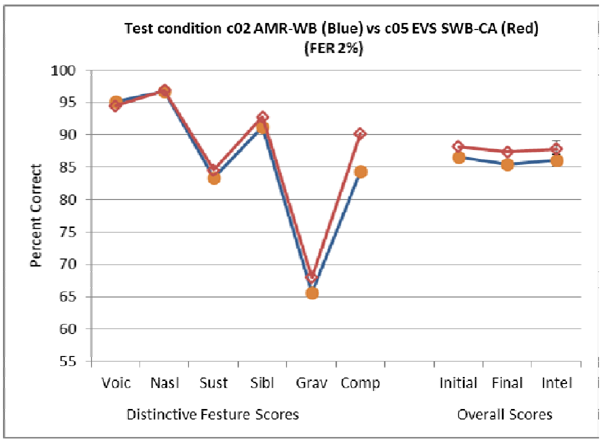
The 3GPP EVS JBM was used for testing in EVS selection and characterization tests. During the tests, the MTSI profiles 1 through 10 were used that are representative of VoLTE and HSPA, covering a wide range of jitter and packet losses. The informative 3GPP EVS JBM was never evaluated for other profiles such as MBMS. While investigating the MBMS downlink packet scheduler, a bug in the 3GPP EVS de-jitter buffer handling was uncovered in case of low jitter scenarios. In particular, the JBM was not exercising the partial copy recovery logic as it was locking up in a low jitter, high FER scenario. This scenario is a direct consequence of using the informative-only 3GPP JBM, which was developed specifically for EVS on VoLTE (not MBMS) using 3GPP VoLTE delay-loss profiles. In Figure **5.1.2.1-2**, we are simulating the test case using an approximation of the jitter based on modelling the MBMS downlink packet scheduler, which triggers the problem. This does not happen for the delay-loss profiles derived from commercial VoLTE field testing. In the subjective evaluation testing that we conducted on Jan 07, 2016 the 3GPP EVS JBM included a bug fix provided by Fraunhofer IIS. A CR that addresses the JBM issue was agreed at the January SA4#87 meeting.

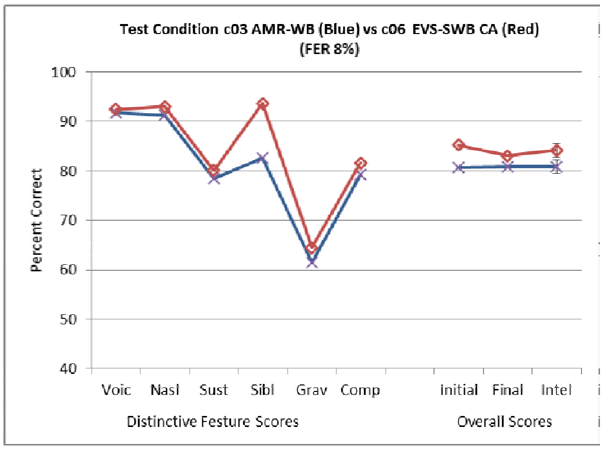
Test Results – Summary

Table 5.1.2.1-2 shows the summary of speech intelligibility test results (Means, Standard Deviations) for each of the test conditions involved in the experiment). Each value shown in the Table is based on 128 samples (32 subjects x 4 talkers). P.INTELL Scores are expressed as Percent Correct. Figure 5.1.2.1-3 shows the P.INTELL profile plots (i.e., Distinctive Feature scores) for the Test conditions c02 vs c05 FER 2%, c03 vs c06 FER 8%, and c04 vs c07 FER 20%.

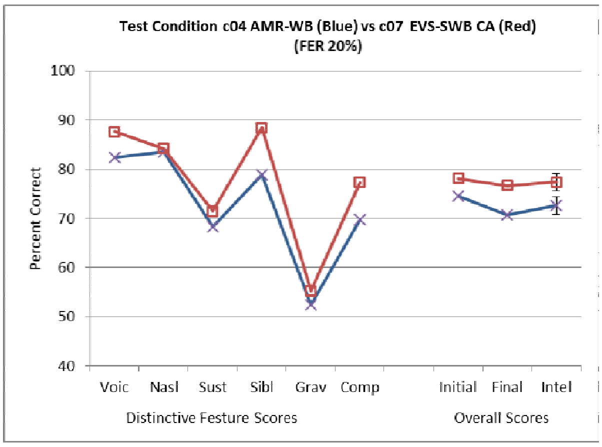
Table 5.1.2.1-2. Test Results - summary

	Condition	Bit Rate (kb/s)	Mean Intel	StdDev
C1	Uncoded (with noise)	-	92.87	5.54
C2	AMRWB 2%FER	12.65	86.07	7.25
C3	AMRWB 8%FER	12.65	80.83	8.09
C4	AMRWB 20%FER	12.65	72.59	10.69
C5	EVS-SWB-CA 2% FER	13.2	87.83	6.85
C6	EVS-SWB-CA 8% FER	13.2	84.18	8.12
C7	EVS-SWB-CA 20% FER	13.2	77.41	10.31





(a) (b)



(c)

Figure. 5.1.2.1-3 P.INTELL profile plots for test conditions (Red: EVS-SWB CA, blue: AMR-WB) (a) c02 vs c05 FER 2%, (b) c03 vs c06 FER 8%, and (c) c04 vs c07 FER 20%.

Statistical Analysis (AMR-WB vs EVS SWB codec)

Figure 5.1.2.1-4 shows the P.INTELL speech intelligibility scores for AMR-WB and EVS SWB-CA at FERs 2%, 8%, and 20%. It can be noted that,

- At a given FER, the EVS SWB CA is "statistically significantly better than (BT)" AMR-WB (as further elaborated below).
- EVS-SWB CA at FER 8% is "statistically no worse than (NWT)" AMR-WB at FER 2% (as further elaborated below).
- AMR-WB at FER 8% MBMS bearer is "statistically worse than" AMR-WB at FER 2% (as further elaborated below).

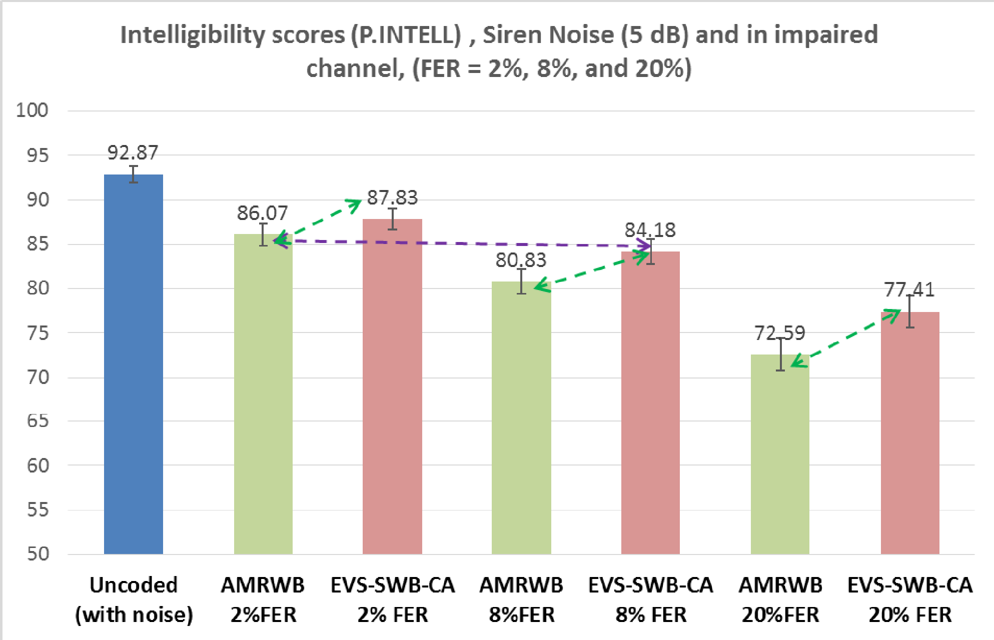


Figure 5.1.2.1-4 Statistical analysis

AMR-WB vs EVS-SWB CA at a given FER

Table 5.1.2.1-3 presents the statistical significance test results comparing AMR-WB at FER 2%, 8%, and 20% vs EVS-SWB CA at FER 2%, 8%, and 20%. T-stat is estimated as (mean_ref – mean_cut)/SEmd, where SEmd = sqrt((stddev_ref^2+stddev_cut^2)/N). The t-scores for a two-sided significance test at 95% confidence interval (CI) is, $t_{(.05,254)}$=1.97. As shown in Table 5.1.2.1-3, for the three FERs under test, 2%, 8% and 20%, the EVS-SWB CA is "statistically significantly better than (BT)" the AMR-WB.

Table 5.1.2.1-3. Statistical significance test (two sided t-test at 95% CI)

AMR-WB at 12.65 kbps vs EVS SWB-CA at 13.2 kbps (at FERs 2%, 8%, and 20%)

Ref (AMR-WB @12.65 kb/s)	Mean (ref)	Std. Dev (ref)	CuT (EVS-SWB-CA @13.2 kb/s)	Mean (CuT)	Std. Dev (CuT)	T-stat (N=128)	CuT vs Ref
c2 (2% FER)	86.07	7.25	c5 (2% FER)	87.83	6.85	1.996	C5 **"BT"** C2
c3 (8% FER)	80.83	8.09	c6 (8% FER)	84.18	8.12	3.307	C6 **"BT"** C3
c4 (20% FER)	72.59	10.69	c7 (20% FER)	77.41	10.31	3.671	C7 **"BT"** C4

AMR-WB at FER 2% vs EVS-SWB CA at FER 8%

Table 5.1.2.1-4 presents the statistical significance test results comparing AMR-WB at FER 2% vs EVS-SWB CA at FER 8%. As shown in Table 4, the EVS-SWB CA at FER 8% is "statistically no worse than (NWT)" the AMR-WB at FER 2%.

Table 5.1.2.1-4. Statistical significance test (two sided t-test at 95% CI)

AMR-WB at 12.65 kbps at FER 2% vs EVS SWB-CA at 13.2 kbps at FER 8%

Ref (AMR-WB @12.65 kb/s)	Mean (ref)	Std (ref)	CuT (EVS-SWB-CA @13.2 kb/s)	Mean (cut)	Std (cut)	T-stat (N=128)	CuT vs Ref
c2 (2% FER)	86.07	7.25	c6 (8% FER)	84.18	8.12	-1.964	**c6 "NWT" c2**

AMR-WB at FER 2% vs 8%

Table 5.1.2.1-5 presents the statistical significance test results comparing AMR-WB at FER 2% vs AMR-WB at FER 8%. As shown in Table 5, the AMR-WB at FER 8% is "statistically worse than (WT)" the AMR-WB at FER 2%.

Table 5.1.2.1-5. Statistical significance test (two sided t-test at 95% CI)

AMR-WB at 12.65 kbps at FER 2% vs FER 8%

Ref (AMR-WB @12.65 kb/s)	Mean (ref)	Std (ref)	CuT (AMR-WB @12.65 kb/s)	Mean (cut)	Std (cut)	T-stat (N=128)	CuT vs Ref
c2 (2% FER)	86.07	7.25	c3 (8% FER)	80.83	8.09	5.457	**c3 "WT" c2**

ANOVA Statistical Analysis (AMR-WB vs EVS-SWB)

Two sets of conditions, i.e., AMR-WB set 1: c02, c03, c04 and EVS-SWB CA set 2: c05, c06, c07 are analyzed based on the Analysis of Variance (ANOVA) metric. The appropriate statistical model is a three factor Analysis of Variance (ANOVA) for Sets (2) x Conditions (3) x Scores (128 - 32 Subjects x 4 Talkers). Table 5.1.2.1-6 shows results of the ANOVA.

Table 5.1.2.1-6. Results of ANOVA for Sets x Conditions x Scores

Source of Variation	df	SS	MS	F	Prob.
Sets	1	2104.8	2104.83	**41.84**	0.0000
Conditions	2	18672.0	9336.00	**117.37**	0.0000
Scores	127	18728.8	147.47		
Sets x Conditions	2	298.7	149.33	**3.15**	0.0434
Sets x Scores	127	6389.2	50.31		
Conditions x Scores	254	20204.2	79.54		
Sets x Cond. X Scores	254	12048.3	47.43		
Total	767	78446.0			

The main effect for **Sets** factor is highly significant (p<0.0001), which means that the average scores for the two sets of conditions are significantly different (AMR-WB: 79.830 and EVS-SWB CA: 83.140). The main effect for **Conditions** factor is also highly significant (p<0.0001), which means that there is significant variation among the average scores for the three conditions (86.950, 82.505, 75.000). Finally, the interaction effect, **Sets x Conditions**, is significant (p<0.05), which indicates that the pattern of scores for the two Sets across the three Conditions is statistically different. Figure 5.1.2.1-5 illustrates the significant interaction between the AMR-WB and the EVS-SWB CA intelligibility scores.

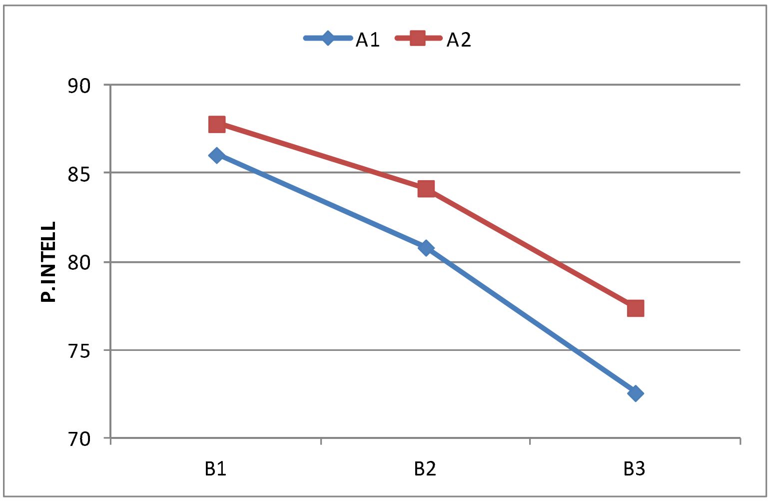


Figure 5.1.2.1-5 Illustration of interaction between Sets and Conditions in the ANOVA

5.1.2.1.1 MCPTT bearers – speech intelligibility

Based on the results in clause 5.1.2.1, it can be concluded that EVS-SWB CA offers a statisically significant improvement in speech intelligibilty over AMR-WB (HD voice). The improved robustness to background noise and resiliency to errors are particularly relevant to the MCPTT service. For example, results for car noise at 20 dB and music and mixed content in clean channel conditions in Clause 5.1.1.4 show robustness to low levels of background noise. Results with other types of noise (e.g., street noise and office babble @ 15-20dB SNR) and in impaired channels (e.g., at FERs in the range of 3-10%) are elaborated in TR 26.952. The test results in Clause 5.1.1.4 and TR 26.952 show voice quality improvements based on the P.800 ACR/DCR test methodology, but may not be fully indicative of speech intelligibility. The speech intelligibility test results based on the P.INTELL test presented in Clause 5.1.2.1 evaluates under high background noise (siren at 5dB) and impaired channel conditions.

Unicast bearer

By definition of the reference summarized in clause 5.1.1.6.1, *AMR-WB meets the reference performance.*

From the results in clause 5.1.2.1, it can be seen that *EVS-SWB CA speech intelligiblity exceeds that of the reference.*

It is important to understand the coverage improvement provided by the improved speech intelligibility of EVS-SWB CA. Based on the analysis in clause 5.1.2.1 showing that speech intelligiblity of EVS-SWB CA @ 8% FER is similar to that of AMR-WB @ 2%, the field test results identified in clause 5.1.1.6.1 demonstrate that for FER<=2%, the VoLTE system achieves 90% coverage while for FER <=8%, the system achieves 99.5% coverage.

Table 5.1.2.1.1-1. AMR-WB and EVS-SWB CA Coverage Over MCPTT Unicast Bearer

Codec and mode	Operating FER	VoLTE Coverage
AMR-WB 12.65 kbps	2%	90.0%
EVS 13.2 kbps SWB CA mode	8%	99.5%

This demonstrates how *EVS-SWB CA can provide similar speech intelligibilty as AMR-WB with better coverage in VoLTE networks.*

MBMS bearer

Based on the results in clause 5.1.2.1 it is expected that *EVS-SWB CA speech intelligiblity exceeds that of AMR-WB under all the MBMS bearer conditions tested.*

To understand the coverage improvement provided by the improved speech intelligibility of EVS-SWB CA, Figure 5.1.2.1.1-1 below shows a CDF of the coverage in an MBSFN cell embedded in a 57-cell system simulation at different FERs on the MBMS downlink[1]. Each curve is constructed by looking at the SNR trace for each terminal in the cell and determining the lowest SNR level below which the terminal experiences at least the target FER, then plots the CDF of all the users against the SNR values. The x-axis then represents the SNR that the MBMS Modulation and Coding Scheme (MCS) needs to work with to achieve the target FER of each curve for the percentage of users represented on the y-axis. Realistically, the most robust MCS (MCS 0) needs an SNR of at least -4dB to operate, so the curves to the left of -4dB SNR are not achievable.

Based on the data provided by this simulation, the MBMS system provides less than 90% coverage at 2% end-to-end FER (1% uplink, 1% downlink). On the other hand, the MBMS system can provide an MCS that would allow for at least 90% coverage at 8% end-to-end FER (1% FER uplink, 7% FER downlink).

This clearly demonstrates how a codec that is more resilient to FER can have improved coverage area compared to a codec that is less resilient to FER.

Therefore, neither AMR-WB nor EVS-SWB CA at 2% end-to-end FER (1% uplink, 1% downlink) can meet the reference coverage KPI because the underlying MBMS system can not provide at least 90% coverage at 2% FER. AMR-WB speech intelligibility performance at 8% FER is statistically lower than AMR-WB speech intelligibility performance at 2% FER. However it is possible for EVS-SWB CA with 8% end-to-end FER (7% downlink and 1% uplink) to exceed the reference coverage KPI without a statistically significant intelligibility performance decrease compared to the reference intelligibility performance.

[1] These results are based on MBMS simulation data.

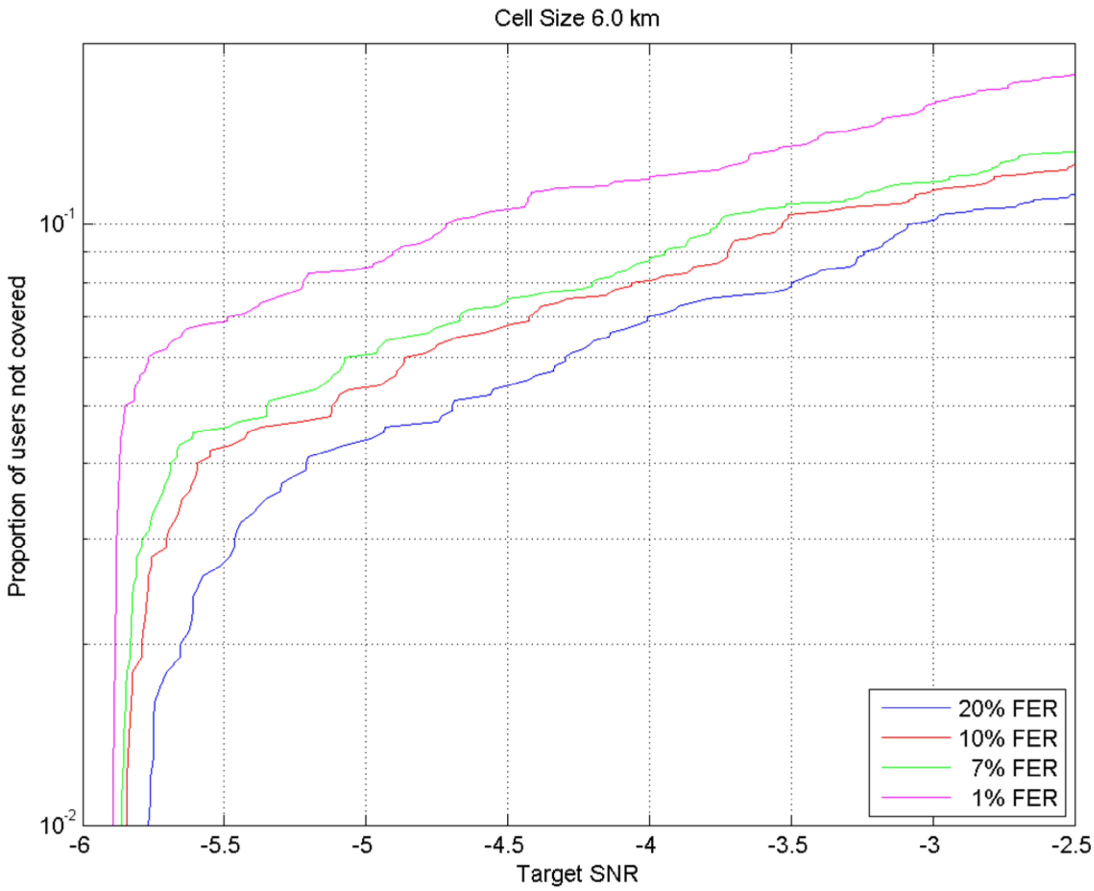


Figure 5.1.2.1.1-1. Coverage in an MBSFN cell (embedded in a 57-cell system simulation) at different FERs on the MBMS downlink

Comparing the 1% and 7% FER curves at any particular SNR value shows how *EVS-SWB CA can provide similar speech intelligibilty as AMR-WB with better coverage over the MBMS bearer.*

LTE-D bearer

The similarities in speech intelligibility between AMR-WB @ 2% FER and EVS-SWB CA @ 8% FER demonstrated in clause 5.1.2.1 can be used to make preliminary estimates about the coverage, power consumption, and capacity improvement provided by the improved speech intelligibility of EVS-SWB CA as was done in clause 5.1.1.6.4.3. The gains based on similarities of subjective speech intelligibility are similar to the gains based on equivalence of voice quality (based on objective P.OLQA measure, Clause 5.1.1.6.3.4), and are listed in Tables 5.1.2.1.1-2 to 5.1.2.1.1-4.

Table 5.1.2.1.1-2. Link budget gains and coverage/range extension offered by EVS codec

Model	BLER Target	Gain	Distance Gain	Area Gain
AMR 12.2 kbps/AMR-WB 12.65 kbps	2%	-	-	-

EVS – 13.2 kbps channel aware mode	~8%[2]	~2.8 dB	~17%	~38%

Table 5.1.2.1.1-3. Power consumption gains offered by EVS modes

Model	TX Power	TX Power Gain	RX Power	RX Power Gain
AMR 12.2 kbps/AMR-WB 12.65 kbps	3.01 units	-	0.16 units	-
EVS - 13.2 kbps CA and non-CA modes	2.91 units	3.3%	0.155 units	3.1%

Table 5.1.2.1.1-4. Increase in capacity -- fraction of successful links offered by EVS codec (3TX/cell and 4TX/cell)

		3 TX/cell		4TX/cell		
Model	BLER Target	Fraction of Successful links	3TX/cell Gain	Fraction of Successful link	4TX/cell Gain	4TX Gain Relative gain w.r.t. 3TX
AMR-WB 12.65kb/s	2%	80%	-	70%	-	**16%**
EVS SWB CA 13.2 kb/s	~8%	~96%	~20%	~91%	~30%	**~51.6%**

A comparison of the EVS-SWB CA and AMR-WB performance with respect to the reference can be performed using Figure 5.1.1.6.3.4-2 as was done in Clause 5.1.1.6.3.4. From the graph, it can be seen that AMR-WB can only get 80% coverage at 2% FER, and to get to the reference 90% coverage would introduce up to 5% FER. Given the trend that AMR-WB speech intelligibility degrades with increasing FER beyond 2%, it is questionable whether AMR-WB can meet the reference coverage for LTE-D bearers. However, more data is needed to show this with statistical significance.

On the other hand, as illustrated in Figure 5.1.1.6.3.4-2, EVS-SWB CA (Option 1) can operate with 90% coverage and 4.5% FER while exceeding the reference speech intelligibility. EVS-SWB CA can also achieve ~96% coverage (at about 8% FER) with the intelligibility similar to the reference. Therefore *EVS-SWB CA exceeds the reference over LTE-D bearers in these three KPIs.*

The above simulations and calculations of coverage focused on the LTE-D data channel and assumed the lower coding rate and lower RB usage of the control (SA) channel would allow the data channel not to exceed 8% BLER.

A further analysis to account for data channel losses caused by errors on the SA channel is calculated using the simulation data illustrated in Figure 5.1.2.1.1-2 on the SA channel link performance and Figure 5.1.1.6.3.4-1 on the link budget performance for the data channel. A link budget of 0.63 dB on the data channel provides 7.4% BLER if the SA channel does not introduce more errors. Since the SA channel uses only 1 RB, the effective SNR is 3.63 dB. The SA channel with 2 HARQs reduces the SA channel BLER to 0.65% which results in an effective BLER 8.0% for the data channel carrying the voice packets which represents the BLER needed for EVS-CA. Performing the same analysis for the BLER needed for AMR-WB, a link budget of 3.56 dB on the data channel results in a data channel BLER of 1.91% and a SA channel BLER of 0.1%, resulting in an effective BLER of 2.0%.

[2] "~X" in this and the subsequent tables indicates "nearly X", i.e., very close to the value X but not exceeding it.

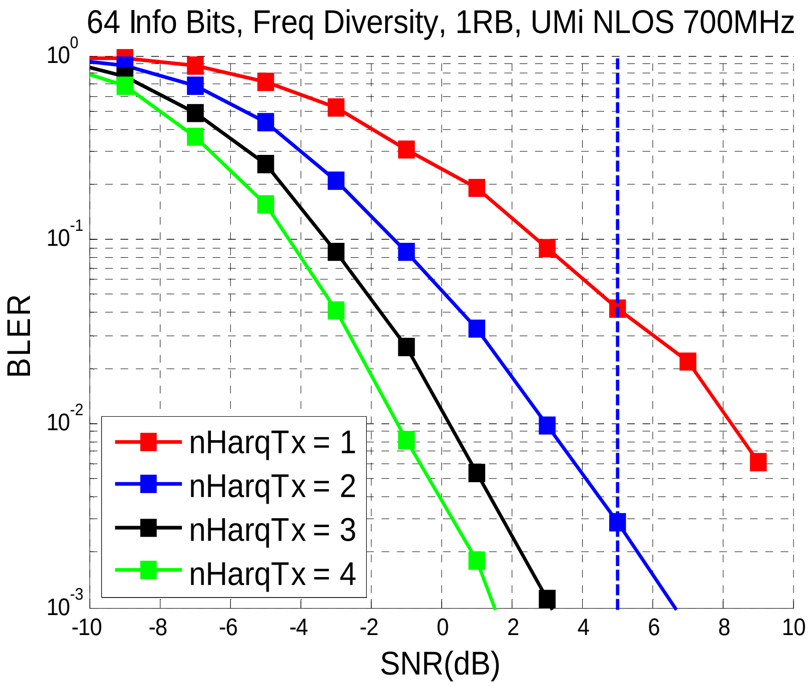


Figure 5.1.2.1.1-2 SA Channel Link Performance

This provides the following results on coverage and capacity over LTE-D bearers.

Table 5.1.2.1.1-5. Link budget gains and coverage/range extension offered by EVS codec

Model	Target BLER of data channel	Effective BLER with SA channel errors	Link Budget of data channel	dB Gain	Distance Gain	Area Gain
AMR 12.2 kbps/AMR-WB 12.65 kbps	1.91%	2.0 %	3.56 dB	-	-	-
EVS – 13.2 kbps channel aware mode	~7.4%	~8.0 %	~0.63 dB	~2.93 dB	~18%	~40%

Table 5.1.2.1.1-6. Increase in capacity -- fraction of successful links offered by EVS codec (3TX/cell and 4TX/cell)

		3 TX/cell		4TX/cell		
Model	Target BLER of TX channel	Fraction of Successful links	3TX/cell Gain	Fraction of Successful link	4TX/cell Gain	4TX Gain Relative gain w.r.t. 3TX
AMR-WB 12.65kb/s	1.91%	79.5%	-	69%	-	**15.7%**
EVS SWB CA 13.2 kb/s	~7.4%	~95%	~19.5%	~90%	~30%	**~50.9%**

5.1.2.1.2 Conclusions

Speech Intelligibility:

- EVS-SWB offers significant speech intelligibility improvements over the reference. It also offers even more improvements over AMR-WB under the same bearer conditions. The improved robustness to background noise and resiliency to errors are particularly relevant to the MCPTT service and result in better speech intelligibly compared to the reference and AMR-WB across all the MCPTT bearers.
- AMR-WB meets the reference speech intelligibility for unicast bearers.
- A downward trend in intelligibility performance with increasing FER raises a question about whether AMR-WB can meet the reference speech intelligibility for the MBMS and LTE-D bearers. However more data is needed to show this with statistical significance.

In summary, based on the select speech intelligibility test results presented in clause 5.1.2.1,

- EVS-SWB outperforms AMR-WB in the selected KPIs and across all the MCPTT bearers.
- There are some MCPTT bearers where some KPIs cannot be meaningfully compared to the reference "HD Voice" experience:
    - In the cases that allow a comparison to the reference (call capacity on unicast, coverage and error-resiliency/intelligibility on all MCPTT bearers), EVS exceeds the performance of the reference for all the bearers. In these cases, AMR-WB meets the reference for unicast bearers. A downward trend in speech intelligibility performance with increasing FER raises a question about whether AMR-WB can meet the reference KPI for the MBMS and LTE-D bearers. However more data is needed to show this with statistical significance.
    - In the cases that do not allow a comparison to the reference (capacity on MBMS and LTE-D), the EVS codec outperforms AMR-WB across all MCPTT bearers. Based on a similar analysis as in Clause 5.1.1.6.3.2-4, the call capacity of EVS-SWB CA exceeds that of AMR-WB across all MCPTT bearers.

The Table 5.1.2.1.2-1 below provides a comparison based on the above conclusions and highlights some key results.

Table 5.1.2.1.2-1: Conclusions and Key Results of Speech Intelligibility Testing

Key Performance Indicator (KPI)	**EVS Compared to AMR-WB**	**EVS Compared to Reference**	**AMR-WB Compared to Reference**[3]
Coverage	Exceeds for all MCPTT bearers ~38% better coverage for LTE-D bearers	Exceeds for all MCPTT bearers	Meets for unicast bearers Questionable whether meets for MBMS and LTE-D bearers.[4]
Error resiliency/Speech Intelligibility	Exceeds for all MCPTT bearers. Similar intelligibility performance between AMR-WB at 2% FER and	Exceeds for all MCPTT bearers	Meets for unicast bearers Questionable whether meets for

[3] The Reference is "HD Voice" AMR-WB performance over 3GPP networks as defined in clause 5.1.1.6.1. This is different from the Analog-FM codec performance which is used as a reference by some members of the Public Safety community.

[4] More data is required for AMR-WB speech intelligibility between 2 to 8% FER to confirm that AMR-WB is unable to meet the reference KPI.

	EVS-SWB at 8% FER		MBMS and LTE-D bearers.[4]
Call Capacity	Exceeds for all MCPTT bearers An estimated 20-30% better capacity than AMR-WB for LTE-D bearers Can also support more group than AMR-WB for LTE-D bearers	Exceeds for unicast bearers Cannot compare for other MCPTT bearers	Meets for unicast bearers Cannot compare for other MCPTT bearers

5.1.3 Review of Codec Alternatives and their Relative Complexity

The complexity of the EVS codec is described in subclause 13.2 of TR 26.952 [3]. Tables 13.2a, 13.2c and 13.2d from [3], reproduced below as Tables 3, 4,and 5, provide a comparison of the mean of the per-audio-frame complexity for EVS and AMR-WB, analysed with a source file comprising 8.5 minutes of mixed speech and music. In the case of the decoders the complexity is measured in the presence of 30% frame erasures and the EVS operation is analysed by bandwidth and sample rate.

Table 3: Highest values of the mean complexity for different sample rates and coded bandwidths

Coded Bandwidth	Sample Rate (kHz)	Encoder Complexity (WMOPS)	Decoder Complexity (WMOPS)	Balanced Combined Complexity (WMOPS)	Worst/Worst Combined Complexity (WMOPS)
NB	**8**	**29.98**	**15.40**	**43.97**	**45.38**
	16	31.52	16.38	47.35	47.90
	32	37.34	19.37	56.71	56.71
WB	**16**	**38.25**	**18.29**	**51.38**	**56.54**
	32	43.32	22.88	60.92	66.20
SWB	**32**	**45.12**	**22.67**	**65.56**	**67.80**

Table 4: Highest values of the mean complexity for AMR-WB

Coded Bandwidth	Sample Rate (kHz)	Encoder Complexity (WMOPS)	Decoder Complexity (WMOPS)	Balanced Combined Complexity (WMOPS)	Worst/Worst Combined Complexity (WMOPS)
WB	**16**	**32.30**	**7.80**	**38.73**	**40.11**

Table 5: Incremental mean complexity for EVS relative to AMR-WB

Coded Bandwidth	Sample Rate (kHz)	Encoder Complexity (WMOPS)	Decoder Complexity (WMOPS)	Balanced Combined Complexity (WMOPS)	Worst/Worst Combined Complexity (WMOPS)
NB	**8**	**-7.18%**	**+97.4%**	**+13.5%**	**+13.1%**
	16	-2.41%	+110%	+22.3%	+19.4%
	32	+15.6%	+148%	+46.4%	+41.4%
WB	**16**	**+18.4%**	**+134%**	**+32.7%**	**+41.0%**
	32	+34.1%	+193%	+57.3%	+65.0%
SWB	**32**	**+39.8%**	**+191%**	**+69.3%**	**+69.0%**

From Table 5 it can be seen that under the same conditions, the EVS codec requires 69% more processing on average to reproduce twice the audio bandwidth in a SWB capable UE than the AMR-WB codec providing WB. It can also be seen that in a WB-only capable UE, the EVS codec requires 41% more processing on average..

The worst-case observed complexity of EFR and AMR is given in [7] as 15.21 and 15.33 WMOPs respectively but both the bit-exact instruction weights and methodology have changed in arriving at the complexity figures for EVS and for AMR-WB above. The complexity of the 3GPP Codecs, beginning with GSM EFR in 1995 [7 & 8], followed by AMR in 1999 [7], AMR-WB in 2001 [13] and EVS in 2015 [3] is plotted in Figure 1.

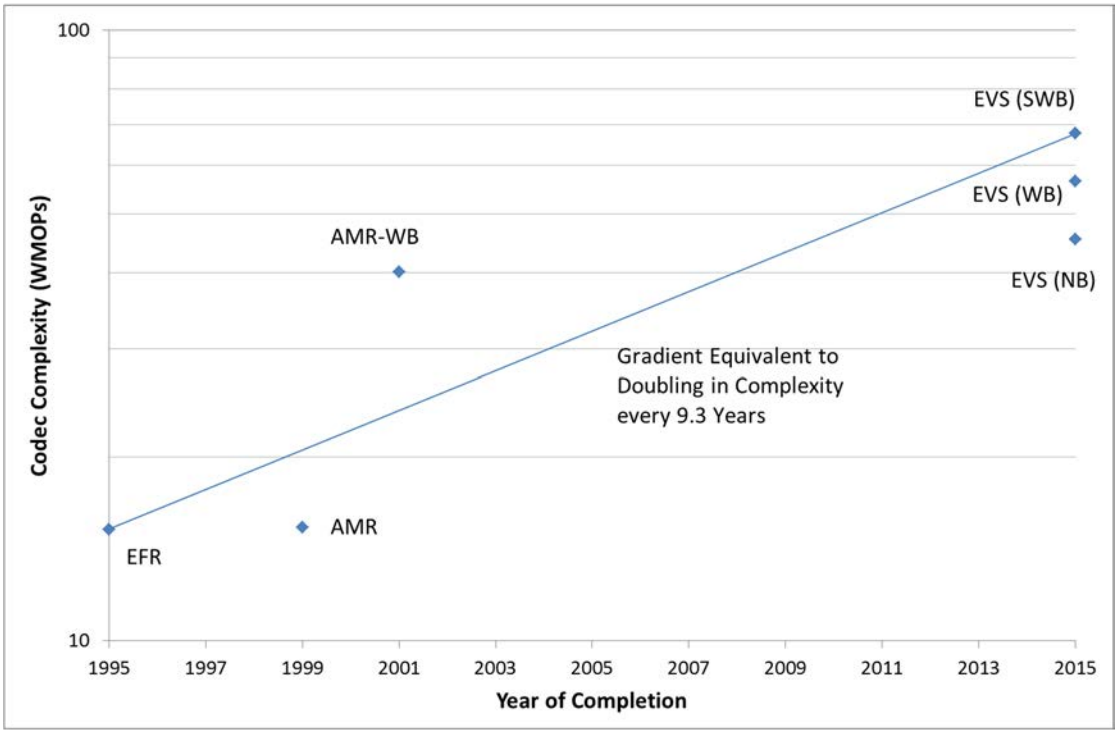


Figure 15: Complexity of 3GPP Codecs over time

The computational complexity is one metric which can be considered amongst others. It is acknowledged that there are other important complexity metrics, such as the physical footprint or area, but these are not readily available and not provided in the document.

5.1.4 Recommended requirements

It is important that at least one of the codecs supported in all MCPTT terminals is capable of providing equivalent or better performance than 3GPP wideband voice ("HD Voice") in terms of coverage, error-resiliency, speech quality, speech intelligibility, and call capacity across all MCPTT radio bearers.

It is recommended that a codec that is important for MCPTT communications be mandated for MCPTT terminals.

It is recommended that a codec that has features that are only "nice to have", but not essential, for MCPTT communications be recommended for MCPTT terminals. The network transcoding functions have to support this codec if the codec is to be used between terminals in MCPTT sessions.

Requirements for interworking with legacy public safety systems is outside the scope of this release.

5.1.5 GAP Analysis and Evaluation

5.1.5.1 Requirements on Audio/Voice Quality

With regard to Audio/Voice Quality (TS 22.179 [2] subclause 6.15.5), several narrowband quality benchmarks, met by the P25 codecs and perhaps also the TETRA codec, should be surpassed. Whilst these benchmarks are expressed in terms of objective predictions of subjective MOS (ACR) test scores (Recommendations ITU-T P.862 and ITU-T P.863), it is well known that such techniques have shortcomings when comparing codecs of different technologies, not to mention codecs with different audio bandwidths. It is nevertheless clear that high audio quality is to be preferred.

Interestingly there is no mention in TS 22.179 about requirements for speech intelligibility but it is quite obvious that for the MCPTT application, speech intelligibility is a key requirement. Neither audio quality nor intelligibility in noise is mentioned, apart from in the context of the inherent noise reduction capabilities of the coding algorithm. Again though, high quality and intelligibility in background noise would appear to be self-evident requirements.

5.1.5.2 Discrete/Ambient Listening and Remotely Initiated Monitoring

For Discrete/Ambient Listening (TS 22.179 [2] subclauses 6.16.1 & 6.16.2) and Remotely Initiated Monitoring (TS 22.179 [2] subclause 6.16.3) it is clearly desirable to encode the audio/speech signals present in the foreground and any background signal.

These requirements clearly favour high quality audio codecs able to cope with non-speech signals. It also seems self-evident that wider audio bandwidths, able to capture more details, are to be preferred. It may also be preferred for noise suppression algorithms present in UE's and DTX to be disabled in discrete/ambient listening and remotely initiated monitoring for optimum performance.

5.1.5.3 Noise Reduction

From examination of the Noise Reduction requirements described in subclause 5.14 of [2] it is clear that the contributors to TS 22.179 wish to emphasize the inherent noise suppression capabilities of the low bit rate coding algorithms of P25 and TETRA. This emphasis, whilst one possible approach to the problem of background noise, is however in conflict with the requirements for high audio/voice quality and with the requirements for ambient listening which are best delivered by accurate rendering at the decoder of the input signal to the codec.

3GPP is well aware of the need for high quality audio in the presence of background noise and, rather than rely on the inherent weaknesses of the coding algorithms, has developed and encouraged the adoption of adaptive noise suppression technologies in UE's prior to the audio encoding stage to address these challenges. See [14] and [15]. As mentioned above though, a method of disabling the noise suppression algorithms present in the UE's would be advantageous for the Discrete/Ambient Listening and Remotely Initiated Monitoring applications.

5.1.5.4 Common Codec Constraints of MCPTT

MCPTT has the following scenarios/use cases that impact the use of mandatory/recommended codecs:

- The group communication can be off-network using the D2D physical layer, in which case a transcoding function is not available. The lack of transcoding requires that the codec selected has to be supported by all terminals in the session. Furthermore, the D2D physical layer for group call uses a broadcast channel that is received by the group members. This also requires that the codec selected has to be supported by all terminals in the session.

- The group communication can be on-network using a broadcast bearer. This also requires that the codec selected has to be supported by all terminals in the session.

The need to use a common codec among all the participants impacts the ability to use any recommended codecs due to the following:

1) If one of the participants does not support the recommended codec then the call set-up will fail or require codec re-negotiation. This raises the following issues:

a) Codec renegotiation delays the call set-up and may not be acceptable in all scenarios, especially for mission-critical communications.

b) Codec re-negotiation for MCPTT is more cumbersome than for typical point-to-point calls due to the following:

i) The responses to the session initiator regarding its selected recommended codec can come at different times from the other participants. If some terminals are in poor radio conditions their responses may be lost or delayed due to transport layer retransmissions. Thus the need to re-negotiate the codec may not be known until later into the call. If the initiator waits for confirmation from all participants before sending media then the media start could be very delayed. If the initiator starts sending media immediately, it will mean that some terminals (the ones unable to use the codec) will still experience much delay before being able to render media to the user. The others will be interrupted when the codec is re-negotiated, likely down to a lower quality codec, thus causing a poor user experience.

ii) If the initiating terminal does not get any additional information about the other session participants (aside from a call rejection), then the initiator may have to try different recommended codecs multiple times before selecting one that all the terminals can use.

iii) If the initiating terminal receives additional codec capability information from the session participants in response to its proposed codec, this requires transmission of more information (e.g. codec capabilities/profile of the terminal) to the initiator in the reverse direction. This will require more time to send the additional information from multiple participants and require that the initiator wait even longer before deciding on what to include in the codec renegotiation proposal, thus delaying the transmission of media.

2) To avoid setting up a session that might require codec re-negotiation the terminal can resort to the following:

a) Not use any recommended codecs and only use a mandatory codec

b) Attempt to determine out-of-band and beforehand, the codec capabilities of all the terminals it wishes to add to the call

i) This could be done by pre-provisioning the terminals with a profile that is shared among a group of callers, i.e., the Rennes Police Department; the members of the MBS SWG.

ii) Or this could be obtained through some capabilities exchange performed out-of-band which may be application- or lower-layer based.

3) Even when all the terminals in a session can support the selected recommended codec, this still places constraints or impacts performance if the group wants to add another caller to the existing session. This raises the following issues:

a) The participants would have to know beforehand that the new caller can support the codec they are using. How can this be easily done in a user-friendly way?

b) If the above is not known, there is a chance they would have to renegotiate their codec and "dumb-down" their media to match that of the new caller.

c) If there is a re-negotiation to another codec there will be disruption in the call and the existing callers will most likely notice a degradation in call quality, e.g. going from SWB to WB, or from WB to NB audio. This results in a very poor user experience.

5.1.5.5 Requirements on Transcoding Functions in the Network

In on-network point-to-point communications which allow use of a transcoder function, there are some challenges that need to be considered when attempting to use a recommended codec for MCPTT.

1) There needs to be a transcoder function to support use of recommended codecs. How does the terminal know that there is a transcoder function in the network?

2) The transcoder function has to support the recommended codec in order for the terminal to be able to use the codec, even if all the other terminals in the call support the recommended codec. An MRFC cannot allow a recommended codec that its MRFP does not support to be included in the SDP Offers relayed to the called

participants. If some of the participants answer using the recommended codec while others do not, the MRFP will not have the proper codecs to support the session.

Therefore, for a MCPTT group to be able to use a recommended codec for its own-network sessions without codec-renegotiation **requires that the network transcoding function supports the recommended codec**. This becomes complicated to ensure when the MCPTT group is not closely coordinated with the MNO who owns and operates the transcoder function in the network. For example, how does the MBS SWG user group ask/guarantee that the Orange network in Rennes supports EVS transcoding?

5.1.6 Criteria with respect to MCPTT codec selection

Many factors were considered in the evaluation of codecs for use in MCPTT. In addition to the analysis in the preceding clauses some of these factors and trade-offs are described in this clause.

For example, public safety grade communications (e.g. MCPTT) are required to operate with high levels of reliability and availability (e.g. minimum downtown time). Another example is the importance of conservation of battery across a first responder's long shift.

While the following list is not meant to be exhaustive (there are selection factors that may not be captured here), these factors give additional information.

1) Availability of devices that include the codec –

This can be described as time to market. The sooner a large number of UEs that utilize the codec are available for MCPTT the better the timeline of early MCPTT deployments can be met. Related to this factor are:

a) use of the MCPTT application on existing devices

b) large volume of devices that utilize the codec may help drive down cost

2) Maturity of the codec vs. the benefits of the latest standardized codec –

This can be described as a trade-off balancing the benefits in performance afforded from the latest standardized codec vs. the known performance and reliability of an existing codec that has been deployed on millions of devices for several years. Related to this factor:

a) stability – trade-off balancing the benefits from improvements afforded from the latest standardized codec vs. the stability of using a codec that is less likely to change over time

3) Flexibilty, complexity, and configuration –

This can be described as a trade-off balancing the benefits in flexibility, and improved performance under certain conditions, afforded from the latest standardized codec vs. code complexity and additional configuration options. Related to this factor are:

a) use of the codec on non-3GPP terminals (wireline), e.g. consoles and dispatch terminals

b) power consumption for terminals with chipsets containing the codec and processing demands for terminals without chipsets containing the codec (e.g. software downloadable)

c) integration effort to effectively utilize a codec in terminals

NOTE: Despite disagreement on the importance of the factors in this clause and the previous clauses, a codec decision is documented in clause 6.

5.1.7 Solution

After consideration of all factors in Release 13 the currently widely deployed and available AMR-WB codec meets the needs of MCPTT services and is recommended as the mandatory codec for MCPTT. Based on operator / MCPTT service provider policy, it is recommended that EVS operating in super-wideband mode be supported as an optional codec that can be used for MCPTT services.

5.2 Key Issue#2: User Experience

5.2.1 Description

A building in the business area, which is covered by cellular network, is on fire, an MCPTT group call is set up and fire fighters involved in the mission join this MCPTT group. Most MCPTT group members, who stay inside of building, receive the MCPTT group call over broadcast channel, some fire fighters, who stay outside of building, receive the MCPTT group call over unicast channel. All fire fighters have the same experience of mouth-to-ear latency and floor control regardless the audio is transferred over unicast or broadcast delivery. A fire fighter carrying an injured person moves out of building and waits the ambulance in a safe zone. After the injured person is transferred away, he reports this information to the MCPTT group while he is moving back into the building and helps other wounded people. He notices that the fire in the room will be out of control; he immediately reports this information to the MCPTT group and starts to move the wounded person away from this area. After several hours of hard work, the fire in the building is finally put out and the MCPTT group is dismissed.

5.2.2 Recommended requirements

The following recommended requirements are derived from the uses case.

- For on-network mode case, it is recommended that the service interruption is minimized when the MCPTT UE moves into/out of MBMS coverage.
- For on-network mode case, it is recommended that all MCPTT users will have the same experience, e.g. access delay regardless the MCPTT call is transferred over UC or BC.
- For on-network mode case, it is recommended that the difference of Mouth-to-ear of latency of audio payload between MCPTT users using unicast delivery method and MCPTT users using broadcast delivery method is minimized.

NOTE: In TS 22.179, the user experience related requirement is also copied below:

"[R-6.15.3.2-002] The MCPTT Service shall provide the MCPTT Access time and Mouth-to-ear latency specified in this subclause to all MCPTT Users related to an MCPTT call regardless of call type (e.g. group, Private Call), group size and/or user density.

NOTE: This ensures that all MCPTT Users experience the same performance regardless of whether the audio is transferred over unicast or multicast delivery."

5.2.3 GAP Analysis and Evaluation

5.2.3.1 Longer e2e delay over BC bearer issue

Per TR 36.868, The minimum end to end delay for media transport for Group Communications over eMBMS is about160 ms (refers to TR 36.868, clause 5.2.1.1.3), and end to end delay for media transport over unicast bearers is about 40 ms (refers to TR 36.868, clause 5.1.1.3).

The RTP payload transmission timing sequence is depicted in figure 15a.

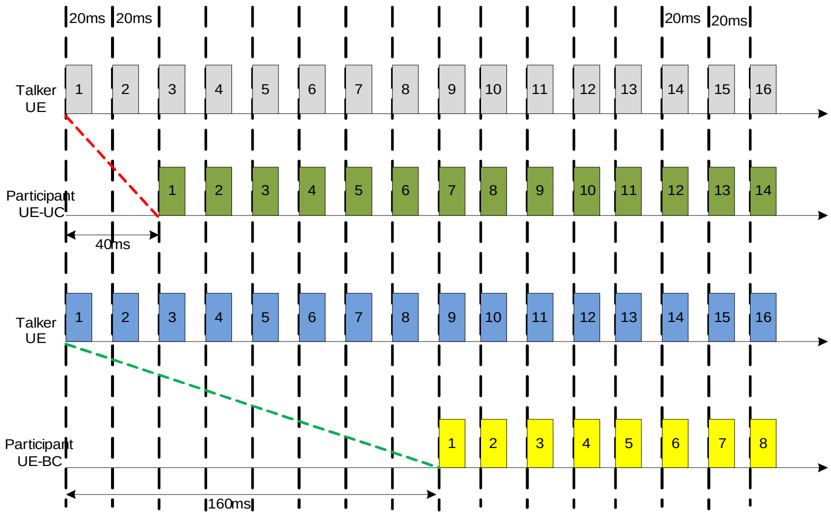


Figure 15a: The RTP payload timing sequence

The longer delay over BC degrades MCPTT user's experience. Consequently, the delay for unicast reception and for MBMS reception may be different. This identified gap needs to be addressed.

5.2.3.2 Mobility issue

Per TS [, the GCS AS acts as the media source, the BM-SC receives the RTP payload from the GCS-AS. The received RTP payload will be put on the MBMS bearer without modifying any RTP headers like SSRC, SN and TS, as this information is filled by GCS AS.

Figure 16 presents a case that the UE in a MCPTT call moving in/out of MBMS coverage.

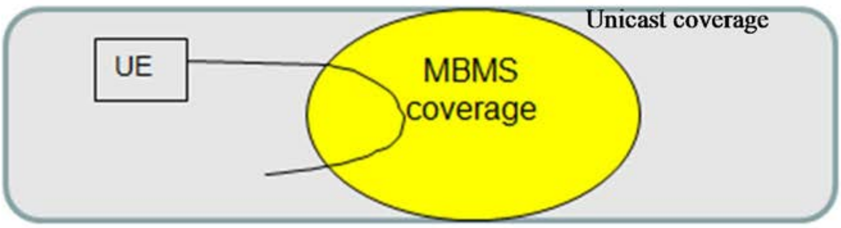


Figure 16: Handoff scenarios between UC bearer and BC bearer

When the distribution delay of unicast and broadcast is different, then the quality may be distorted during handoffs.

5.2.3.2.1 BC handoff to UC

When the UE moves out of MBMS coverage, for the break before make case (refers to TS 23.468 [12], clause 5.3.3), the UE starts receiving DL data over unicast after it has stopped receiving data over eMBMS. As showed in figure 17, UE receives RTP payload (SN=1,2,...,m) at MBMS coverage, BC HO to UC occurs at time t0, the UE receives the RTP payload with SN=m+n due to longer transmission delay over MBMS bearer.

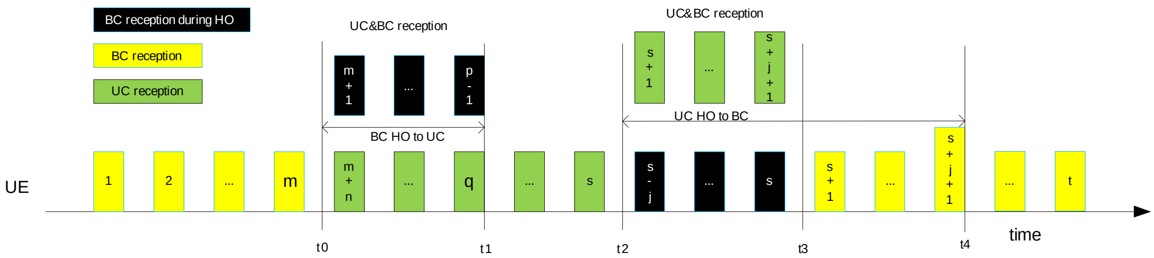


Figure 17: RTP advance and rewind caused by handoff between BC and UC

This SN number advancing issue impacts user's experience.

For the make before break case, when the UE notifies the GCS AS that the UE is moving out of MBMS coverage and the GCS AS starts to send the data over UC. The UE simultaneously receives data by Unicast Delivery and MBMS Delivery. As showed in figure 17, During HO period from t0 to t1, the UE receives the RTP payload with SN=m+1 to p-1 over MBMS bearer, the UE receives the RTP payload with SN=m+n to q over UC bearer. UE has two sequences of RTP payloads with discontinuous SN number.

The RTP sequence selection and SN number advancing issue needs to be fixed.

5.2.3.2.2 UC handoff to BC

For switching from unicast delivery to MBMS delivery case (refers to TS 23.468, clause 5.3.2), the UE simultaneously receives data by Unicast Delivery and MBMS Delivery. As shown in Figure 17, from t2 to t3 period, the UE receives duplicated RTP payloads (SN number from s-j to s) over MBMS bearer. At the same time, the UE continues receiving new RTP payloads (SN number from s+1 to s+j+1) over unicast bearer.

The UE notifies the GCS AS via GC1 that it is in MBMS coverage and receiving the MBMS bearer service. At time t3, the GCS AS stops sending the data over by Unicast Delivery to this UE. The UE now receives the content only by MBMS Delivery. The UE receives RTP payload with SN starting from s+1 over broadcast bearer. During time t3 to t4, the UE receives RTP payload with SN starting from s+j+1 over BC bearer.

The RTP sequence selection issue needs to be fixed during t2 to t4.

5.2.4 Assumptions

The following list of working assumptions is derived from the uses case.

-- For on-network mode case, an MCPTT user's service experience is not interrupted by the movement of the MCPTT UE and the change of delivery method of audio.

- For on-network mode case, the MCPTT service is able to grant the floor control regardless of the MCPTT UE's reception mode (unicast reception, broadcast reception).

5.2.5 Solution

5.2.5.1 Transport delay difference adjustment

To mitigate delay differences between MBMS bearer and unicast bearers, the transport delay difference of MBMS bearer is proposed to be reported back to GCS AS. The GCS AS adjusts the timing of RTP payload over unicast to minimize the transport delay between UC and BC.

The reporting method has 3 options.

- Option A: RTCP method (RFC 3550)
- Option B: QoE procedure of MBMS
- Option C: GC1 interface enhancement

Option A requires RFC enhancement or 3GPP defined extensions, RTCP reporting interval is another concern. Option B only requires QoE enhancement, however, whether MBMS QoE procedure is applied to MCPTT is open now. Option C requires SA6 co-ordination.

It is proposed to consider option B considering re-using existing MBMS mechanism as much as possible.

5.2.5.2 RTP payload treatment

Once UE receives 2 RTP flows from both UC and BC with same and/or different SN value (but the same SSRCs), UE should accept those 2 RTP flows, discard redundant RTP payloads, reorder those RTP payloads and submit ordered RTP payloads to the upper layer.

If RTP SN advancing is detected, the UE should reports advancing information back to GCS AS to allow transport delay adjustment if the majority of participants use MBMS reception. The GCS may not perform any delay adjustment if the majority of participants use unicast delivery.

5.3 Key Issue#3: MCPTT over MBMS support

5.3.1 Description

The main purpose of MCPTT is one to many communications. The speech contributions from a single user is distributed to many receivers. A central floor control arbitrator (here, logically located in the MCPTT server) [2] controls the floor.

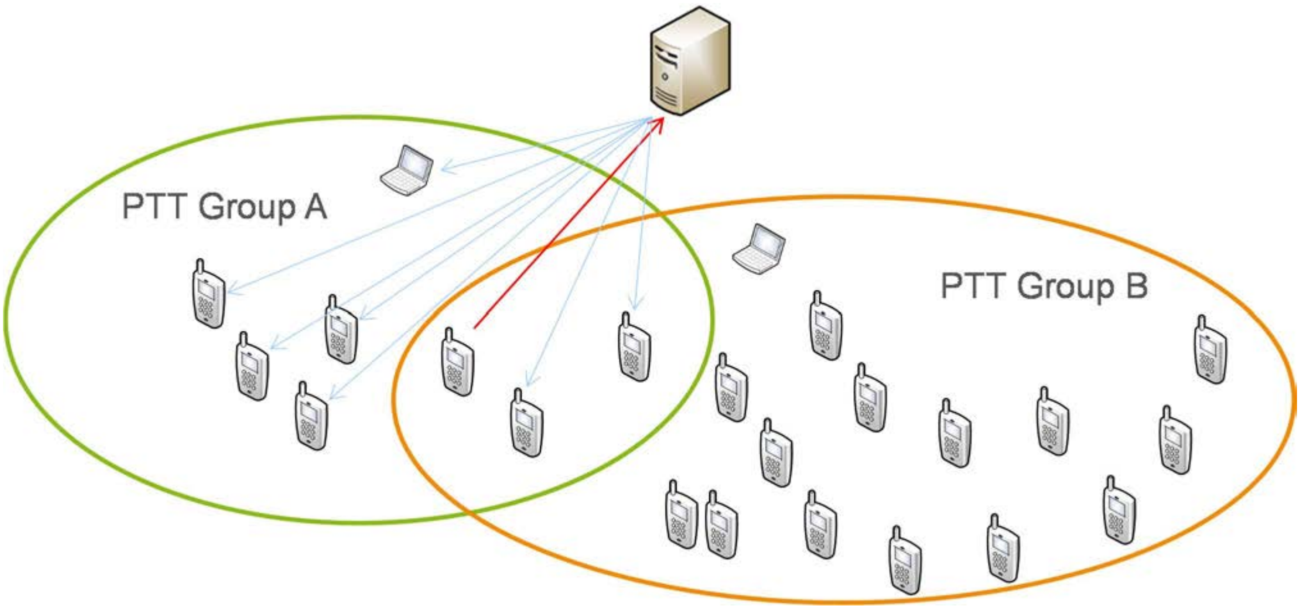


Figure 18: PTT Groups

MCPTT needs to support many simultaneous and possibly large groups. A MCPTT user may be affiliated to one or more groups at a time. The groups may be used by different organizations such as police and fire brigade. MCPTT users are typically affiliated to the groups of a single organization. MCPTT users may also be affiliated to groups of multiple organizations. It should be assumed that MCPTT users are generally affiliated to more than one group at a time.

MCPTT Group call media may be carried over unicast and/or MBMS bearer services. For the media plane, the MCPTT server decides on one of the three cases:

1: which MCPTT group call media is forwarded over unicast-only,

2: which MCPTT group call media is forwarded over broadcast-only or

3: which MCPTT group call media is forwarded over unicast for some users and broadcast at the same time for other users.

It is unclear, whether case 2 can be regarded as a sub-case of case 3 or needs to be explicitly supported.

Note, service continuity between unicast and broadcast needs to be supported for case 3 (above). MCPTT devices may move into the broadcast coverage or may leave the broadcast coverage without or without significant (ffs) service interruption.

An MCPTT group is defined in a group management system. The group definition includes a list of members of the group. When a MCPTT user is authenticated and registered to the MCPTT system the user can affiliate to one or more MCPTT groups that the user is a member of. The definition of being affiliated is found in TS 22.179 [2]:

Affiliated MCPTT Group Member: An MCPTT Group Member who has indicated an interest in the group and is prepared to receive and/or transmit Group Communications from/to the particular MCPTT Group.

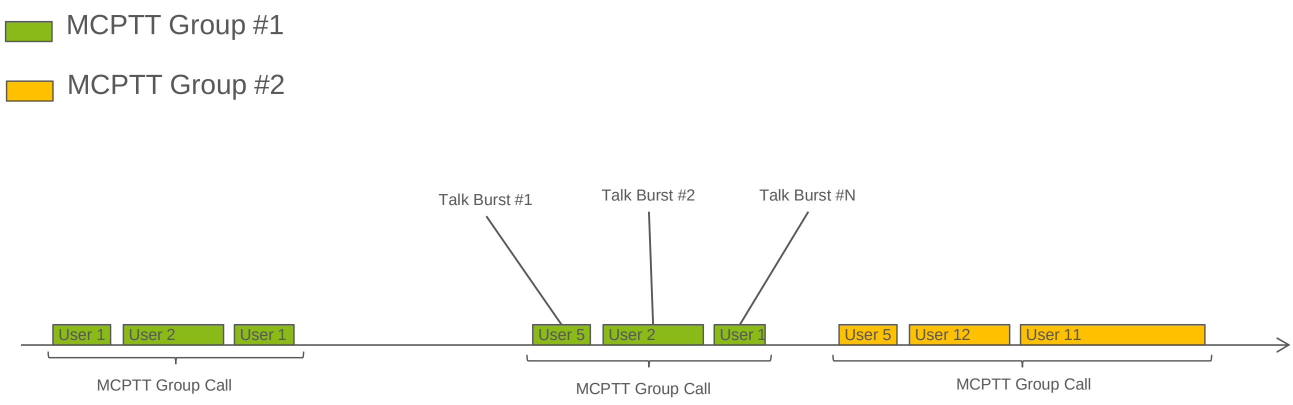


Figure 19: Example MCPTT calls (downlink only)

The figure above illustrates some MCPTT calls of two different groups as example. Note that user 5 belongs to both groups. Each MCPTT Group Call consists of one or more speech contributions from different users. Each speech contribution is called a Talk Burst. Note that User 1 in the above example had two Talk Bursts in the first MCPTT Group Call.

An MCPTT Group Call is typically short and lasts for typically not more than 30 sec (these are only considerations, a talk burst can be 10 sec, but may be longer or shorter) . Resources are released based on idle timers. The MCPTT Group Call setup (Access Time) time will be below 300 ms.

In order to achieve low Group Call Setup times, the MBMS bearer is already established and devices inside of the MBMS coverage have already activated the reception of the bearer. MBMS MCPTT Call setup is triggered by the first floor request. Devices will continuously receive the MBMS bearer traffic, when in broadcast coverage, in order to receive the MBMS MCPTT floor control messages.

The MBMS Bearers used for MCPTT will be pre-established. In order to increase efficiency, each MBMS Bearer Service is provisioned to carry the traffic of multiple MCPTT groups. For example, an MBMS Bearer Service is provisioned to carry 20 different MCPTT groups, but allow only talk bursts of up to 4 groups at a time. In this example, the MBMS bearer is provisioned to carry 4 simultaneous talk bursts (e.g. GBR set to 80kbps, while each talk bursts needs 20kbps). Devices become aware about the different, possible MBMS MCPTT groups during service announcement.

Devices will locally discard data, which belongs to not-affiliated groups (the device may receive data, which belongs to other groups). One MBMS bearer service should typically only carry MCPTT group data for groups within one organization, to minimize that data that is discarded by the MCPTT client.

Speech contributions (called Talk Bursts) from one or more users can be distributed during the MCPTT Group Call. The MCPTT server decides which user gets the floor (floor control). For each MCPTT Group Call (or for each talk burst), the MCPTT server decides whether to map the Call onto unicast or broadcast. If another MCPTT Group Call with a higher priority or with a larger receiver size becomes active (or increases in priority), the MCPTT changes the decision and moves an ongoing MCPTT Group call to unicast. There may be multiple different reasons why the MCPTT server decides to reallocate an on-going MCPTT call to broadcast or to unicast. The service continuity realization should allow the reallocation of on-going group call from unicast to broadcast and vice versa without or with minor (ffs) service interruptions.

The MCPTT system may utilize multiple MBMS Bearers based on network operator policy. Note, per 3GPP Rel 12 standards, existing MBMS UEs are only required to receive a single MBMS bearer at a time. Further, per 3GPP Rel 12 standard, existing MBMS UEs are not required to support unicast interactions while receiving MBMS bearers (simultaneous usage of unicast & MBMS bearers). Note, the on-going 3GPP Rel 13 PRoSE work assumes that at least relay UEs are capable of handle unicast and multiple MBMS bearer services simultaneously.

There are requirements to allow overriding in MCPTT. An on-going MCPTT talk burst may be overridden. Thus, the MCPTT client will monitor the reception while talking in order to detect overrides (i.e. receive MCPTT data while transmitting MCPTT data).

Audio rendering details (i.e. when the device needs to render contributions from multiple sources) are left to implementation.

Any service announcement will be done well in advance, likely at the time, when the device affiliates with the group. The precise time and procedure for service announcement is open.

5.3.2 Deployment Considerations

MBMS is currently deployed in several networks. The present MCPTT architecture assumes, that all MBMS related functions are re-implemented and realized in the MCPTT server and the MCPTT Client.

The deployment option should be studied, where the MCPTT Server can delegate functions into the BM-SC and whether the MCPTT client can leverage functions from a separately implemented MBMS Client. The present MEPRO-API work illustrates the possible realization of APIs between an App (here the MCPTT client) and the MBMS client.

5.3.3 Realization (Stage 3) Considerations (On-Network)

From the general MCPTT description, a couple of Stage 3 considerations and requirements can be derived.

The MCPTT client may be affiliated to multiple MCPTT groups.

The MCPTT clients have always established unicast uplink EPS bearer for media for a given group (e.g. for requesting the floor and providing speech contributions) simultaneously with downlink MBMS bearer for media plane.

The MCPTT client has established downlink EPS bearer for media plane while it is able to receive the MBMS bearer for that group.

RTP session for broadcast may have also RTCP uplink associated to it.

The MCPTT client notifies the MCPTT server, once it is able to receive the MBMS bearer which is associated with the group. The MCPTT server may not provide the MCPTT Group Call over unicast EPS bearer to that MCPTT client until the MCPTT client informs the MCPTT server that the MCPTT client has left or is leaving the MBMS bearer coverage. Note that the MCPTT server may not always provide the MCPTT Group Call over MBMS bearers, if other Group Calls are given higher priority for that MBMS bearer.

Per indication by MCPTT server that MBMS bearer is established, the MCPTT client will continuously receive and monitor the MBMS Bearer, when inside of the broadcast coverage, in order to understand, whether the content of an affiliated group is sent over the associated MBMS bearer. The MCPTT client notifies the MCPTT server over GC1 interface, once it does not finding the MBMS bearer anymore.

It is assumed, that the MCPTT client will select receiving MCPTT group call over MBMS if it is available. When MBMS access information (see clause 5.3.7) is provided for an MCPTT group, then the MCPTT client will activate MBMS reception through the time of affiliation with the group.

MCPTT clients have to be prepared that traffic from non-affiliated groups is received on MBMS bearers. The MCPTT client will silently discard the traffic of non-affiliated groups, when receiving MCPTT via MBMS bearers.

Usage of UDP sessions and destination UDP ports (in downlink direction from MCPTT server to the MCPTT client): In case of unicast, the UDP destination port (scoped with the IP unicast address of the device) is allocated by the receiver and communicated back to the MCPTT server.

Each MBMS Bearer Service may be provisioned to carry the traffic of multiple MCPTT groups, not all Groups are active at a time. The device may need to monitor multiple MBMS bearers.

5.3.4 Media Handling

In unicast delivery mode, the UE address is assigned by the P-GW. The UE registers its IP address to the MCPTT server. The MCPTT server encapsulates with UE's IP address as destination IP address of RTP payload. The P-GW routes IP packets from the MCPTT server to the UE.

In broadcast delivery mode, the group's multicast IP-address is provided by the MCPTT server to the UEs. Therefore the MCPTT server encapsulates with the multicast IP address as used by the UEs as destination IP address of RTP packets and floor control packets. Once the BM-SC receives UDP/IP packets over MB2-U interface, the received packets will be put on the MBMS bearer associated with TMGI/flow identifier for the MCPTT group without any modification.

5.3.5 QoE for MCPTT over MBMS

5.3.5.1 QoE for both MNO and MCPTT service provider

In TS 26.346 [11], the MBMS reception reporting server is configured by the BM-SC, the UE reports the QoE result to the MBMS reception reporting server. The QoE report is used by MNO to evaluate the MBMS user service and MBMS transport network performance.

The MCPTT server may be located outside of MNO network; it will benefit MNO if the current QoE reception report mechanism is kept to allow MNO evaluating the eMBMS transport layer performance. Besides the benefit to MNO, it is foreseen that the QoE result also helps MCPTT service provider evaluating the MCPTT service QoE via MBMS bearer.

It is proposed that:

- The MBMS client reports QoE to the network to benefit MNO for MBMS transport network performance evaluation.

- The MCPTT UE reports QoE to the MCPTT provider for MCPTT service quality evaluation with maximum re-use of existing MBMS QoE support.

5.3.6 eNB Scheduling on the MBMS Bearer

When carrying traffic over MBMS/MBSFN, the eNB schedules and transmits the packets received from the MCPTT AS according to the timestamp within the SYNC packet (see clause 15.3.3 in TS 36.300 [19]) corresponding to the MCH Scheduling Period (MSP), e.g., 40 or 80 ms, on the MBMS bearer.

It is recommended that a MCPTT UE that receives traffic over an MBSFN bearer uses a de-jitter buffer that can manage this type of jitter.

5.3.7 Needed information to describe an MCPTT User Plane

Rel 13 MCPTT is limited to support to Speech related media streams. Other media streams may be added in later releases.

In order to describe a speech related media stream over MBMS bearers, the device needs to receive at least the following information

- The number of media in the session.

- The destination IP address and port number for each RTP sessions in the MBMS bearer service.

- The protocol ID (e.g. RTP/AVPMedia type(s) and fmt-list.

- Mode of MBMS bearer per media.

- Multi-Carrier deployments should be supported. The UE needs to interpret SIB15 correctly.

- The start time and end time of the session.

- In case Source Specific IP Multicast (SSM) is required for MCPTT, also the Sender IP Address will be present.

- In order to support nation wide services through multiple PLMN-IDs, a list of alternative TMGIs may be present. Note, it is unclear whether this case is relevant for MCPTT.

Note, all above information except the information for Multi-Carrier support can be encoded in SDP.

6 Conclusion

In this document, many issues related to MCPTT, from media handling and codec perspective have been documented.

On the key issue "codec for MCPTT", analysis of various 3GPP codecs is provided, along with a recommendation for Release 13 to mandate that the AMR-WB codec be supported for MCPTT applications and, based on operator / MCPTT service provider policy, have EVS in super-wideband mode supported as an optional codec. The EVS codec was shown to provide statistically significant performance improvements relative to the AMR-WB codec for some of the MCPTT KPIs as studied and reported in this document

On the key issue "User Experience", it is recommended that the UE discards redundant RTP payloads in the case that the received RTP payloads by the UE contain identical SN and SSRC.

On the key issue "MCPTT over MBMS support", the following points have been made

- Realization considerations for MCPTT client operation in support of MBMS have been documented
- The media handling for unicast delivery mode and broadcast delivery mode has been documented
- It is recommended MCPTT UE uses a de-jitter buffer, when receiving traffic over an MBMS bearer
- The session information needed to describe the MCPTT user plane is provided.

Annex A: Simulation Models and Parameters

A.1 MBMS Bearer Simulation Model

A.1.1 Coverage

As described in clause 5.1.1.6.2.1, coverage using an MBMS bearer cannot guarantee a target FER throughout the entire coverage area. In SC-PTM, the limitations of geometry/SNR and the interference from adjacent cells limits the throughput and error rate that can be achieved.

Simulations of a single cell surrounded by two rings of adjacent cells that transmit at only 50% of their total unicast traffic were performed using the following simulation parameters:

Table A.1 Simulation parameters

	Freq (MHz)	Cell Radius (km)	Ant Height (Hb) (m)	Avg Clutter Height (m)	Dhb (m)	Slope	I	Avg EIRP (dBW) in 5 MHz	eNB Tx Pwr (dBW)
D1	2000	0.288	30	15	15	37.6	128.1	33	13

	UE Ant Loss (dB)	Impl. Loss (dB)	Log normal Shadowing	Down tilt (deg)	Noise Figure (dB)	Penetration Loss (dB)	Receive Height (m)	Vert Beam width (deg)	Horiz Beam width (deg)
D1	6	3	8	10	6	20	1.5	10	70

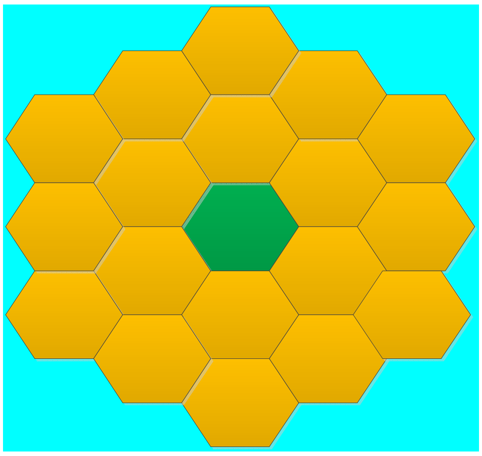


Figure A.1.1.-1 Serving cell (green) in the center is surrounded by the two rings of (yellow) cells that are only transmitting at 50% to reduce their interference on the serving cell.

The results of the simulation showed that even with the two rings of neighbouring cells reducing their load to 50%, the serving cell could not guarantee an FER <=1% for the entire cell (the coverage was less than 95% of the cell). The coverage in the serving cell would be even worse when the neighbouring cells are transmitting closer to their full load. Furthermore, other challenging environments such as indoor, basement, elevator, stairwell, etc.… reception would further impair the MBMS bearer error rate.

Therefore the MBMS error traces used to evaluate the performance of 3GPP speech codecs were chosen with FER values in the range of 1 to 5%.

A.1.2 Error Traces

During the Application Layer FEC work of Rel-12 an MBMS bearer model was defined and documented in [20] , clause 5.3. In communication with RAN1 and RAN2, it was agreed to use a two-state Markov model for the simulation of LTE RLC-PDU losses as shown in Figure A.1.2-1:

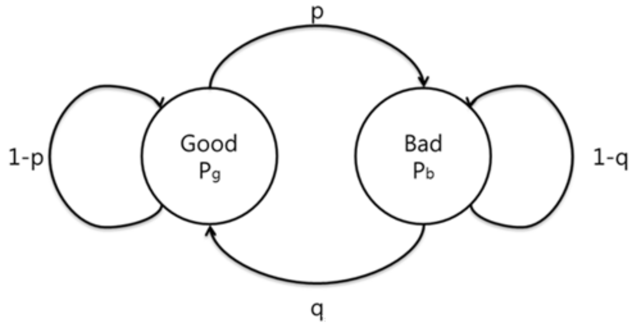


Figure A.1.2-1 Markov model for LTE RLC-PDU losses

The model was parametrized based on the D1 simulation settings of 3GPP TR 36.942..

Table A.1.2-1 Parameter Settings for MBMS LTE simulations

Parameter	Setting
Center Frequency (MHz)	2000
Cell radius (m)	288
Bandwidth (MHz)	5
Penetration Loss (dB)	20
Speed (km/h)	3
Antenna Down tilt (degree)	15
Antenna Height (m)	30
Antenna Clutter Height (m)	15
Dhb (m)	15
Slope	37.6
I	128.1
Average EIRP (dBW, 5MHz)	33
eNB Tx Power (dBW)	13
UE Antenna Loss (dB)	6
Implementation Loss (dB)	3
Noise Figure (dB)	6
Penetration Loss (dB)	20
Receiver Height (m)	1.5
Vertical Beamwidth (degree)	10
Horizontal Beamwidth (degree)	70

The simulation was carried out with a 19 sites configuration as shown in Figure A.1.2-2:. Each site has 3 cells. All sites have 100% SFN operation. 30 UEs are uniformly dropped into the center site (dark green one) in each simulation run of 50 sec. In total 900 UEs are dropped and the SNR is sampled accordingly. The overall SNR distribution is also shown in Figure A.1.2-2.

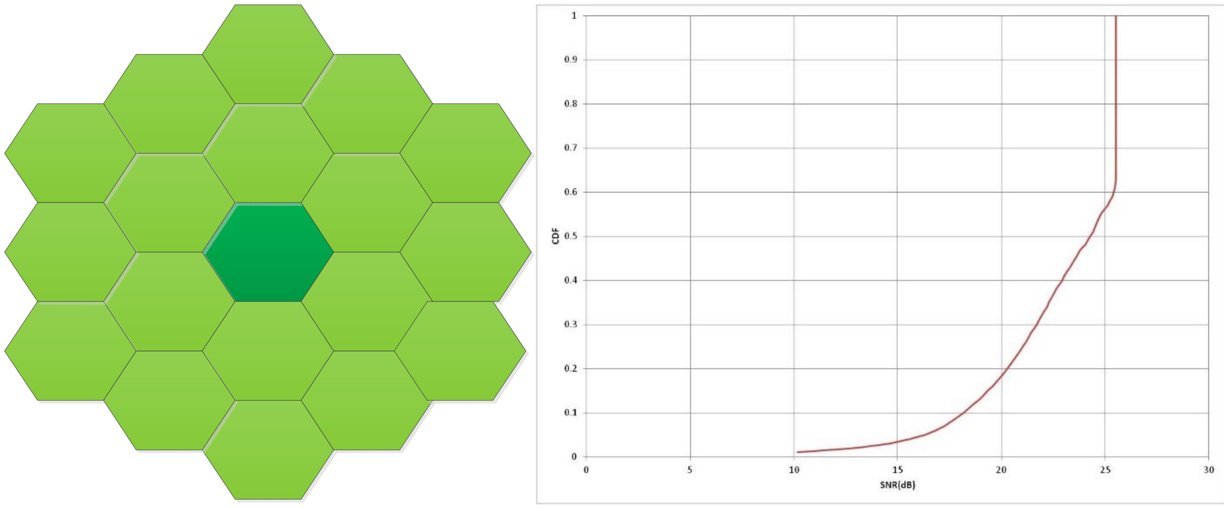


Figure A.1.2-2 Simulation Grid and SNR distribution

Based on those SNR traces, two representative traces were selected that in combination with MCS24 result in a 1%, 5%, 10% and 20% target BLER.

The parametrization of the Markov model is as follows:

- each state persists for 10ms, and
- a state is good if it has:
  - less than 10% packet loss probability for the 1% and 5% BLER simulations,
  - less than 40% packet loss probability for the 10% and 20% BLER simulations.
- MCS=24 was used for all cases and then users at different 'MBMS geometry' were picked to get the different average error rate.

The parameters for Markov channel modelling are provided in Table A.1.2-2.

Table A.1.2-2 Markov channel parameters

Parameter	Meaning
p	transition probability from Good state to Bad state
q	transition probability from Bad state to Good state
p_g	BLER in Good state
p_b	BLER in Bad state
$\frac{1}{\mathrm{p}}$	Average Length of Bad state segment
$\frac{1}{\mathrm{q}}$	Average length of Good state segment

The time in a good state Tg or time in a bad state Tb may be computed by multiplying the average length of a good (bad) segment by the sampling period. The probability of the good state and probability of a bad state may be computed as q/(p+q) and p/(p+q), respectively.

Specifically, the following parameters for the LTE MBMS channel simulations:

- MCS=9 and MCS=21 with 498 byte RLC-SDU size and 1332 byte RLC-SDU size.
- RLC-SDU distance of 10ms and 40ms for MCS=21

- RLC-SDU distance of 10ms for MCS=9
- Channel model with Markov model loss rate of 1%, 5%, 10% and 20% target BLER with speed 3 kph in Table A.1.2-3.
- Channel model with Markov model loss rate of 1%, 5%, 10% and 20% target BLER with speed 120 km/h in Table A.1.2-4.

Table A.1.2-3: Markov parameters for 3 km/h

Table 1 3 km/h				
	BLER = 1%	**BLER = 5%**	**BLER = 10%**	**BLER = 20%**
p	0.58%	1.80%	2.79%	4.61%
q	36.13%	24.01%	20.90%	16.80%
sg	98.42%	93.02%	88.23%	78.48%
sb	1.58%	6.98%	11.77%	21.52%
pg	0.03%	0.06%	0.56%	1.16%
pb	59.47%	70.54%	82.30%	89.20%
BLER	0.97%	4.98%	10.19%	20.12%
Tg (ms)	1724	555	359	217
Tb (ms)	28	42	48	60

Table A.1.2-4: Markov parameters for 120 km/h

Table 2 120 km/h				
	BLER = 1%	**BLER = 5%**	**BLER = 10%**	**BLER = 20%**
p	6.06%	27.07%	46.48%	35.60%
q	94.30%	70.95%	50.95%	63.29%
sg	93.97%	72.39%	52.29%	64.00%
sb	6.03%	27.61%	47.71%	36.00%
pg	0.00%	0.00%	0.00%	9.72%
pb	17.31%	19.54%	22.33%	40.40%
BLER	1.05%	5.40%	10.66%	20.77%
Tg (ms)	165	37	22	28
Tb (ms)	11	14	20	16

Regarding the MCS selection, the optimum operating MCS strongly depends on the deployment scenario, including site-to-site distance, operating frequency, interference conditions at MBSFN area boundaries, etc. Therefore, one specific value is not suitable. Using two different MCS cases can give some diversity in the assumptions, hence a good approach to use the following two values:

- higher value MCS=21 resulting in RLC-SDU size of 1332 byte.
- lower value corresponding to 1 bit/s/Hz, with MCS=9 resulting in RLC-SDU size of 498 byte.

It is additionally from the following list of available simulation conditions the following were selected as a good candidate representative:

- RLC-SDU distance of 10 ms and 40 ms for MCS=21
- RLC-SDU distance of 10 ms for MCS=9

However, in call cases above, the RLC-SDU size is sufficiently large to contain multiple speech frames. Therefore, the focus in the simulation is on the loss patterns and delay. We focus on 10ms and 40ms RLC-SDU distance in the following. Also, the focus for the scenario here is on error rates of at most 5%, as it is believed that higher error rates are unrealistic if no Application Layer FEC is applied, which is the case for MCPTT speech due to tight latency requirements.

In order to generate appropriate error patterns, the tools attached to TR 26.947 and as described in Annex B.2 are used. Specifically the following error patterns are generated.

```
java LossVectorGenerator 0.0058 0.3613 0.0003 0.5947 1 44000 0 0 errortrace_v3_b1_10ms_44000_0.txt
java LossVectorGenerator 0.0058 0.3613 0.0003 0.5947 1 44000 1 0 errortrace_v3_b1_10ms_44000_1.txt
java LossVectorGenerator 0.0058 0.3613 0.0003 0.5947 1 44000 2 0 errortrace_v3_b1_10ms_44000_2.txt
java LossVectorGenerator 0.0058 0.3613 0.0003 0.5947 1 44000 3 0 errortrace_v3_b1_10ms_44000_3.txt
java LossVectorGenerator 0.0058 0.3613 0.0003 0.5947 1 44000 4 0 errortrace_v3_b1_10ms_44000_4.txt
java LossVectorGenerator 0.0606 0.9430 0.0000 0.1731 1 44000 0 0 errortrace_v120_b1_10ms_44000_0.txt
java LossVectorGenerator 0.0606 0.9430 0.0000 0.1731 1 44000 1 0 errortrace_v120_b1_10ms_44000_1.txt
java LossVectorGenerator 0.0606 0.9430 0.0000 0.1731 1 44000 2 0 errortrace_v120_b1_10ms_44000_2.txt
java LossVectorGenerator 0.0606 0.9430 0.0000 0.1731 1 44000 3 0 errortrace_v120_b1_10ms_44000_3.txt
java LossVectorGenerator 0.0606 0.9430 0.0000 0.1731 1 44000 4 0 errortrace_v120_b1_10ms_44000_4.txt
java LossVectorGenerator 0.0180 0.2401 0.0006 0.7054 1 440000 0 errortrace_v3_b5_10ms_44000_0.txt
java LossVectorGenerator 0.0180 0.2401 0.0006 0.7054 1 44000 1 0 errortrace_v3_b5_10ms_44000_1.txt
java LossVectorGenerator 0.0180 0.2401 0.0006 0.7054 1 44000 2 0 errortrace_v3_b5_10ms_44000_2.txt
java LossVectorGenerator 0.0180 0.2401 0.0006 0.7054 1 44000 3 0 errortrace_v3_b5_10ms_44000_3.txt
java LossVectorGenerator 0.0180 0.2401 0.0006 0.7054 1 44000 4 0 errortrace_v3_b5_10ms_44000_4.txt
java LossVectorGenerator 0.2707 0.7095 0.0000 0.1954 1 44000 0 0 errortrace_v120_b5_10ms_44000_0.txt
java LossVectorGenerator 0.2707 0.7095 0.0000 0.1954 1 44000 1 0 errortrace_v120_b5_10ms_44000_1.txt
java LossVectorGenerator 0.2707 0.7095 0.0000 0.1954 1 44000 2 0 errortrace_v120_b5_10ms_44000_2.txt
java LossVectorGenerator 0.2707 0.7095 0.0000 0.1954 1 44000 3 0 errortrace_v120_b5_10ms_44000_3.txt
java LossVectorGenerator 0.2707 0.7095 0.0000 0.1954 1 44000 4 0 errortrace_v120_b5_10ms_44000_4.txt
```

Subsampling of the traces from 10ms to 40ms is done by dropping 3 out of 4 packets from the error trace.

A.1.3 eNB Scheduling

The eNB has scheduling opportunities on the MBMS bearer every 40ms and transmits all the packets it has received from the MCPTT AS at the next scheduling opportunity. The eNB is not specified to have a de-jitter buffer and can be considered to forward whatever packets it has received in the last 40ms to the UE, without re-ordering or further buffering of the packets.

Although a de-jitter buffer is not specified at the eNB, the simulations used introduced two cases: with and without de-jitter buffering at the eNB.

Case1 eNB dejitter buffer: The uplink channel conditions were simulated with a 1% delay-error profile and the eNB includes a dejitter buffer to compensate for jitter in the uplink.

The downlink MBMS bearer channel was simulated using the profiles described in the previous clause for 120 km/h and 3 km/h velocities, and copied into Table A.1.3.1. The channel models for error rates of 3.22% and 3.85% at 120km/h were produced using a random error model since [20] did not provide traces at these error rates and it was confirmed that the errors are very uncorrelated as this velocity.

Table A.1.3-1: Downlink MBMS Error Profiles for different coverage areas [20]

Downlink Channel Profile	Speech / Fading conditions	Error rate
v120_b1	120 kmph	0.88 %
v120_b3	120 kmph	3.22 %
v120_b4	120 kmph	3.85 %
v120_b5	120 kmph	5.70 %
v3_b1	3 kmph	0.81 %
v3_b5	3 kmph	4.73 %

The downlink MBMS channel packets are scheduled at 40 ms intervals and there is a playout buffer of length 60ms in the MCPTT UE. If the scheduler does not have any of the two 20 ms VoIP packets that are supposed to send in a particular 40 ms interval, it does not substitute the missing packet with a future packet.

Case2 no eNB dejitter buffer: Two uplink scenarios are considered here under this case.

a) uplink channel conditions were simulated with same 1% delay-error profile as in case 1,

b) uplink channel conditions were simulated with 3% error profile

In both cases, no de-jitter buffer is present in the eNB, all packets arriving with variable delay jitter in the uplink channel are scheduled on the downlink MBMS bearer channel at the next 40ms scheduling opportunity interval. The downlink MBMS bearer channel was simulated using the same profiles described in Table A.1.3-1 with various error rates.

The simulations use a de-jitter buffer at the MCPTT UE to compensate for the jitter introduced in both the uplink and downlink path from the talker to the listener.

A.2 Correlation between Subjective MOS and P.OLQA

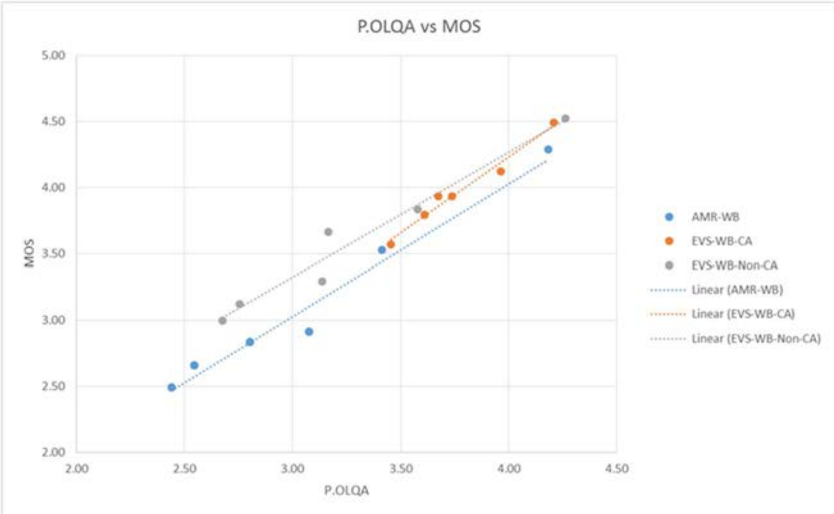


Figure A.2-1: Correlation between Subjective MOS and P.OLQA for AMR-WB, EVS Channel-Aware, and EVS Non-Channel-Aware

Annex B: Change history

TR 26.879 (WITHDRAWN)

Date	TSG #	TSG Doc.	CR	Rev	Cat	Subject/Comment	New version
2015-09	69	SP-150453				Presented to TSG SA#69 (for information)	1.0.0
2015-12	70	SP-150665				Presented to TSG SA#70 (for approval)	2.0.0
2015-12	70	SP-150813				Updates in Clauses 2 and 5.1.1.1 done during TSG SA#70. Presented to TSG SA#70 (for approval)	2.0.1
2015-12	70					Version approved for Release 13	13.0.0
2016-06	72	SP-160265	0004	2	F	Evaluation of listening effort of MCPTT candidates	13.1.0
2016-06	72	SP-160265	0006		F	Removing References not in the Public Domain	13.1.0
2016-06	72	SP-160265	0007	1	F	Reflect the Final Codec Choices for MCPTT	13.1.0
2016-06	72	SP-160265	0008	2	F	Criteria to be considered with respect to MCPTT codec selection	13.1.0
2016-06	72	SP-160265	0009	1	F	Update to Technical Report to include Speech Intelligibility Results	13.1.0

TR 26.989

Change history							
Date	**Meeting**	**TDoc**	**CR**	**Rev**	**Cat**	**Subject/Comment**	**New version**
2016-06	72					Decision at TSG SA#72: TR 26.879 is withdrawn and a new TR number is allocated as TR 26.989	13.1.0
2017-03	75					Version for Release 14	14.0.0
2018-06	75					Version for Release 15	15.0.0

History

Document history		
V15.0.0	July 2018	Publication