Technical Specification Group Services and System Aspects **TSGS#23(04)0356**
Meeting #24, Seoul, South Korea
7-10 June 2004

**Source:** **TSG-SA WG4**

**Title: CRs TS 26.235 and TS 26.236 on the introduction of the DSR codec (Release 6)**

**Document for:** **Approval**

**Agenda Item:** **7.4.3**

The following CRs, agreed at the TSG-SA WG4 meeting #31, are presented to TSG SA #24 for approval.

| Spec | CR | Rev | Phase | Subject | Cat | Vers | WG | Meeting | S4 doc |
|---|---|---|---|---|---|---|---|---|---|
| 26.235 | 006 | 4 | Rel-6 | Introduction of the DSR codec | B | 6.0.0 | S4 | TSG-SA WG4#31 | S4-040360 |
| 26.236 | 010 | 3 | Rel-6 | Introduction of the DSR codec | B | 5.4.0 | S4 | TSG-SA WG4#31 | S4-040359 |

**3GPP TSG-S4#31 meeting**
**May 17 - 21, 2004, Montreal, Canada**

***Tdoc S4 (04)0359***

*CR-Form-v3*

# CHANGE REQUEST

⌘ **26.236** **CR** **010** ⌘ rev **3** ⌘ Current version: **5.4.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

***Proposed change affects:*** ⌘ UICC apps ☐ ME ☒ Radio Access Network ☐ Core Network ☐

| | | |
|---|---|---|
| ***Title:*** | ⌘ | Introduction of the DSR Codec |
| ***Source:*** | ⌘ | TSG SA WG4 |
| ***Work item code:*** | ⌘ | SRSES codec |
| ***Date:*** | ⌘ | 08.06.2004 |
| ***Category:*** | ⌘ | **B** |
| ***Release:*** | ⌘ | REL-6 |

*Use one of the following categories:*
***F*** *(essential correction)*
***A*** *(corresponds to a correction in an earlier release)*
***B*** *(Addition of feature),*
***C*** *(Functional modification of feature)*
***D*** *(Editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

*Use one of the following releases:*

| | |
|---|---|
| *2* | *(GSM Phase 2)* |
| *R96* | *(Release 1996)* |
| *R97* | *(Release 1997)* |
| *R98* | *(Release 1998)* |
| *R99* | *(Release 1999)* |
| *REL-4* | *(Release 4)* |
| *REL-5* | *(Release 5)* |

| | | |
|---|---|---|
| ***Reason for change:*** | ⌘ | Additions for the DSR codec |
| ***Summary of change:*** | ⌘ | Changes needed to introduce SES payload formats |
| ***Consequences if not approved:*** | ⌘ | The DSR feature is not introduced |

| | | |
|---|---|---|
| ***Clauses affected:*** | ⌘ | Introduction, 2, 3.2, 4, 5.4 |
| ***Other specs Affected:*** | ⌘ | ☐ Other core specifications ⌘<br>☐ Test specifications<br>☐ O&M Specifications |
| ***Other comments:*** | ⌘ | |

# Introduction

The present document contains a specification for required protocol usage within 3GPP specified Conversational Packet Switched Multimedia Services [5] which is based IP Multimedia Subsystem (IM Subsystem). IM Subsystem as a subsystem includes specifically the conversational IP multimedia services, whose service architecture, call control and media capability control procedures have been defined in 3GPP TS 24.229 [7], and are based on the 3GPP adopted version of IETF Session Initiated Protocol (SIP) [1].

In conversational packet switched multimedia service depends on IM Subsystem. The individual media types are independently encoded and packetized to appropriate separate Real Time Protocol (RTP) packets. These packets are then transported end-to-end inside UDP datagrams over real-time IP connections that have been negotiated and opened between the terminals or between a terminal and a server during the SIP call as specified in 3GPP TS 24.229 [7].

The UEs operating within IM Subsystem need to provide encoding/decoding of the derived codecs, and perform corresponding packetization/depacketization functions. Logical bound between the media streams is handled in the SIP session layer, and inter-media synchronization in the receiver is handled with the use of RTP time stamps.

Next changed section

## 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] IETF RFC 2543: "SIP: Session Initiation Protocol".

[2] IETF RFC 2327: "SDP: Session Description Protocol".

[3] IETF RFC 3550: "RTP: A Transport Protocol for Real-Time Applications", Schulzrinne H. et al, July 2003.

[4] IETF RFC 3551: "RTP Profile for Audio and Video Conferences with Minimal Control", Schulzrinne H. and Casner S., July 2003.

[5] 3GPP TS 26.235: "Packet switched conversational multimedia applications; Default codecs".

[6] 3GPP TS 24.228: "Signalling flows for the IP multimedia call control based on SIP and SDP; stage 3".

[7] 3GPP TS 24.229: "IP multimedia call control protocol based on SIP and SDP".

[8] 3GPP TS 23.228: "IP Multimedia Ssubsystem (IMS); Stage 2".

[9] 3GPP TS 23.107: "Quality of Service (QoS) concept and architecture".

[10] 3GPP TS 23.207: "End to end quality of service concept and architecture".

[11] 3GPP TS 23.060: "General Packet Radio Service (GPRS); Service description; Stage 2".

[12] 3GPP TS 26.071: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; General description".

[13] 3GPP TS 26.090: "AMR speech Codec; Transcoding Functions".

[14] 3GPP TS 26.073: "AMR speech Codec; C-source code".

[15] 3GPP TS 26.104: "ANSI-C code for the floating-point Adaptive Multi-Rate AMR speech codec".

[16] 3GPP TS 26.171 (Release 5): "AMR speech codec, wideband; General description".

[17] 3GPP TS 26.190 (Release 5): "Mandatory Speech Codec speech processing functions AMR Wideband speech codec; Transcoding functions".

[18] 3GPP TS 26.201 (Release 5): "AMR speech codec, wideband; Frame structure".

[19] IETF RFC 3267: "RTP payload format and file storage format for the Adaptive Multi-Rate (AMR) Adaptive Multi-Rate Wideband (AMR-WB) audio codecs", March 2002.

~~3GPP TS 26.235: "Packet switched conversational multimedia applications; Default codecs ". Annex B: "RTP payload format and storage format for AMR and AMR-WB audio".~~

[20] ITU-T Recommendation H.263: "Video coding for low bit rate communication".

[21] IETF RFC 2429: "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)".

[22] ISO/IEC 14496-2 (1999): "Information technology - Coding of audio-visual objects - Part 2: Visual".

[23] IETF RFC 3016: "RTP Payload Format for MPEG-4 Audio/Visual Streams".

[24] ITU-T Recommendation H.263 (annex X): "Annex X: Profiles and levels definition".

[25] 3GPP TS 26.235: "Packet Switched Conversational Multimedia Applications; Default Codecs ". Annex C: "ITU-T H.263 MIME media type registration".

[26] ITU-T Recommendation T.140 (1998): "Protocol for multimedia application text conversation" (with amendment 2000).

[27] IETF RFC 2793: "RTP Payload for Text Conversation".

[28] IETF RFC 3556: "Session Description Protocol (SDP) Bandwidth Modifiers for RTP Control Protocol (RTCP) bandwidth", Casner S., July 2003.

[29] RTP Payload Formats for European Telecommunications Standards Institute (ETSI) European Standard ES 202 050, ES 202 211, and ES 202 212 Distributed Speech Recognition Encoding draft-ietf-avt-rtp-dsr-codecs-00.txt

CR Editor's note: The above document cannot be formally referenced until it is published as an RFC.

# 3 Definitions and abbreviations

## 3.1 Definitions

For the purposes of the present document, the following term and definition applies:

**3G PS multimedia terminal:** terminal based on IETF SIP/SDP internet standards modified by 3GPP for purposes of 3GPP IM Subsystem services

## 3.2 Abbreviations

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

| | |
|---|---|
| AMR | Adaptive MultiRate codec |
| DSR | Distributed Speech Recognition |
| IETF | Internet Engineering Task Force |
| IM Subsystem | Internet protocol Multimedia Subsystem |
| ITU-T | International Telecommunications Union-Telecommunications |
| RFC | IETF Request For Comments |
| RTPCP | RTP Control Protocol |
| RTP | Real-time Transport Protocol |
| SDP | Session Description Protocol |
| SES | Speech Enabled Service |
| SIP | Session Initiation Protocol |

# 4 General

3G PS multimedia terminals provide real-time video, audio, SES or data, in any combination, including none, over 3GPP IM Subsystem. Terminals are based on IETF defined multimedia protocols SIP, SDP, RTP and RTCP. Communication may be either 1-way or 2-way. Such terminals may be part of a portable device or integrated into an automobile or other non-fixed location device. They may also be fixed, stand-alone devices; for example, a video telephone or kiosk. Multimedia terminals may also be integrated into PCs and workstations.

In the case of SES then uplink communication is from the terminal to a server containing speech recognition.

In addition, interoperation with other types of multimedia telephone terminals, such as 3G-324M may be possible, however in such case a media gateway functionality supporting 3G-324M - IM Subsystem interworking will be required within or outside the IM subsystem.

Figure 1 presents the user plane protocol stack of a 3G PS conversational multimedia terminal explaining the transport of different media types and QoS reports.

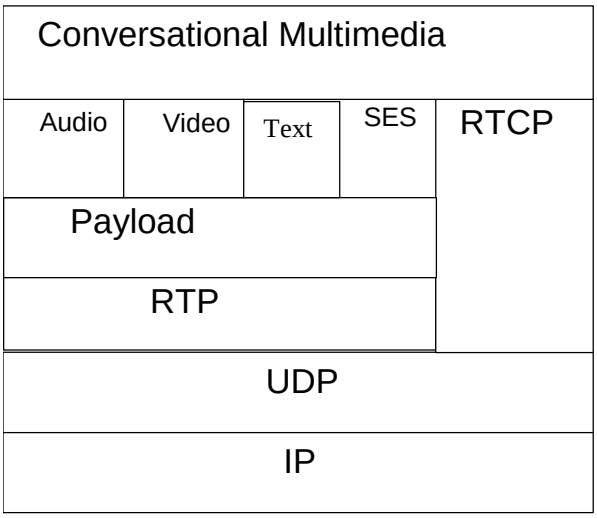


**Figure 1 – User plane protocol stack for 3G PS conversational multimedia terminal**

# 5 Media type requirements

Media type RTP payload usage is specified in this clause. The media types and corresponding codecs are specified in 3GPP TS 26.235 [5]. The continuous media type RTP payloads are mapped to RTP packets according to IETF RTP Profile for Audio and Video Conferences with Minimal Control in RFC 3551 [4].

## 5.1 Audio

### 5.1.1 RTP session description parameters

The IETF AMR and AMR-WB RTP payload format [19] offers different options. Here is the list of options and how they should be used by the transmitter. The receiver shall at least support the options as they are listed:

- the bandwidth efficient operation shall be used,
- only one speech frame shall be encapsulated in each RTP packet,
- the multi-channel session shall not be used,
- interleaving shall not be used,
- internal CRC shall not be used.

## 5.2 Video

Video packets should not be large to allow better error resilience and to minimize the transmission delay in conversational service. The size of each packet shall be kept smaller than 512 bytes.

## 5.3 Real time text

Real time text media type RTP payload format for ITU-T Recommendation T.140 is specified in [27]. Redundant transmission provided by the RTP payload format is recommended in error prone channel.

## 5.4 SES

The RTP payload for the DSR codec and AMR or AMR-WB used for SES are specified in [29, 19].

**3GPP TSG-S4#31 meeting**
**May 17-21, 2004, Montreal, Canada**

***Tdoc S4 (04)0360***

*CR-Form-v3*

# CHANGE REQUEST

⌘ **26.235** **CR** **006** ⌘ rev **4** ⌘ Current version: **6.0.0** ⌘

*For **HELP** on using this form, see bottom of this page or look at the pop-up text over the ⌘ symbols.*

***Proposed change affects:*** ⌘ UICC apps ☐ ME **X** Radio Access Network ☐ Core Network ☐

| | | |
|---|---|---|
| ***Title:*** | ⌘ | Introduction of the DSR Codec |
| ***Source:*** | ⌘ | TSG SA WG4 |
| ***Work item code:*** | ⌘ | SRSES-Codec |
| ***Date:*** | ⌘ | 08.06.2004 |
| ***Category:*** | ⌘ | **B** |
| ***Release:*** | ⌘ | REL-6 |

*Use one of the following categories:*
***F*** *(essential correction)*
***A*** *(corresponds to a correction in an earlier release)*
***B*** *(Addition of feature),*
***C*** *(Functional modification of feature)*
***D*** *(Editorial modification)*
Detailed explanations of the above categories can be found in 3GPP TR 21.900.

*Use one of the following releases:*

| | |
|---|---|
| *2* | *(GSM Phase 2)* |
| *R96* | *(Release 1996)* |
| *R97* | *(Release 1997)* |
| *R98* | *(Release 1998)* |
| *R99* | *(Release 1999)* |
| *REL-4* | *(Release 4)* |
| *REL-5* | *(Release 5)* |

| | | |
|---|---|---|
| ***Reason for change:*** | ⌘ | Additions for the DSR codec |
| ***Summary of change:*** | ⌘ | Changes needed to add the DSR codec into the codec list |
| ***Consequences if not approved:*** | ⌘ | The DSR feature is not introduced |

| | | |
|---|---|---|
| ***Clauses affected:*** | ⌘ | 2, 3.2, 6.5, 9.1and 9.2, Annex D |
| ***Other specs Affected:*** | ⌘ | ☐ Other core specifications ⌘<br>☐ Test specifications<br>☐ O&M Specifications |
| ***Other comments:*** | ⌘ | |

# 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] IETF RFC 3261: "SIP: Session Initiation Protocol".

[2] IETF RFC 2327: "SDP: Session Description Protocol".

[3] IETF RFC 2429: "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)".

[4] IETF RFC 1889: "RTP: A Transport Protocol for Real-Time Applications".

[5] IETF RFC 3016: "RTP Payload Format for MPEG-4 Audio/Visual Streams".

[6] ITU-T Recommendation H.263: "Video coding for low bit rate communication".

[7] 3GPP TS 26.110: "Codec for Circuit Switched Multimedia Telephony Service; General Description".

[8] 3GPP TS 26.111: "Codec for Circuit Switched Multimedia Telephony Service; Modifications to H.324".

[9] 3GPP TS 26.071: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; General description".

[10] 3GPP TS 26.090: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Transcoding functions".

[11] 3GPP TS 26.073: "Adaptive Multi-Rate (AMR); ANSI C source code".

[12] 3GPP TS 26.104: "ANSI-C code for the floating-point AMR speech codec".

[13] ISO/IEC 14496-2 (1999): "Information technology - Coding of audio-visual objects - Part 2: Visual".

[14] 3GPP TS 24.228: "Signalling flows for the IP multimedia call control based on SIP and SDP".

[15] 3GPP TS 24.229: "IP Multimedia Call Control Protocol based on SIP and SDP".

[16] 3GPP TS 26.171 (Release 5): "AMR speech codec, wideband; General description".

[17] 3GPP TS 26.190 (Release 5): "Mandatory Speech Codec speech processing functions AMR Wideband speech codec; Transcoding functions".

[18] 3GPP TS 26.201 (Release 5): "AMR speech codec, wideband; Frame structure".

[19] ITU-T Recommendation H.263 (annex X): "Annex X, Profiles and levels definition".

[20] 3GPP TS 23.228: "IP multimedia subsystem; stage 2".

[21] 3GPP TS 23.107: "QoS Concept and Architecture".

[22] 3GPP TS 23.207: "End to end quality of service concept and architecture".

[23] 3GPP TS 23.060: "General Packet Radio Service (GPRS); Service description; Stage 2".

[24] IETF RFC 2793: "RTP Payload for Text Conversation".

[25] ITU-T Recommendation T.140 (1998): "Protocol for multimedia application text conversation" (with amendment 2000).

[26] 3GPP TS 26.101: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Frame Structure".

[27] IETF RFC 2119: "Key words for use in RFCs to Indicate Requirement Levels".

[28] 3GPP TS 26.093: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Source Controlled Rate operation".

[29] 3GPP TS 46.060: "Enhanced Full Rate (EFR) speech transcoding".

[30] TIA/EIA -136-Rev.A, part 410 - "TDMA Cellular/PCS – Radio Interface, Enhanced Full Rate Voice Codec (ACELP). Formerly IS-641. TIA published standard, 1998".

[31] ARIB, RCR STD-27H, "Personal Digital Cellular Telecommunication System RCR Standard".

[32] IETF draft-westberg-realtime-cellular-01.txt, "Realtime Traffic over Cellular Access Networks".

[33] IETF draft-larzon-udplite-03.txt, "The UDP Lite Protocol".

[34] 3GPP TS 26.092: "Mandatory Speech Codec speech processing functions; AMR Speech Codec; Comfort noise aspects".

[35] IETF RFC 3267: "RTP payload format and file storage format for the Adaptive Multi-Rate (AMR) Adaptive Multi-Rate Wideband (AMR-WB) audio codecs", March 2002.

[36] IETF RFC 2833: “RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals”, May 2000.

[37] 3GPP TS 26.243: “TS Software documentation for fixed-point DSR Extended Front End ”

[38] RTP Payload Formats for European Telecommunications Standards Institute (ETSI) European Standard ES 202 050, ES 202 211, and ES 202 212 Distributed Speech Recognition Encoding draft-ietf-avt-rtp-dsr-codecs-00.txt

CR Editor's note: The above document cannot be formally referenced until it is published as an RFC.

# 3 Definitions and abbreviations

## 3.1 Definitions

For the purposes of the present document, the following terms and definitions apply.:

**3G PS multimedia terminal:** terminal based on IETF SIP/SDP internet standards modified by 3GPP for purposes of 3GPP packet switched network based multimedia telephony

## 3.2 Abbreviations

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

| | |
|---|---|
| AMR | Adaptive MultiRate codec |
| DSR | Distributed Speech Recognition |
| IETF | Internet Engineering Task Force |
| IM Subsystem | Internet protocol Multimedia Subsystem |
| ITU-T | International Telecommunications Union-Telecommunications |

| | |
|---|---|
| RFC | IETF Request For Comments |
| RTCP | RTP Control Protocol |
| RTP | Real-time Transport Protocol |
| SDP | Session Description Protocol |
| SES | Speech Enabled Services |
| SIP | Session Initiated Protocol |

**Next changed section**

## 6.4 Interactive and background data

SIP signalling offers initialisation of packet switched interactive or background class reliable data services as well. However specification of such data services are outside the scope of the present document.

## 6.5 Speech Enabled Service

3G PS multimedia terminals offering speech enabled services should support the DSR Extended Advanced Front-end codec [37]

Speech enabled services may also be supported with AMR or AMR-WB audio codecs, however it is noted that there is a substantial performance advantage from DSR [ see Annex D ].

**Next changed section**

# 9 Multimedia stream encapsulation

## 9.1 MIME media types

The terminal shall declare the mandatory and any optional media streams using the codec specific MIME media types in the associated SDP syntax. The MIME media types for the mandatory and optional codecs shall be according to the corresponding types registered by IANA.

- AMR narrowband speech codec MIME media type as specified in annex B.
- AMR wideband speech codec MIME media type is specified in annex B.
- H.263 [6] video codec MIME media type is specified in annex C.
- MPEG-4 visual simple profile level 0 MIME media type as specified in RFC 3016 [5].
- ITU-T Recommendation T.140 [25] Text Conversation MIME media type as specified by RFC 2793 [24].
- Telephone-event MIME media type as specified by RFC 2833 [36].
- DSR MIME media type as specified in draft-ietf-avt-rtp-dsr-codecs-00.txt [38]

## 9.2 RTP payload

RTP payload formats specified by IETF shall be used for real time media streams.

RTP payload format for the AMR narrowband speech codec is specified in annex B.

RTP payload format for the AMR wideband speech codec is specified in annex B.

RTP payload format for the ITU-T Recommendation H.263 [6] video codec is specified in IETF RFC 2429 [3].

RTP payload format for the MPEG-4 visual simple profile level 0 is specified in IETF RFC 3016 [5].

RTP payload format for the ITU-T Recommendation T.140 [25] text conversation coding is specified in IETF RFC 2793 [24].

RTP payload format for the telephone-event is specified in IETF RFC 2833 [36].

RTP payload format for the DSR Extended Advanced Front-end is specified in draft-ietf-avt-rtp-dsr-codecs-00.txt [38]

# Annex D (Informative): Performance results from SES selection

The following tables show the results from the SES selection evaluations conducted in SA4 by ASR vendors Scansoft and IBM.

There are three tables applicable for comparison at different channels bandwidths and sampling rates:

- Low data rate comparison (which is at 8kHz) [AMR 4.75 compared to DSR]
- High data rate comparison at 8kHz [AMR 12.2 compared to DSR]
- High data rate comparison at 16kHz [AMR-WB 12.65 compared to DSR]

For each test database the absolute performance in terms of word error rate is given and the relative improvement of DSR compared to AMR is shown.

The relative improvement is computed for each database as (word error rate for AMR – word error rate for DSR)/word error rate for AMR.

The tasks are split into the task categories shown in the tables: ie digits, subword, tone confusability and channel errors.

The overall word error rate improvement is the weighted average of the improvement for each of the categories. Ie Sum of task category weight x improvement for category.

# Low Data Rate comparison

Sampling rate = 8kHz
AMR mode = AMR-NB 4.75

| | | word error rate AMR-NB 4.75 | word error rate DSR | Relative Improvement |
|---|---|---|---|---|
| Digits | Aurora-2 (result B) | 11.73 | 9.62 | 17.99% |
| | Aurora-2 (result A) | 16.1 | 12.4 | 22.98% |
| | Aurora-3 German | 18.27 | 13.83 | 24.30% |
| | Aurora-3 Spanish (Result A) | 9.23 | 4.86 | 47.35% |
| | Aurora-3 Spanish (Result B) | 13.93 | 4.86 | 65.11% |
| | Aurora-3 Italian | 21.68 | 6.15 | 71.63% |
| | US English In-Car (digits test) | 19 | 12 | 36.84% |
| | German In-Car (digit test) | 11.4 | 8.3 | 27.19% |
| | Japanese In-Car (digit test) | 16.2 | 9 | 44.44% |
| | US English In-Car (digits test) | 4.49 | 2.44 | 45.66% |
| | Mandarin Embedded PDA (digit test) | 2.57 | 1.66 | 35.41% |
| 0.3 | **Average improvement on digits tasks** | | | **39.90%** |
| Subword | Mandarin Embedded PDA | 4.09 | 2.52 | 38.39% |
| | US English In-Car | 4.25 | 2.78 | 34.59% |
| | US English In-Car | 14.2 | 9.5 | 33.10% |
| | German In-Car | 12 | 10.1 | 15.83% |
| | Japanese In-Car | 18 | 13 | 27.78% |
| | Mandarin Name dialling (baseform test) | 0.83 | 0.58 | 30.12% |
| 0.4 | **Average improvement on subword tasks** | | | **29.97%** |
| Tone Confusability | Mandarin Name dialling (tone confusion test) | 3.59 | 3.06 | 14.76% |
| 0.1 | **Average improvement on tone confusability** | | | **14.76%** |
| Channel errors | 1% BLER (result A) | 5.67 | 2.39 | 57.85% |
| | 1% BLER (result B) | 9.4 | 6.7 | 28.72% |
| | 3% BLER (result A) | 6.51 | 2.38 | 63.44% |
| | 3% BLER (result B) | 17.6 | 6.8 | 61.36% |
| 0.2 | **Average improvement with channel errors** | | | **52.84%** |

## OVERALL RELATIVE REDUCTION IN WORD ERROR RATE 36%

# High Data Rate comparison at 8kHz

Sampling rate = 8kHz
AMR mode = AMR-NB 12.2

| | | word error rate AMR-NB 12.2 | word error rate DSR | Relative Improvement |
|---|---|---|---|---|
| Digits | Aurora-2 (result B) | 10.28 | 9.62 | 6.42% |
| | Aurora-2 (result A) | 14.2 | 12.4 | 12.68% |
| | Aurora-3 German | 15.9 | 13.83 | 13.02% |
| | Aurora-3 Spanish (Result A) | 7.7 | 4.86 | 36.88% |
| | Aurora-3 Spanish (Result B) | 11.95 | 4.86 | 59.33% |
| | Aurora-3 Italian | 19.04 | 6.15 | 67.70% |
| | US English In-Car (digits test) | 15.6 | 12 | 23.08% |
| | German In-Car (digit test) | 8.6 | 8.3 | 3.49% |
| | Japanese In-Car (digit test) | 11 | 9 | 18.18% |
| | US English In-Car (digits test) | 3.37 | 2.44 | 27.60% |
| | Mandarin Embedded PDA (digit test) | 2.57 | 1.66 | 35.41% |
| 0.3 | **Average improvement on digits tasks** | | | **27.62%** |
| Subword | Mandarin Embedded PDA | 3.14 | 2.52 | 19.75% |
| | US English In-Car | 3.29 | 2.78 | 15.50% |
| | US English In-Car | 12.9 | 9.5 | 26.36% |
| | German In-Car | 9.7 | 10.1 | -4.12% |
| | Japanese In-Car | 12.8 | 13 | -1.56% |
| | Mandarin Name dialling (baseform test) | 0.84 | 0.58 | 30.95% |
| 0.4 | **Average improvement on subword tasks** | | | **14.48%** |
| Tone Confusability | Mandarin Name dialling (tone confusion test) | 3.81 | 3.06 | 19.69% |
| 0.1 | **Average improvement on tone confusability** | | | **19.69%** |
| Channel errors | 1% BLER (result A) | 4.73 | 2.39 | 49.47% |
| | 1% BLER (result B) | 7.1 | 6.7 | 5.63% |
| | 3% BLER (result A) | 6.33 | 2.38 | 62.40% |
| | 3% BLER (result B) | 12.6 | 6.8 | 46.03% |
| 0.2 | **Average improvement with channel errors** | | | **40.88%** |

## OVERALL RELATIVE REDUCTION IN WORD ERROR RATE 24%

# High Data Rate comparison at 16kHz

Sampling rate = 16kHz
AMR mode = AMR-WB 12.65

| | | word error rate AMR-WB | word error rate DSR | Relative Improvement |
|---|---|---|---|---|
| **Digits** | Aurora-3 Spanish (Result A) | 7.5 | 4.6 | 38.67% |
| | Aurora-3 Spanish (Result B) | 7.39 | 3.47 | 53.04% |
| | Aurora-3 Italian | 14.77 | 5.62 | 61.95% |
| | US English In-Car (digits test) | 17.8 | 12.3 | 30.90% |
| | German In-Car (digit test) | 9.2 | 7.3 | 20.65% |
| | Japanese In-Car (digit test) | 11.3 | 8.4 | 25.66% |
| | US English In-Car (digits test) | 2.04 | 1.78 | 12.75% |
| | Mandarin Embedded PDA (digit test) | 1.8 | 1.14 | 36.67% |
| 0.35 | **Average improvement on digits tasks** | | | **35.04%** |
| **Subword** | Mandarin Embedded PDA | 2.29 | 1.63 | 28.82% |
| | US English In-Car | 2.35 | 2.31 | 1.70% |
| | US English In-Car | 13.2 | 7.8 | 40.91% |
| | German In-Car | 10.7 | 7.1 | 33.64% |
| | Japanese In-Car | 12.3 | 10.8 | 12.20% |
| 0.45 | **Average improvement on subword tasks** | | | **23.45%** |
| **Channel errors** | 1% BLER (result A) | 2.74 | 1.84 | 32.85% |
| | 1% BLER (result B) | 7.4 | 4.8 | 35.14% |
| | 3% BLER (result A) | 3.44 | 1.84 | 46.51% |
| | 3% BLER (result B) | 10.9 | 5 | 54.13% |
| 0.2 | **Average improvement with channel errors** | | | **42.16%** |

## OVERALL RELATIVE REDUCTION IN WORD ERROR RATE 31%

It is noted that for the evaluations at 16kHz one of the ASR vendors performed downsampling to 8kHz.