Technical Specification Group Services and System Aspects **TSGS#10(00)0555**
Meeting #10, Bangkok, Thailand, 11-14 December 2000 Agenda Item: 7.4.3

Source: **TSG-SA WG4**

Title: **Results of AMR Wideband (AMR-WB) Codec Selection Phase**

Document for: **Approval**

Agenda Item: **7.4.3**

1. Introduction

TSG-S4 has carried out development and standardisation of Adaptive Multi-Rate Wideband (AMR-WB) codec since 1999. The work item on wideband coding was approved at TSG-SA#2 (March 1999) [1]. The effective start of the work was pending on the results of Wideband Feasibility Study carried out in ETSI SMG11 by June 1999 [2]. The work was then carried out jointly by S4 and SMG11 under a common S4/SMG11 work item. The common harmonised WI description was approved in ETSI SMG#29 (June 1999) and in TSG-SA#5 (October 1999) [3]. After more than one year, the AMR-WB codec development and selection have now been completed in time, and the results are brought for approval at TSG-SA#10.

The AMR-WB codec brings quality improvement over the existing narrowband telephony through the use of extended audio bandwidth. The Adaptive Multi-Rate (AMR) narrowband codec, standardised for GSM Release 98 and 3GPP Release 99, provides good performance for telephone bandwidth speech (audio bandwidth limited to 3.4 kHz). However, the introduction of a wideband speech service (audio bandwidth extended to 7 kHz) brings improved voice quality especially in terms of increased voice naturalness. Wideband coding brings speech quality exceeding that of (narrowband) wireline quality to 3G and GSM systems.

The AMR-WB codec, like the AMR narrowband codec, is an adaptive codec consisting of several modes one of which is chosen based on the operating conditions on the radio channel. By adapting coding depending on the channel quality, high robustness against transmission errors is obtained. The codec also includes a source controlled rate operation mechanism, which allows it to encode speech at a lower average rate by taking speech inactivity into account. The AMR-WB codec is designed to operate in multiple applications: in GSM full-rate radio traffic channel (with and without an additional constraint of compatibility with 16 kbit/s A-ter sub-multiplexing), in Circuit Switched EDGE/GERAN 8-PSK Phase II radio channels (full- and half-rate), and in 3G UTRAN radio channel.

The AMR-WB codec selection has been carried out in S4 as a competitive selection process consisting of two phases: a Qualification (Pre-Selection) Phase and a Selection Phase. Seven AMR-WB candidate codecs were submitted for the Qualification Phase. One candidate was later withdrawn and the remaining six were accepted at TSG-SA#8 in June 2000 to proceed into the Selection Phase [4]. Since then two proponents have joined their codec development reducing the number of codec candidates to five for the Selection Phase. The codecs that participated into the Selection Phase came from Ericsson, FDNS consortium (consisting of France Télécom, Deutsche Telekom, Nortel Networks and Siemens), Motorola, Nokia and Texas Instruments.

The Selection Phase was carried out from July to October 2000. The results were reviewed, analysed and debated during S4#13 in October 2000. Recommendation for the codec to be selected as the AMR-WB codec was made, and this is brought for approval at TSG-SA#10. At S4#13, a Verification Phase was started to verify some of the key characteristics of the best solution. Also, a set of draft AMR-WB specifications have been prepared. These were reviewed at S4#14 in November 2000 and are brought for information to TSG-SA#10 [11-19]. The AMR-WB specifications will be brought for approval at TSG-SA#11.

2. Selection Phase Testing

The five candidate codecs were tested in a variety of test conditions in six independent test laboratories. Testing was carried out using 5 languages (French, Japanese, Mandarin Chinese, North American English, and Spanish). The tests took place during a period from September to October 2000 and they were completed on schedule. During the testing, the performances of the candidate codecs were evaluated in multiple of test conditions consisting of 6 experiments and 19 sub-experiments [5, 6].

The testing covered the four applications specified for the AMR-WB codec[1]:

Application A: GSM full-rate traffic channel with an additional constraint of 16 kbit/s A-ter sub-multiplexing
Application B: GSM full-rate traffic channel
Application C: Circuit Switched EDGE/GERAN 8-PSK Phase II radio channels
Application E: 3G UTRAN WCDMA radio channel

[1] Letter "D" was reserved for an intended GSM multi-slot application. However, this was not found necessary and was withdrawn later during standardisation.

The experiments and sub-experiments included in the selection tests are as follows [5]:

1. Experiment 1: Input Level and tandeming performance for clean speech (ACR-test[2])
   - 1a: Applications A and B
   - 1b: Applications C and E
2. Experiment 2: Clean Speech performance with static errors (ACR)
   - 2a: Clean Speech and in Static Errors for GSM FR Channel (Application A)
   - 2b: Clean Speech and in Static Errors for GSM FR Channel (Application B)
   - 2c: Clean Speech and in Static Errors for Higher-Rate Channels (Application C)
   - 2d: Clean Speech and in Static Errors for Higher-Rate Channels (Application E)
   - 2e: Clean Speech and in Static Errors for GSM EFR and wideband to narrowband tandeming
3. Experiment 3: Car and Street noise (15 dB SNR) performance for the GSM FR channel (DCR-test)
   - 3a: GSM FR channel (Application A) in Car noise
   - 3b: GSM FR channel (Application A) in Street noise
   - 3c: GSM FR channel (Application B) in Car noise
   - 3d: GSM FR channel (Application B) in Street noise
   - 3e: GSM EFR performances in Car and Street noise
4. Experiment 4: Car and Street noise (15 dB SNR) performance for higher-rate channels (DCR-test)
   - 4a: Higher-rate channels (Application C) in Car noise
   - 4b: Higher-rate channels (Application C) in Street noise
   - 4c: Higher-rate channels (Application E) in Car noise
   - 4d: Higher-rate channels (Application E) in Street noise
5. Experiment 5: Performance in Dynamic Conditions (ACR-test)
   - 5a: Performance in Dynamic Conditions for AMR-WB (Application A)
   - 5b: Performance in Dynamic Conditions for EFR
6. Experiment 6: VAD/DTX in GSM FR channel for Application B (CCR-test)

The listening test laboratories participating into the AMR-WB selection tests were: ARCON (North American English), AT&T (Mandarin Chinese, North American English, Spanish), Dynastat (North American English, Spanish), France Télécom (French), Lockheed-Martin Global Telecommunications (North American English, Spanish), and NTT-AT (Japanese). Each experiment in the tests was carried out with two languages to avoid any bias due to a particular language. The allocation of experiments to listening laboratories, and the languages used for each experiment, are shown in Table 1.

Processing of speech samples through the candidate algorithms was carried out by the candidate organisations themselves and was cross-checked for correctness by other candidates. Two host laboratories, ARCON and Lockheed-Martin Global Telecommunications processed the samples through reference codecs. A blind procedure was followed to ensure that the listening test laboratories and the test subjects had no knowledge of the codec algorithms. The test results from the individual laboratories were combined by a Global Analysis Laboratory (ARCON) and were presented at S4#13 in October 2000.

Before the processing of speech samples started the candidates had to deliver, in early August, an executable of their codec software to ETSI freezing the algorithm development. The processing of speech samples was carried out during August and early September 2000. Listening tests started in mid-September. The listening test results and deliverables from the codec proponents (technical descriptions of the codec algorithms) were reviewed at S4#13. Nortel Networks provided the error patterns required in the testing for Applications A, B and C. The error patterns for testing of Application E were provided by Ericsson (Uplink) and Nokia (Downlink). The seed-values of the error patterns were kept secret during testing.The key milestones of the listening tests and the relating selection phase activities are shown in Table 2.

[2] Experiments 1, 2 and 5 are Absolute Category Rating (ACR) tests, experiments 3 and 4 are Degradation Category Rating (DCR) tests, and experiment 6 is a Comparison Category Rating (CCR) test. The results are given as Mean Opinion Scores (MOS), Differential MOS (DMOS), or Comparison MOS (CMOS), respectively. ACR tests ask the listeners to assess the quality of each speech sample under test while DCR and CCR tests ask the listeners to assess the quality differences between two samples. The difference between DCR and CCR tests is that in DCR tests the listeners assess the degradation in the second sample compared to the first one, while in CCR tests the listeners assess the quality difference between the samples. (ACR, DCR and CCR tests are all well established and recognised speech quality testing methodologies. These methodologies are used within the experiments, depending on which is the most suitable one for each test.)

Experiment	ARCON	AT&T	Dynastat	FT	LMGT	NTT-AT	Total of languages
1a	NAE			FR			2
1b	NAE			FR			2
2a			NAE			JP	2
2b			NAE			JP	2
2c			NAE			JP	2
2d			NAE			JP	2
2e			NAE			JP	2
3a		SP			NAE		2
3b		SP			NAE		2
3c		MCH			NAE		2
3d		MCH			NAE		2
3e			SP		NAE		2
4a		NAE			SP		2
4b		NAE			SP		2
4c			NAE		SP		2
4d			NAE		SP		2
5a		NAE		FR			2
5b		NAE		FR			2
6	NAE					JP	2
Total of sub-experiments	3	8	8	4	9	6	38

Note: NAE: North American English; MCH: Mandarin Chinese; SP: Spanish; FR: French; JP: Japanese

***Table 1**: Allocation of Experiments to the Listening Laboratories.*

Responsible	Action Description	Deadline
Test laboratories	Delivery of the speech samples to the host laboratories for processing	July 31st
Candidates	Receipt of executables for AMR-WB candidates by ETSI	August 6th
Candidates	Send executables, processed material etc to the cross-checking candidate, and to the host laboratory (without the executable).	August 24th
Candidates	Completion of processing and verification of correctness	August 28th
Host Laboratories	Sending of final set of speech material to test laboratories	September 13th
Candidates	Delivery of all remaining Selection Deliverables (technical descriptions of candidate algorithms, analysis of compliance to design constraints etc.) to ETSI	October 18th
Candidates	Delivery of complete IPR declaration to ETSI	October 8th
Test laboratories	End of listening tests	October 9th
Test laboratories	Delivery of test results (test raw data) to ETSI and Global Analysis Laboratory	October 9th
Global Analysis Laboratory	Preparation and delivery of test results summary / technical report to the S4-reflector	October 16th
Host and listening laboratories	Presentation of test results to S4	S4#13 (October 23rd –27th)
S4	Review of the selection test results, recommendations for the codec to be chosen	S4#13 (October 23rd –27th)
S4	Review of draft specifications and first verification results	S4#14 (Nov 27th – Dec 1st)
S4	Presentations of Selection Test results and AMR-WB codec selection for approval. Presentation of AMR-WB draft specifications for information.	TSG-SA#10
S4	Presentation of AMR-WB specifications for approval.	TSG-SA#11

***Table 2**: Key milestones of the AMR-WB Selection Phase.*

3. Review of AMR-WB Performance Requirements

In Application A, the speech coding rate is restricted below 14.4 kbit/s, while in Application B rates up to the GSM FR transmission channel bit-rate of 22.8 kbit/s are possible. Due to this restriction, Application B can provide better maximum quality (at low error-rate conditions) than Application A.
In Application A, the general quality requirement is to be better than ITU-T G.722 wideband codec at 48 kbit/s (G.722-48k). In Application B, quality equal to G.722-56k is required. For applications C and E an even higher quality requirement is set requiring quality to be equal to G.722-64k. These are general requirements for clean channel performance (no transmission errors). Under the impact of background noise, relaxation is allowed in some cases (e.g., in Application A quality equal to G.722-48k is required in tandem conditions under background noise). In erroneous transmission, the codec should be robust against transmission errors. An illustrative diagram of the setting of quality requirements is given in Figure 1 [4].
The requirements are explained in detail in the following sections. A full description of the performance requirements can be found in [7].

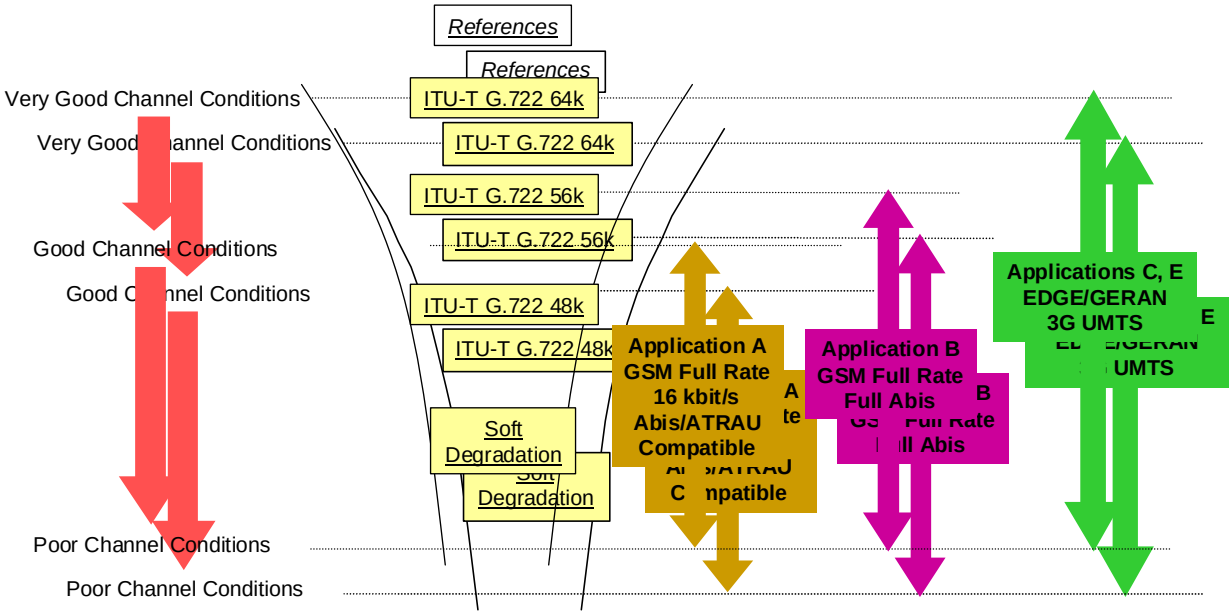


Figure 1: Quality requirements for the AMR-WB codec for the various applications [4].

3.1. GSM FR Channel (Applications A and B)

For clean speech, at 19 dB C/I and above, the AMR-WB codec is required to provide in Application A quality better than (error-free) G.722-48k, and in Application B quality equal to G.722-56k. At 13 dB C/I, quality should still be equal to (error-free) G.722-48k in both applications. Under 13 dB C/I, graceful degradation comparable to the performance demonstrated by GSM EFR (Enhanced Full Rate) codec is required. Table 3a shows the requirements for clean speech.

Clean speech	**Application A: GSM FR with 16 kbit/s subultiplexing**		**Application B: GSM FR**	
C/I	**Performance requirement**	**Performance objective**	**Performance requirement**	**Performance objective**
no errors	better than G.722-48k	G.722-56k	G.722-56k	G.722-64k
19 dB	better than G.722-48k		G.722-56k	
16 dB	G.722-48k		G.722-48k	
13 dB	G.722-48k		G.722-48k	
< 13dB	(see Note 1)		(see Note 1)	

Note 1: The degradation in subjective performance shall not be greater than the degradation in subjective performance demonstrated by EFR over the same C/I interval. The specific intervals of interest are 13dB to 10dB, 13dB to 7dB, and 13dB to 4dB.

Table 3a: *Clean speech requirements under static error conditions for Applications A and B.*

For background noise conditions (speech in background noise), the requirements are given in Table 3b. The requirements are the same as for clean speech except that quality equal to G.722-48k is required for Application A at C/I ≥ 19 dB. (Also, a different testing methodology, Poor or Worse, considered more suitable for background noise testing, was adopted[3].)

[3] Poor or Worse methodology is employed, where "with 10% PoW" is interpreted as no more than 10 additional percentage points of annoying degradation with respect to the reference codec (in terms of annoying or very annoying quality scores in the listening tests: "1" and "2" out of votes ranging from "1" to "5").

Speech in background noise	Application A: GSM FR with 16 kbit/s submultiplexing		Application B: GSM FR	
C/I	Performance requirement	Performance objective	Performance requirement	Performance objective
no errors	G.722-48k (with 10% PoW)	G.722-56k	G.722-56k (with 10% PoW)	G.722-64k
19 dB	G.722-48k (with 10% PoW)		G.722-48k (with 10% PoW)	
16 dB	G.722-48k (with 10% PoW)		G.722-48k (with 10% PoW)	
13 dB	G.722-48k (with 10% PoW)		G.722-48k (with 10% PoW)	
< 13dB	See Note 1 (in Table 3a)		See Note 1(in Table 3a)	

Table 3b: *Background noise requirements under static error conditions for Applications A and B.*

In tandem (2 asynchronous encodings), the requirement for AMR-WB for both clean speech and background noise is to be equal to G.722-48k in tandem for Application A and equal to G.722-56k in tandem for Application B. For input level dependency, for clean speech, the general requirement is to be better than G.722-48k for Application A and equal to G.722-56k for Application B. For talker and language dependency, the requirement is to provide in Application A the same quality as G.722-48k and in Application B the same quality as G-722-56k.
For Applications A and B, requirements were set also for dynamic conditions (codec operated with mode adaptation on). Under typical dynamic error conditions, the requirement is to be better than EFR under the same error conditions. For difficult error conditions (6 dB worse than typical C/I-conditions), the requirement is to be at least as good as the EFR codec in the same conditions.

3.2. Higher Rate Channels (Applications C and E)

In the EDGE HR-channel, for clean speech and speech in background noise, AMR-WB should give at 25 dB C/I and above quality equal to (error-free) G.722-56k. At 19 dB C/I, quality should still be equal to (error-free) G.722-48k. In the EDGE FR-channel, the same quality as in the HR-channel should be obtained at 3 dB worse C/I conditions.
In the 3G UTRAN channel, AMR-WB should give in error-free transmission quality equal to (error-free) G.722-64k. Quality equal to (error-free) G.722-48k is required at FER=1.0% / RBER=0.1%.
The requirements for Application C are given in Table 4a and for Application E in Table 4b.

Clean speech and speech in background noise	Application C: Half-Rate Circuit Switched EDGE Phase II channel	Application C: Full-Rate Circuit Switched EDGE Phase II channel)
C/I	Performance requirement	Performance requirement
25 dB	G.722-56k	
22 dB	G.722-48k	G.722-56k
19 dB	G.722-48k	G.722-48k
16dB		G.722-48k

Table 4a: *Requirements for clean speech and background noise under static test conditions for Application C.*

Clean speech and speech in background noise	Application E: 3G UTRAN channel	
Error Condition [FER, RBER]	Performance requirement	Performance objective
No errors	G.722-64k	
[0.5%, -]	G.722-56k	
[1.0%, 0.1%], Uplink (Note 1)	G.722-48k	
[1.0%, 0.1%], Downlink (Note 1)	G.722-48k	
[1.0%, 0.1%], Uplink (Note 2)		G.722-48k

Note 1: The least significant bits shall be subjected to the residual error profile. The number of bits in this class shall be 25% of the total bits per frame.
Note 2: The least significant bits shall be subjected to the residual error profile. The number of bits in this class shall be 50% of the total bits per frame.

Table 4b*: Requirements for clean speech and background noise under static test conditions for Application E.*

Application E includes all bit rates. The requirements are however only tested for the highest modes. The error performance for Application E is specified and evaluated using error protection schemes from the UTRAN toolbox. Each error condition is defined using two error profiles, one FER profile (single indicator per frame) and one residual BER profile (bit-level residual error channel). The requirement for the no error case applies to modes with higher bit rates, i.e., those not tested in Applications A and B.
For both Application C and E, in tandem (2 asynchronous encodings), the requirement for clean speech is to be equal to G.722-64k in tandem, and in background noise to be equal to G.722-56 in tandem. For input level

dependency, for clean speech, the general requirement is to be equal to G.722-64k. For talker and language dependency, equal performance to G.722-64k is required.

3.3. Other Requirements and Objectives

The following Tables summarise some additional requirements set for the AMR-WB codec: source controlled operation in the DTX mode (discontinuous transmission), non-speech inputs and music.

Condition	Requirement
Switching between different AMR-WB bit-rates	No annoying artefacts
Clean speech with DTX enabled	Performance with DTX disabled
Speech and background noise with DTX enabled	Performance with DTX disabled

Table 5a: *Additional performance requirements for speech signals in source controlled operation (all applications).*

Condition	Requirement	Objective
DTMF		Transparent transmission of DTMF
Information tones	Recognisable as given information tone.	
Idle noise	-66dBm0 (unweighted)	

Table 5b*: Requirements and objectives for speech codec performance with non-speech inputs (all applications).*

Condition	Requirement	Objective
Music	No annoying effects	G.722-56k

Table 5c: *Requirements and objectives with music for Applications C and E.*

3.4. Testing of Performance Requirements in the Selection Tests

The selection tests were extensive consisting of altogether 6 experiments and 19 sub-experiments and covering all the four applications defined for AMR-WB. All above mentioned performance requirement conditions were included in the testing except only a few ones considered less critical for the selection (e.g., testing in tandem under background noise, switching between different AMR-WB bit-rates, and testing with non-speech signals and music). These were excluded for practical reasons to keep the selection tests within a reasonable size and will be covered during the verification phase testing and/or in the characterisation tests.

4. Review of the Selection Procedure

The selection procedure consisted of comparing the performances of the candidate codecs against a set of performance requirements and ranking the candidate performances using a number of Figures of Merit (FoM). Technical descriptions and other deliverables from the proponents were also reviewed and compliance with a set of mandatory design constraints was analysed.

4.1. Selection Procedure

The Selection Procedure followed the pre-defined selection rules described in [8]. The selection procedure consisted of the following steps:

1. The selection test results will be presented and analyzed while keeping secret the identity of the candidates. Each candidate will be informed of the code used for its own solution and its solution only. The selection rules 2a, 2b and 3 will be applied at this stage.
2. After the review and discussion of the test results (as specified for rule 3), TSG-S4 will try to reach a consensus on a quality ranking of the candidates.
3. Each candidate will then present its solution and show the compliance with the design constraints. All candidates not compliant with all design constraints will be excluded according to the selection rule 1.
4. The test results obtained by each candidate will then be revealed.
5. A final discussion and review of the solution characteristics and test results will take place.
6. S4 will then try to reach a consensus on a single candidate to serve as the basis for the AMR-WB standardization.

The first two selection rules are eliminating rules. The first rule excludes all candidates failing to demonstrate full compliance with the AMR-WB design constraints. The second rule excludes all candidates with test results too far below the expected performance level. The third rule consists of a direct comparison between candidates using a set of Figures of Merit.

4.2. Design Constraints (Rule 1)

Design constraints are a set of mandatory requirements that the AMR-WB codec needs to fulfil. Any candidate codec not compliant with all design constraints is excluded from selection. The design constraints include constraints, e.g., for implementation complexity and transmission delay.

The computational complexity of the speech codec (without channel coding) was limited below 40 wMOPS for all applications. For speech coding and channel coding (Applications A and B), the detailed complexity limits are given in Annex 3. For Application C, the definition of the channel is being carried out in TSG-GERAN. However, for the purposes of AMR-WB selection tests, the codec proponents had to provide an example channel codec solution complying with a number of constraints as shown in Annex 3. Application E was tested with residual error patterns (impacting the bit-stream from/to speech codec), and the proponents did not therefore need to provide channel codec as part of the proposal.
The algorithmic transmission delay requirement was set for the GSM FR channel, where the same delay as in AMR narrowband codec was required but with 6.5 ms relaxation. The relaxation is needed because of the increased Abis/Ater delay (caused by the higher speech coding bit-rates) and also due to allowing the use of band-splitting and re-composition filters in the solutions, as felt necessary for wideband coding.
The proponents were required to provide for the Selection Phase, a fixed-point C-code implementation of the proposed AMR-WB codec. This consisted of speech codec (including voice activity detection and source controlled rate mechanism) for all applications, channel coding for the GSM FR channel, and example channel codings for EGDE FR and EDGE HR channels.
The same codec mode and channel measurement signalling scheme as used in AMR narrowband was required to be used. Also, the same source controlled rate scheme with regard to transport format and update frequency as in AMR narrowband was a requirement.
The design constraints are explained in detail in [9].

4.3. Unacceptable Low Performance (Rule 2)

This rule is an eliminating rule to exclude all candidates with performance too far below the expected performance level. The rule consists of two parts: Rule 2a checks that more than 50% of the performance requirements were met for various subsets of the tests. Rule 2b checks that there were no more than 10% of severe failures for each of the subsets.
Selection Rule 2a: Any candidate failing 50% or more of the test conditions contained in any of the following test sets will be excluded. A test is failed if the codec performance (measured MOS score or PoW) does not meet the requirement specification at the 95% confidence level.

List of test sets for Rule 2a:

- Set #1: all conditions (90), including the CCR Tests
- Set #2: all clean conditions (47)
- Set #3: all background noise conditions (43), including the CCR Tests
- Set #4: all conditions of application A (30)
- Set #5: all conditions of application B (26), including the CCR Tests
- Set #6: all conditions of application C, D, E (34)

Selection Rule 2b: Any candidate severely failing more than 10% of the test conditions contained in any of the following test sets will be excluded.

List of test sets for Rule 2b:

- Set #1: all conditions (87), excluding the CCR Tests
- Set #2: all clean conditions (47)
- Set #3: all background noise conditions (40), excluding the CCR Tests
- Set #4: all conditions of application A (30)
- Set #5: all conditions of application B (23), excluding the CCR Tests
- Set #6: all conditions of application C, D, E (34)

4.4. Direct Comparison of the Candidates (Rule 3)

Table 6 lists a number of Figures of Merit that were identified to be used to analyze and compare the performance of the candidates.

Metric	Ranking Provided
Weighted ΔdBq	Per experiment and across all experiments Per lab and across labs Full set of test results (Preferred FoM) and restricted to the failed tests only (ΔdBq computed with reference to the requirement in this case)
Weighted ΔMOS	Per Experiment and Per lab (cannot be computed across labs and experiments) Full set of test results and restricted to failed tests
Number of systematic failures (2 failures out of 2 tests)	Per Experiment and across all experiments Across labs
Unweighted ΔPoW percentages (for the relevant conditions)	Per Experiment and across all relevant experiments
Unweighted ΣCMOS (for the relevant conditions)	Per Experiment and across all relevant experiments

Table 6: *List of FoMs selected for the evaluation of the test results.*

Details on how rules 2 and 3 are applied can be found in [8].

5. Analysis of the Listening Test Results

The Selection Phase test results were presented and analyzed at S4#13 while keeping the identity of the candidates secret. The candidates were referred to as Codec 1, Codec 2,..., Codec 5 during the analysis to prevent any bias. These were later revealed to be:

Codec 1 = Ericsson

Codec 2 = FDNS consortium

Codec 3 = Nokia

Codec 4 = Motorola

Codec 5 = Texas Instruments

During the selection discussion, Codec 4 was withdrawn. The codec development had not been fully completed by the deadline (of August 6th) and the proponent had delivered a non bit-exact version of the software into the testing.

5.1. Performance Requirements (Rule 2)

5.1.1 Rule 2a (less than 50% failures in each set)

The candidate performances were analysed in accordance to the selection Rule 2. The number of failures for each subset of conditions is given in Tables 7a and 7b.

Rule 2A	Candidate Failures in Set#1					Candidate Failures in Set#2					Candidate Failures in Set #3				
Codec #	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Number of failures	17	29	0	13	11	6	5	0	3	3	11	24	0	10	8
Failure-%	10,6	18,1	0,0	8,1	6,9	8,1	6,8	0,0	4,1	4,1	12,8	27,9	0,0	11,6	9,3
Pass / Fail	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass

Table 7a*: Number of failures for sets #1 - #3.*

Rule 2A	Candidate Failures in Set#4					Candidate Failures in Set#5					Candidate Failures in Set#6				
Codec #	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Number of failures	4	8	0	5	3	2	3	0	4	4	11	18	0	4	4
Failure-%	9,1	18,2	0,0	11,4	6,8	4,5	6,8	0,0	9,1	9,1	16,7	27,3	0,0	6,1	6,1
Pass / Fail	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass	Pass

Table 7b: *Number of failures for sets #4 - #6.*

All candidates met the requirement of Rule 2a. For Codec 3, no failures against the performance requirements were found at all in any of the tests.

5.1.2 Rule 2b (10% or less severe failures in each set)

All codec candidates met Rule 2b. None of the candidate codecs had severe failures in any of the sets.

5.2. Direct Comparison of Candidates (Rule 3)

A number of pre-defined Figures of Merit were used to analyse and compare the performance of the candidates. The results are given in Tables 8a-8c. The best FOM for each case is highlighted in the tables with a boldface font.

Rule 3 FOM	Weighted ΔMOS					Weighted ΔdBQ					Unweighted %ΔPOW				
Codec #	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Total	19.0	6.8	**60.4**	19.6	32.0	146.9	47.6	**787.6**	217.7	353.4	36,5%	68,8%	**10,4%**	49,0%	19,8%

Table 8a: *FOM results for weighted □MOS, weighted □dBQ and unweighted %□POW.*

Rule 3 FOM	Number of systematic failures				
Codec #	1	2	3	4	5
Total	3	7	**0**	4	3

Table 8b: *FOM results for systematic failures.*

Rule 3 FOM restricted to failures	Weighted ΔMOS					Weighted ΔdBQ				
Codec #	1	2	3	4	5	1	2	3	4	5
Total	-2.1	-5.6	**0,0**	-1,4	-1.3	-30.4	-65.7	**0,0**	-13,9	-17.0

Table 8c: *FOM results for weighted □MOS and weighted □dBQ when restricted to failures.*

Codec 3 is the best quality codec in all the total FOMs.

6. Selection Phase Deliverables

All proponents provided a full set of selection deliverables in time. The deliverables required for the Selection Phase are explained in detail in [10].

6.1. Technical Descriptions

The proponents provided detailed technical descriptions of their algorithms (including speech and channel coding, source controlled rate scheme, and link adaptation). Each candidate was also required to provide a report showing that the proposal fulfills all the mandatory design constraints. To speed up the preparation of the AMR-WB specifications, all proponents were required to prepare first draft versions of the specifications to be revealed (for the best candidate) immediately after the selection in S4. All proponents provided all the necessary information on their proposals as required by the selection process.

6.2. Compliance to Design Constraints (Rule 1)

All mandatory design constraints are met by all candidates.

7. Conclusions on the AMR-WB Codec Candidates

7.1. Solution Characteristics

All candidate algorithms fulfil the mandatory design constraints.

7.2. Speech Quality Performance

All candidate algorithms meet the Rule 2 requirements for the amount of failures and severe failures. Codec 3 is the only codec candidate that meets all the performance requirements in all of the laboratories in the selection tests. It has no failures at all.
The Figures of Merit show that Codec 3 has the best quality of the candidates. Codec 3 is ranked as the best codec with regard to speech quality. Quality ranking for the remaining codecs was not performed.

7.3. Ranking of Codec Algorithms

Taking into account the listening test results, technical descriptions and other relevant information, Codec 3 is the best candidate and should be chosen as the AMR-WB codec.

8. Highlights of the Performance of the Best Codec Candidate

The quality performance of Codec 3 can be characterised as follows:
Applications A and B (GSM FR channel):

- For clean speech, the codec provides in Application A error-free quality exceeding G.722-48k and in Application B quality equal to G.722-56k.
- Under background noise, the codec provides in Application A error-free quality equal to G.722-48k and in Application B quality equal to G.722-56k.
- In both Applications A and B, at 13 dB C/I, quality is still equal to the quality of error-free G.722-48k, for both clean speech and in background noise. Below 13 dB C/I, smooth degradation (comparable to degradation for GSM EFR) is provided.

Applications C and E (GSM CS EDGE, 3G UTRAN):

- In the EDGE FR-channel, for clean speech and speech in background noise, at 22 dB C/I and above quality equal to error-free G.722-56k is provided. At 16 dB C/I, quality equal to error-free G.722-48k is still produced.
- In the EDGE HR-channel, for clean speech and speech in background noise, at 25 dB C/I and above quality equal to error-free G.722-56k is provided. At 19 dB C/I, quality equal to error-free G.722-48k is still produced.
- In the 3G UTRAN channel, for clean speech and speech in background noise, quality equal to G.722-64k is provided for error-free transmission. Under transmission errors at FER=1.0% / RBER=0.1%, quality equal to G722-48k is given. (The least significant bits are subjected to the residual error profile with the number of bits in this class 25% of the total bits per frame).

9. Recommendations

Taking into account all the factors including the analysis of the subjective test results, the technical descriptions of the AMR-WB candidate algorithms and their compliance against the mandatory design constraints, S4 recommends the following:

- Candidate Codec 3 (Nokia) should be chosen as the AMR-WB codec for 3GPP and GSM

TSG-SA are requested to approve this recommendation.

10. Conclusions

The AMR-WB codec selection phase testing and the analysis of results was completed on schedule by TSG-SA#10 bringing to a close a selection process of the AMR-WB codec among a set of (originally) 7 candidates. The Development and the two Selection Phases were completed in the space of more than one year (21 months from the approval of the WI description at TSG-SA#2). During the selection phase good quality was demonstrated by the codec candidates. The best candidate codec met all the performance requirements, outperformed the others in the Figures of Merit, and is recommended for the selection as the AMR-WB codec.

References:

[1] "Proposed TSG-S4 Work Items for approval", Tdoc SP-99060, 3GPP TSG-SA meeting #2, 2-4 March, 1999 (Fort Lauderdale, USA)

[2] "AMR-WB Feasibility Study Report v. 1.0.0", ETSI TC SMG Tdoc P-99-429, SMG meeting #29, 23rd – 25th June, 1999 (Miami, USA)

[3] "Common WI description for the Wideband Codec", Tdoc SP-99354, 3GPP TSG-SA meeting #5, 11-13 October, 1999 (Kjongju, South Korea)

[4] "AMR Wideband Speech Codec Qualification Phase Report", Tdoc SP-000259, 3GPP TSG-SA#8, 26-28 June, 2000 (Dusseldorf, Germany)

[5] "Permanent Project Document: AMR-WB Selection Test Plan (WB-8b, version 1.0) ", 3GPP TSG-S4 Tdoc S4-000382

[6] "Permanent Project Document: Processing Functions for WB-AMR Subjective Experiments (WB-7, v.1.0)", 3GPP TSG-S4 Tdoc S4-000389

[7] "Permanent Project Document: AMR Wideband Performance Requirements (WB-3, version 2.2)", 3GPP TSG-S4 Tdoc S4-000321

[8] "Permanent Project Document: Selection Rules for AMR-WB (WB-5b, version 1.1)", 3GPP TSG-S4 Tdoc S4-000508

[9] "Permanent Project Document: Design Constraints (WB-4, version 1.3)", 3GPP TSG-S4 Tdoc S4-000340

[10] "Permanent Project Document: AMR Wideband Codec Development Project Deliverables for the Selection Test (WB-6b, version 2.0)", 3GPP TSG-S4 Tdoc S4-000427

[11] "3G TS 26.171 AMR Wideband Speech Codec; General description v.0.0.1", 3GPP TSG-SA Tdoc SP-00xxxx

[12] "3G TS 26.173 AMR Wideband Speech Codec; C-source code v.0.0.1", 3GPP TSG-SA Tdoc SP-00xxxx

[13] "3G TS 26.190 AMR Wideband Speech Codec; Transcoding Functions v.0.0.3", 3GPP TSG-SA Tdoc SP-00xxxx

[14] "3G TS 26.191 AMR Wideband Speech Codec; Error concealment of erroneous or lost frames v.0.0.1", 3GPP TSG-SA Tdoc SP-00xxxx

[15] “3G TS 26.192 AMR Wideband Speech Codec; CN for AMR Speech Traffic Channels v.0.0.2”, 3GPP TSG-SA Tdoc SP-00xxxx
[16] “3G TS 26.193 AMR Wideband Speech Codec; Source Controlled Rate operation v.0.0.1”, 3GPP TSG-SA Tdoc SP-00xxxx
[17] “3G TS 26.194 AMR Wideband Speech Codec; VAD for AMR Speech Traffic Channels v.0.0.2”, 3GPP TSG-SA Tdoc SP-00xxxx
[18] “3G TS 26.201 AMR Wideband Speech Codec; Speech Codec Frame Structure v.0.0.1”, 3GPP TSG-SA Tdoc SP-00xxxx
[19] “3G TS 26.202 AMR-WB speech codec; interface to Iu and Uu v.0.0.1”, 3GPP TSG-SA Tdoc SP-00xxxx

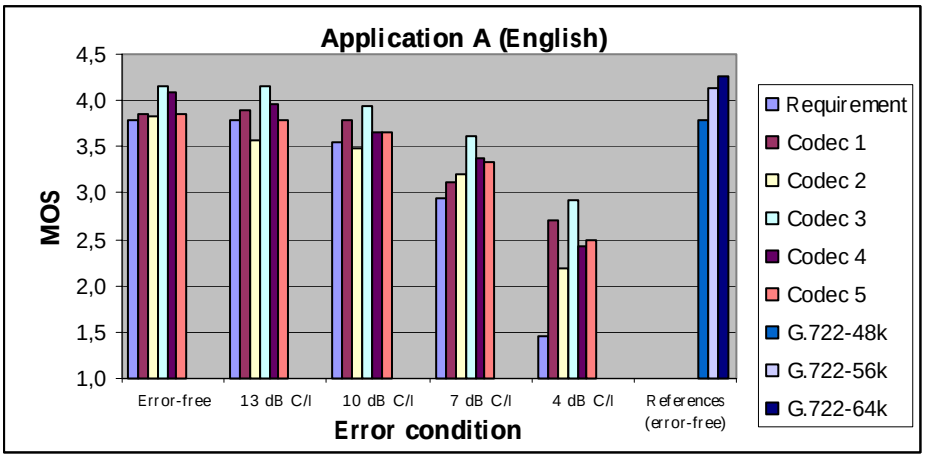
List of Annexes:

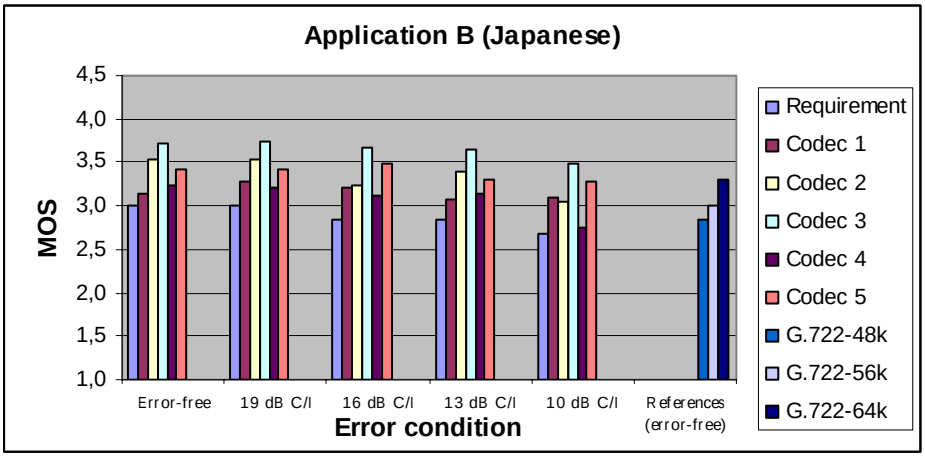
Annex 1: Extracts from the AMR-WB Selection Test Results

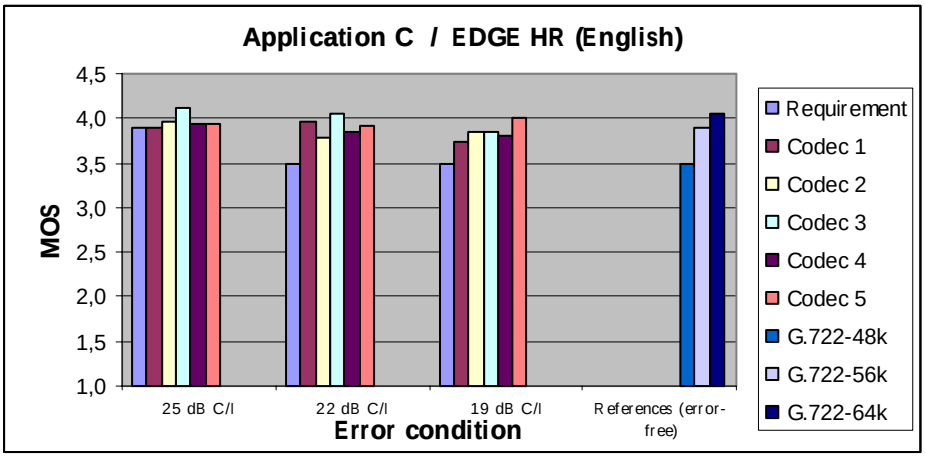
Annex 2: Complete AMR-WB Selection Test Results

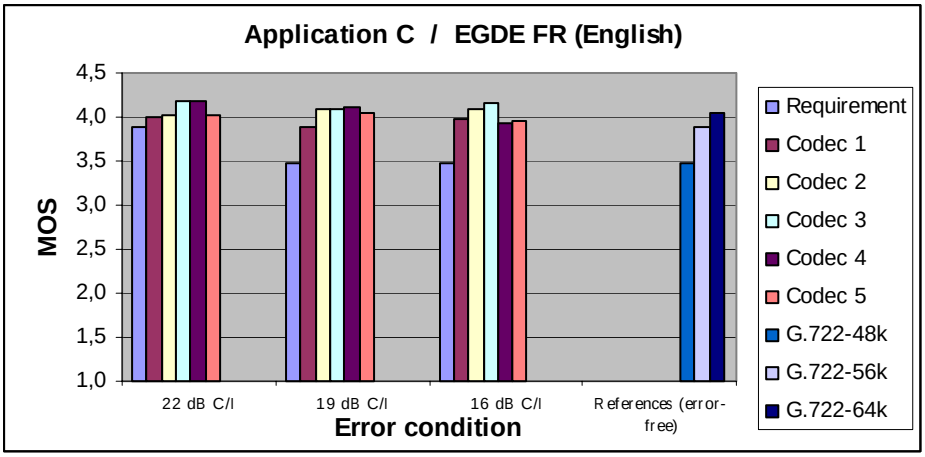
Annex 3: Complexity of the AMR-WB Candidate Codecs

Annex 1: Extracts from the AMR-WB Selection Test Results









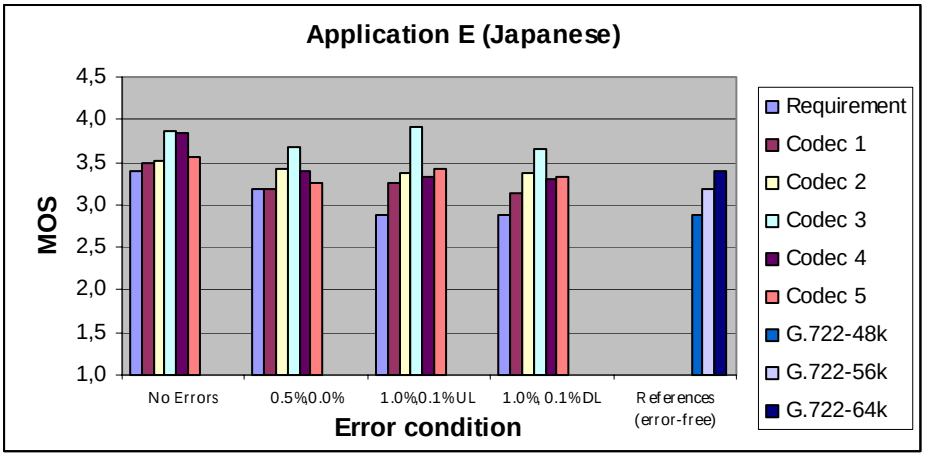
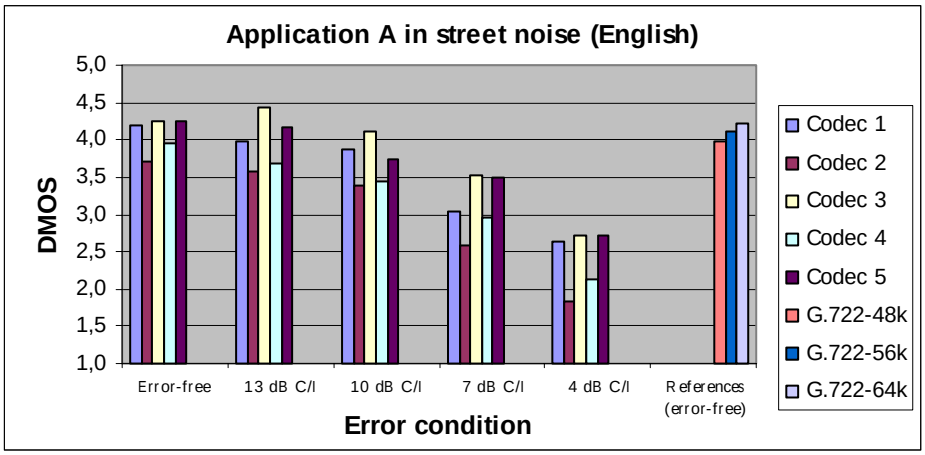
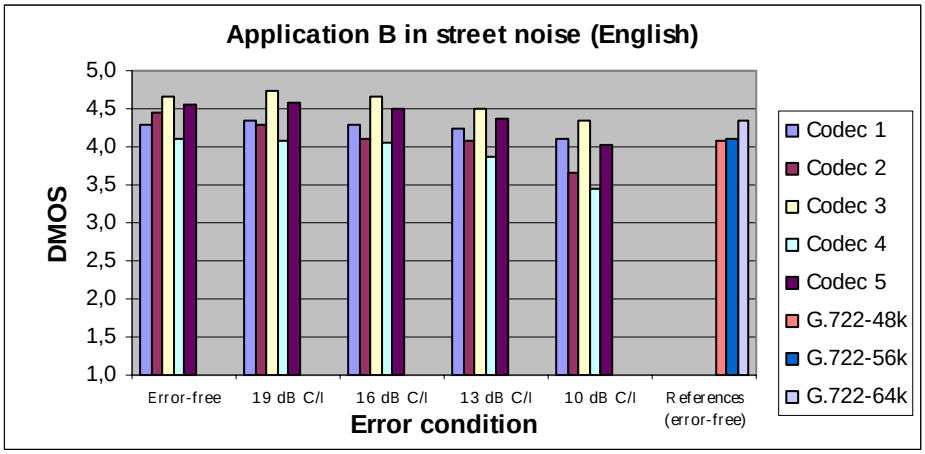


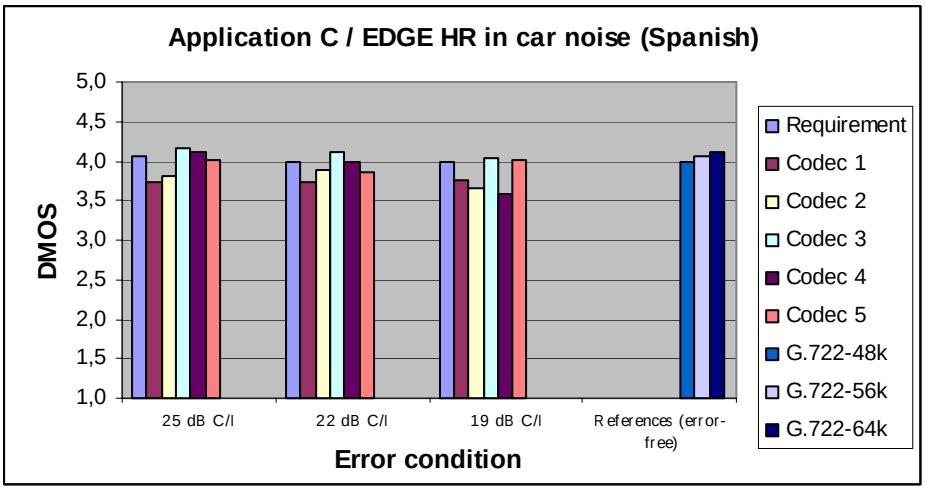
Fig 1: Experiment 2: Clean Speech performance with static errors (ACR)

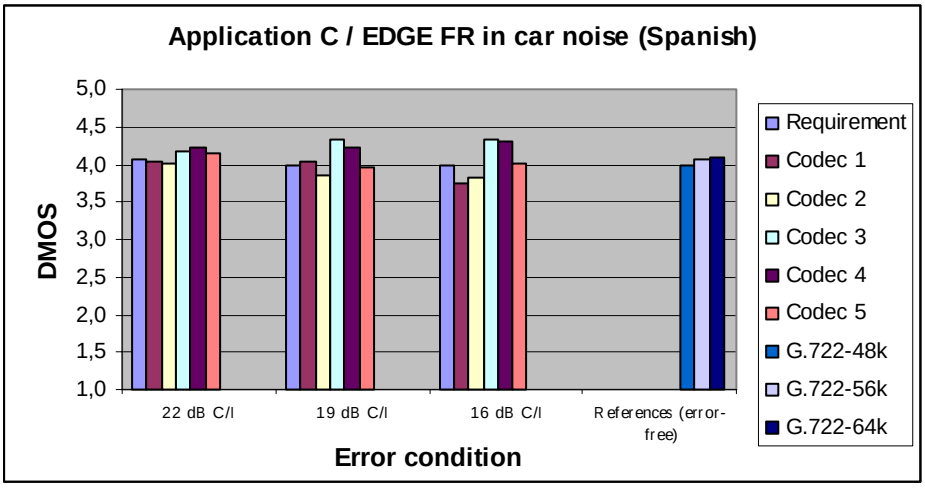
a) Application A (English)
b) Application B (Japanese)
c) Application C / EDGE HR (English)
d) Application C / EDGE FR (English)
e) Application E (Japanese)

Note: The absolute MOS values depend on the test setting and conditions and are not directly comparable between the sub-experiments.


Application A in street noise (English)
DMOS
5,0
4,5
4,0
3,5
3,0
2,5
2,0
1,5
1,0
Error-free
13 dB C/I
10 dB C/I
7 dB C/I
4 dB C/I
References (error-free)
Error condition
Codec 1
Codec 2
Codec 3
Codec 4
Codec 5
G.722-48k
G.722-56k
G.722-64k


Application B in street noise (English)
DMOS
5,0
4,5
4,0
3,5
3,0
2,5
2,0
1,5
1,0
Error-free
19 dB C/I
16 dB C/I
13 dB C/I
10 dB C/I
References (error-free)
Error condition
Codec 1
Codec 2
Codec 3
Codec 4
Codec 5
G.722-48k
G.722-56k
G.722-64k


Application C / EDGE HR in car noise (Spanish)
DMOS
5,0
4,5
4,0
3,5
3,0
2,5
2,0
1,5
1,0
25 dB C/I
22 dB C/I
19 dB C/I
References (error-free)
Error condition
Requirement
Codec 1
Codec 2
Codec 3
Codec 4
Codec 5
G.722-48k
G.722-56k
G.722-64k


Application C / EDGE FR in car noise (Spanish)
DMOS
5,0
4,5
4,0
3,5
3,0
2,5
2,0
1,5
1,0
22 dB C/I
19 dB C/I
16 dB C/I
References (error-free)
Error condition
Requirement
Codec 1
Codec 2
Codec 3
Codec 4
Codec 5
G.722-48k
G.722-56k
G.722-64k

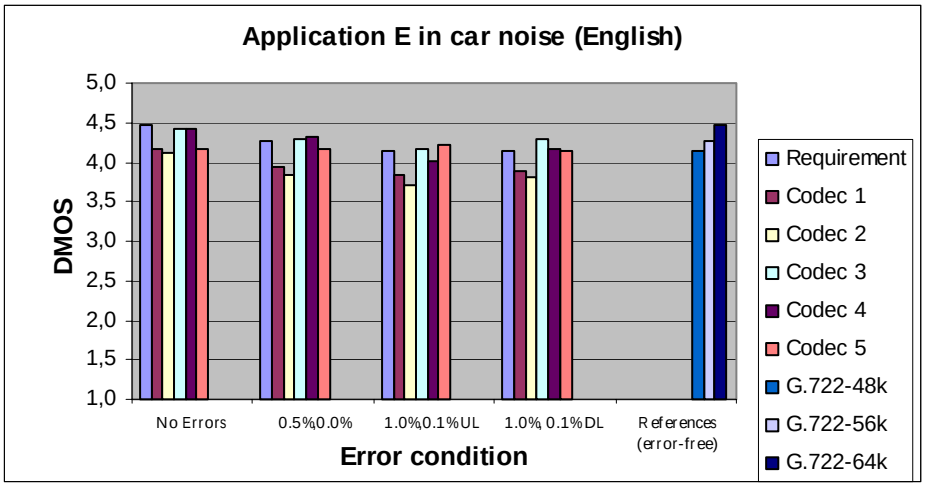


Fig 2: Experiment 3: Car and Street noise (15 dB SNR) performance for the GSM FR channel (DCR-test); and Experiment 4: Car and Street noise (15 dB SNR) performance for higher-rate channels (DCR-test)

a) Application A in street noise (English)
b) Application B in street noise (English)
c) Application C / EDGE HR in car noise (Spanish)
d) Application C / EDGE FR in car noise (Spanish)
e) Application E in car noise (English)

Note: The absolute DMOS values depend on the test setting and conditions and are not directly comparable between the sub-experiments. (Note also that the requirements are not drawn in figures 2a and 2b since they are not given as DMOS-values, but instead as 10% PoW measures.)

Annex 2: Complete AMR-WB Selection Test Results

See the Excel-spreadsheet in the attached file "AMRWB_GAL.zip" (contained in S4 document S4-000485)

Annex 3: Complexity of the AMR-WB Candidate Codecs

This Annex gives estimates of the codec complexities (given by codec proponents)[4].

COMPLEXITY	Requirement	Codec 1	Codec 2	Codec 3	Codec 5
Speech codec complexity A: wMOPS B: RAM C: ROM D: Program ROM	A: wMOPS ≤ 40 wMOPS B: RAM ≤ 15 kwords C: ROM ≤ 18 kwords D: Prog. ROM ≤ 5821 basic operators	A: 38.63 wMOPS B: 13.415 kwords C: 16.279 kwords D: 4798 basic ops	A: 37.091 wMOPS B: 12.066 kwords C: 7.332 kwords D: 5481 basic ops	A: 35.4 wMOPS B: 6.42 kwords C: 9.94 kwords D: 3771 basic ops	A: 38.9 wMOPS B: 5.94 kwords C: 16.02 kwords D: 5512 basic ops
Additional complexity for source controlled rate operation (over speech coding complexity limits) E: wMOPS F: RAM G: ROM H: Program ROM	E: wMOPS ≤ 1.6 wMOPS F: RAM ≤ 149 words G: ROM ≤ 513 words H: Program ROM ≤ 491 basic operators	E: 0.833 wMOPS F: B includes this G: C includes this H: D includes this	E: 0.479 wMOPS F: 107 words G: 7 words H: 131 basic ops	E: 0.73 wMOPS F: 75 words G: 0 words H: 268 basic ops	E: 0.36 wMOPS F: 65 words G: 0 words H: 314 basic ops
Channel codec complexity for Applications A and B: I: wMOPS J: RAM K: ROM L: Program ROM	I: wMOPS ≤ 5.7 wMOPS J: RAM ≤ 3.0 kwords K: ROM ≤ 4.5 kwords L: Program ROM ≤ 2013 basic operators	I: 4.51 wMOPS J: 2722 kwords K: 4075 kwords L: 1346 basic ops	I: 5.42 wMOPS J: 2.359 kwords K: 4.242 kwords L: 360 basic ops	I: 3.45 wMOPS J: 2.88 kwords K: 3.18 kwords L: 579 basic ops	I: 5.5 wMOPS J: 2.787 kwords K: 2.985 kwords L: 910 basic ops
Constraints for channel codec in Application C (example solution used in testing)	• Only the polynomials denoted G1-G7 in 05.03 can be applied. • Recursive Systematic Codes as used in TCH/AFS and TCH/AHS can be used. • Constraint length K=7 can be used in all modes. • Use of a single CRC is allowed up to 16 parity bits. • 24 bits should be reserved to an inband channel in FR and 12 bits in HR.	Requirement is met.	Requirement is met.	Requirement is met.	Requirement is met.

[4] Codec 4 was withdrawn during the Selection Phase and no estimates for complexity were given for it.