3GPP TSG-SA Codec Working Group
TSG-S#4: June 21-23, 1999, Miami, FL-USA

TSGS#4(99)248
Agenda Item: 5.4.3

Source: TSG-S4

3G TS 26.093 V2.0.0 (1999-06)

Technical Specification

3rd Generation Partnership Project (3GPP)
TSG-SA Codec Working Group
Mandatory Speech Codec speech processing functions
AMR Speech Codec; Source Controlled Rate operation

3GPP

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

Keywords
Adaptive Multi-Rate, Mandatory speech coder

3GPP

Postal address

Office address

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

Contents

Intellectual Property Rights

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project, Technical Specification Group Services and System Aspects, Working Group 4 (Codec).
The contents of this informal document may be subject to continuing work within the 3GPP and may change following formal TSG-S4 approval. Should TSG-S4 modify the contents of this document, it will be re-released with an identifying change of release date and an increase in version number as follows:

Version m.t.e
where:

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

1. Scope

This document describes the operation of the Adaptive Multi Rate speech codec during Source Controlled Rate (SCR) operation.

For clarity, the description is structured according to the block diagrams in figures 1 and 3.This structure of distributing the various functions between system entities is not mandatory for implementation, as long as the operation on the speech decoder output remains the same.

The SCR functions described in this technical specification are mandatory for implementation in the UEs. The receiver requirements are mandatory for implementation in all Transcoders , the transmitter requirements only for those links where SCR will be used.

Annex A-E describes the interworking operation of AMR with GSM-EFR, TDMA-EFR, TDMA-US1 and PDC-EFR. This mode of operation is F.F.S.

2. Normative references

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

[1] 3G TS 26.071 : "AMR Speech Codec; General description".

[2] 3G TS 26.073 : "AMR Speech Codec; ANSI-C code".

[3] 3G TS 26.074 : "AMR Speech Codec; Test sequences".

[4] 3G TS 26.090 : "AMR Speech Codec; Transcoding functions".

[5] 3G TS 26.091 : "AMR Speech Codec; Error concealment of lost frames".

[6] 3G TS 26.092 : "AMR Speech Codec; Comfort noise aspects".

[7] 3G TS 26.094 : "AMR Speech Codec; Voice Activity Detector (VAD)".

[8] 3G TS 26.101 : “AMR Speech Codec; Frame structure".

3. Definitions, symbols and abbreviations

3.1 Definitions

For the purpose of this document, the following definitions apply.

frame: Time interval of 20 msec. corresponding to the time segmentation of the Adaptive Multi Rate speech transcoder ,also used as a short term for a traffic frame.

traffic frame: Block of 95..244 information bits transmitted on the speech traffic channels.
SID frame: Frame that conveys information about the acoustic background noise.
speech frame: Traffic frame that has been classified as a SPEECH frame.
VAD flag: Boolean flag, generated by the VAD algorithm indicating the presence ("1") or the absence ("0") of a speech frame.

RX_TYPE: , generated by the de-framing unit , indicating to the RX SCR handler the type of data in the current frame. Refer to Table 2 for an example.
TX_TYPE: flag, generated by the TX SCR handler, indicating to the framin unit the type of data in the current frame. Refer to Table 1 for an example.
hangover period: A period of frames added at the end of a speech burst in which VAD flag ="0" and TX_TYPE is ="00", this period provides the encoder with an extra window to analyze the Comfort Noise parameters .

3.2 Symbols

For the purpose of this document, the following symbols apply.

$N_{elapsed}$ Number of elapsed frames since the last updated SID frame.

3.3 Abbreviations

For the purpose of thisdocument , the following abbreviations apply.

AN	Access Network
SCR	Source Controlled Rate operation
TS	Telecommunication Standard
GSM	Global System for Mobile Telecommunications
GSM-EFR	GSM Enhanced Full Rate speech codec
UE	User Equipment
PDC-EFR	ARIB PDC-EFR 6.7 kbit/s speech coder
RAN	Radio AN
RX	Receive
SID	Silence Descriptor (Background character Descriptor)
TDMA-EFR	TIA IS-641 Enhanced speech coder
TDMA-US1	TIA TDMA-US1 (12.2 kbit/s EFR)
TX	Transmit
VAD	Voice Activity Detector

4. General

Source Controlled Rate operation (SCR) is a mechanism in the AMR Speech Codec which allows the codec to encode speech at a lower average rate by taking speech inactivity into account. The SCR scheme may be used for the following purposes:

- to save power in the User Equipment;
- to reduce the overall load in the networks.

SCR shall be in operation in UEs if commanded so by the network.

4.1 General organisation

The default SCR mechanism described in this document requires the following functions:

- a Voice Activity Detector (VAD) on the transmit (TX) side;
- evaluation of the background acoustic noise on the transmit (TX) side, in order to transmit characteristic parameters to the receive (RX) side;
- generation on the receive (RX) side of a similar noise, called comfort noise, during periods where the transmission is switched off.

The Voice Activity Detector (VAD) is defined in [7] and the AMR-mode comfort noise functions in [6]. Both are based partly on the speech transcoder and its internal variables, defined in [4].
In addition to these functions, if the parameters arriving at the RX side are detected to be seriously corrupted by errors, the speech or comfort noise must be generated from substituted data in order to avoid seriously annoying effects for the listener. These functions for the AMR-mode are defined in [5].
An overall description of the speech processing parts can be found in [1]. An overview of one link SCR operation is shown in Figure 1.

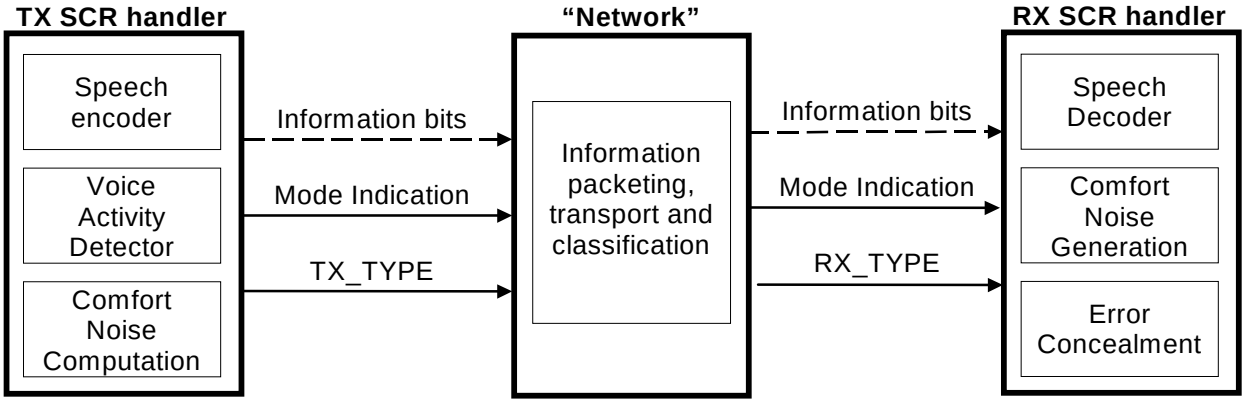


Figure 1: Block diagram of one link SCR operation

5. AMR SCR operation

5.1 Transmit (TX) side

A block diagram of the transmit side SCR functions is shown in Figure 2.

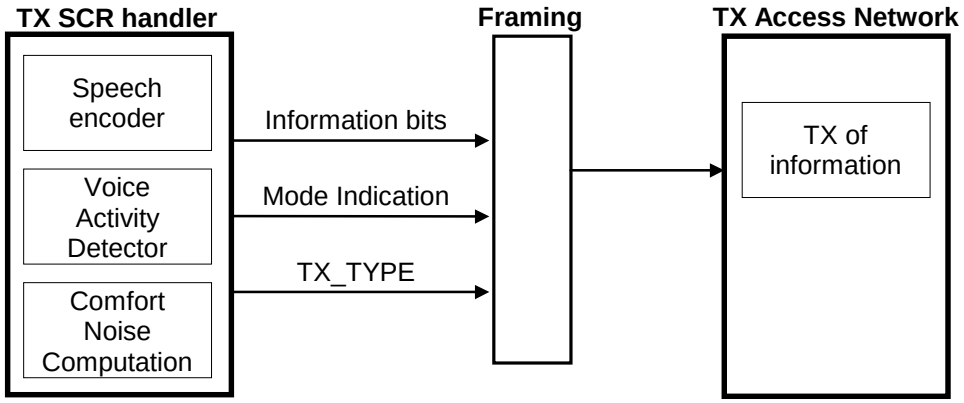


Figure 2: Block diagram of the transmit side SCR functions

5.1.1 General operation

The TX SCR handler passes traffic frames, individually marked by TX_TYPE, to the Framing unit. Each frame consists of bit fields containing the information bits, the codec mode indication, and the TX_TYPE. TX_TYPE shall be used to specify the contents of the frame. The table below provides an overview of the different TX_TYPEs used and explains the required contents in the information bit and the mode indication bit fields.

TX_TYPE	Legend	Information Bits	Mode Indication
00	SPEECH	speech frame, size depending on codec mode	current codec mode
01	SID_FIRST (end of speech marker, start of CN generation)	no useful information,	the codec mode that would have been used if TX_TYPE had been 00 (SPEECH)
10	SID_UPDATE	comfort noise information , information bits	the codec mode that would have been used if TX_TYPE had been 00 (SPEECH)
11	NO_DATA	no useful information	no useful information

Table 1: TX TYPE identifiers for AMR

TX_TYPE = “11” indicates that the Information Bit and Codec Mode fields do not contain any useful data (and should not be transmitted over AN). The purpose of this TX_TYPE is to provide the option to save network transmission between the transcoder and AN.

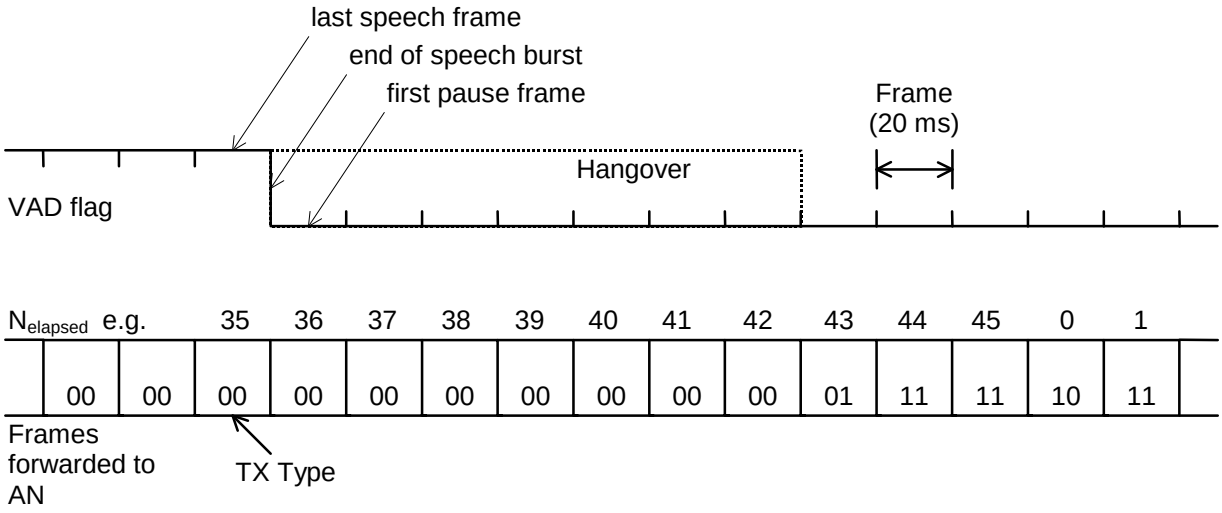
The scheduling of the frames for transmission on the Access Network is controlled by the TX SCR handler by the use of the TX_TYPE field and the given SCR operation mode.

5.1.2 Functions of the TX SCR handler

5.1.2.1 AMR SCR Timing procedures

To allow an exact verification of the TX SCR handler functions, all frames before the reset of the system are treated as if there were speech frames of an infinitely long time. Therefore, the first seven frames after the reset are always marked with TX_TYPE= "00"", even if VAD flag ="0" .
The Voice Activity Detector (VAD) shall operate all the time in order to assess whether the input signal contains speech or not. The output is a binary flag (VAD flag ="1" or VAD flag ="0", respectively) on a frame by frame basis.
The VAD flag controls indirectly, via the TX SCR handler operations described below, the overall SCR operation on the transmit side.
Whenever VAD flag ="1", the speech encoder output frame along with mode information shall be passed directly to the AN), marked with TX_TYPE ="00"

At the end of a speech burst (transition VAD flag ="1" to VAD flag ="0"), it takes eight consecutive frames to make a new updated SID analysis available. Normally, the first seven speech encoder output frames after the end of the speech burst shall therefore be passed directly to the AN, marked with TX_TYPE ="00" (SPEECH) ("hangover period").
The end of the speech is then indicated by passing frame eight after the end of the speech burst to the AN, marked with TX_TYPE = “01” (SID_FIRST) (see figure 2).



TX Types: “00” = SPEECH; “01” = SID_FIRST; “10” = “SID_UPDATE; “11” = NO DATA
$N_{elapsed}$: No. of elapsed frames since last SID_UPDATE

Figure 3: Normal hangover procedure for AMR ($N_{elapsed}$ > 23)

If, however, at the end of the speech burst, less than 24 frames have elapsed since the last SID_UPDATE frame was computed, then this last analysed SID_UPDATE frame should be passed to the AN whenever a SID_UPDATE frame (TX_TYPE=“10”) is to be produced, until a new updated SID analysis is available (8 consecutive frames marked with VAD flag ="0").This reduces the load on the network in cases where short background noise spikes are taken for speech, by avoiding the "hangover" waiting for the SID frame computation.

Once the first SID analysis after the end of a speech burst has been computed and the first SID frame has been passed to the AN, the TX SCR handler shall at regular intervals compute and pass updated SID_UPDATE (Comfort Noise) frames (TX_TYPE = “10”) to the AN) as long as VAD flag = "0". SID_UPDATE frames shall be generated every 8^{th} frame. However for frame loss robustness the first SID_UPDATE frame shall be sent as the third frame after the initial SID_FIRST frame.

Note: The speech encoder is operated in full speech modality if TX_TYPE = "00" and otherwise in a simplified mode, because not all encoder functions are required for the evaluation of comfort noise parameters and because comfort noise parameters are only to be generated at certain times.

5.1.3 The TX part of the AN

The TX part of the AN has the following overall functionality. The transmission is cut after the transmission of a SID frame when the speaker stops talking. During speech pauses the transmission is resumed at regular intervals for transmission of one SID_UPDATE frame, in order to update the generated comfort noise on the RX side. The transcoder decides what frames to send. In the case when nothing is to be transmitted it outputs frames marked with TX_TYPE = ”11”.

Demands on the TX part of the Access Network

The TX part of the AN operates in the following way regarding SCR:

- frames marked with TX_TYPE = “00” (SPEECH) are scheduled for transmission.
- frames marked with TX_TYPE = “01” (SID_FIRST) are scheduled for transmission.
- frames marked with TX_TYPE = “10” (SID_UPDATE) are scheduled for transmission
- for frames marked with TX_TYPE = “11” (NO_DATA) no processing or transmission is carried out.

SPEECH frames shall override possible SID_FIRST or SID_UPDATE frames in these exceptional cases.

5.2 Receive (RX) side

A block diagram of the receive side SCR functions is shown in Figure 3 below.

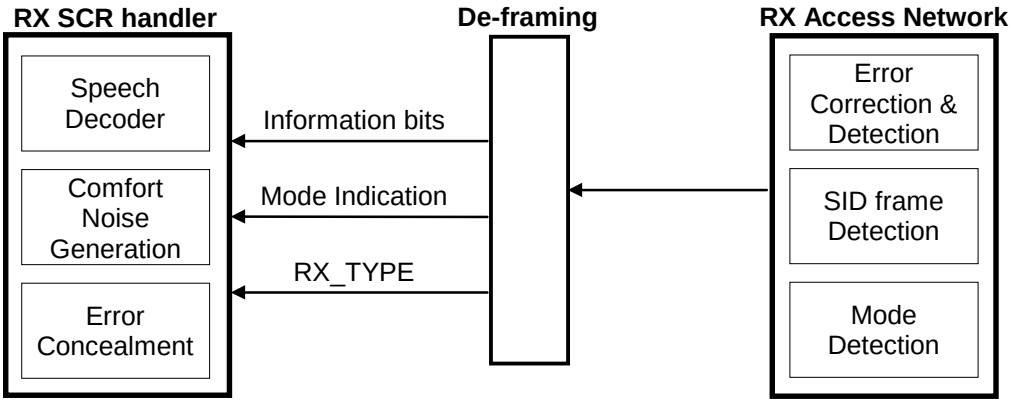


Figure 4: Block diagram of the receive side SCR functions

5.2.1 General operation

Whatever their context (speech, SID, or none), the deframing unit after AN continuously passes the received traffic frames to the RX SCR handler, individually marked by various pre-processing functions with a 3 bit type indicator RX_TYPE described in Table 2, which serve to classify the traffic frame. This classification allows the RX SCR handler to determine in a simple way how the received frame is to be handled.

RX_TYPE	Legend	Description
000	**SPEECH_GOOD**	Speech frame without errors in class A, and most likely not in class B (quality information in the RX part of the AN also OK)
001	**SPEECH_PROBABLY_DEGRADED**	Speech frame with CRC OK (class A OK), but less sensitive bits may be corrupted. At least 1 bit error in the speech coder's class B bits is suspected
010	**SPEECH_BAD**	(likely) speech frame with bad CRC (or estimated to be very bad by the RX part of the AN)
011	**SPARE**	Spare
100	**SID_FIRST**	This SID-frame marks the beginning of a comfort noise period.
101	**SID_UPDATE**	Correct SID update frame
110	**SID_BAD**	Corrupt SID update frame (bad CRC; applicable only for SID_UPDATE frames)
111	**NO_DATA**	Nothing useable was received. The synthesis mode of the previous frame type is used.

Table 2: RX_TYPE identifiers for AMR

5.2.2 RX part of the AN

The RX/de-framing unit uses a combination of measurements from AN , and CRC checks to classify each received frame according to RX_TYPE (Table 2).

5.2.3 Demands on the RX SCR handler

The RX SCR handler is responsible for the overall SCR operation on the RX side.
The SCR operation on the RX side shall be as follows:

- whenever a frame classified SPEECH_GOOD is received the SCR handler shall pass it directly on to the speech decoder;

- when a frame classified as SPEECH_PROBABLY_DEGRADED, SPEECH_BAD, SID_BAD is received the error concealment procedure(s) shall be applied. If not in Comfort Noise Generation mode, the RX SCR handler shall treat NO_DATA frames delivered by the AN as SPEECH_BAD frames without valid speech information.;

- frames classified as SID_UPDATE, SID_FIRST or SID_BAD shall result in Comfort Noise Generation mode, until the next SID_UPDATE frame has arrived or frames classified as SPEECH_OK or SPEECH_PROBABLY_DEGRADED are detected. During this period, the RX SCR handler shall ignore any unusable frames (NO_DATA) delivered by the AN;

- Frames classified as SPARE are handled as NO_DATA frames.

5.3 AMR SID Information format

When the TX SCR handler is ordered by the network to operate in AMR mode with SCR operation enabled the SID_UPDATE frame format is according to [5]. This is the default and only mandatory operating mode of the SCR handler.

Annex A: ETSI-AMR SCR handler

The ETSI-AMR SCR handler is indentical to the AMR SCR handler as described in the main body of this specification. Speech coding interworking aspects ETSI-AMR are described in [TBD].

Annex B: ETSI GSM-EFR SCR handler

The interworking operation of AMR (only 12.2 kbit/s mode) with GSM-EFR is F.F.S. This annex provides the basis for the AMR operation under this condition.

Transmit (TX) side

Speech coding internetworking aspects with ETSI-EFR are described in [TBD].

General operation

The TX SCR handler passes traffic frames, individually marked by TX_TYPE, to the TX Framing Unit part of the Access Network. Each frame passed to the AN consists of bit fields containing the information bits, the codec mode indication, and the TX_TYPE. TX_TYPE shall be used to specify the contents of the frame. The table below provides an overview of the different TX_TYPEs used and explains the required contents in the information bit and the mode indication bit fields.

TX_TYPE	Legend	Information Bits	Mode Indication
00	SPEECH	speech frame, size depending on codec mode	current codec mode
10	SID_UPDATE	comfort noise information , information bits	the codec mode that would have been used if TX_TYPE had been 00 (SPEECH)
11	NO_DATA	no useful information	no useful information

Table 3: TX TYPE identifiers for GSM-EFR

TX_TYPE = “NO_DATA” indicates that the Information Bit and Codec Mode fields do not contain any useful data (and should not be transmitted over AN). The purpose of this TX_TYPE is to provide the option to save network transmission between the transcoder and AN.
The scheduling of the frames for transmission on the Access Network is controlled by the TX SCR handler by the use of the TX_TYPE field and the given SCR operation mode.

Functions of the TX SCR handler

GSM-EFR SCR Timing procedures

To allow an exact verification of the TX SCR handler functions, all frames before the reset of the system are treated as if there were speech frames of an infinitely long time.
The SID_UPDATE timing is according to ETSI GSM 06.81.

The TX part of the AN

Demands on the TX part of the Access Network

The TX part of the AN operates in the following way regarding SCR:
- frames marked with TX_TYPE = (SPEECH) are scheduled for transmission.
- frames marked with TX_TYPE = (SID_UPDATE) are scheduled for transmission
- for frames marked with TX_TYPE = (NO_DATA) no processing or transmission is carried out.

SPEECH frames shall override other frames in these exceptional cases.

Receive (RX) side

Whatever their context (speech, SID, or none), the deframing unit after AN continuously passes the received traffic frames to the RX SCR handler, individually marked by various pre-processing functions with a 3 bit type indicator RX_TYPE described in Table 4, which serve to classify the traffic frame. This classification allows the RX SCR handler to determine in a simple way how the received frame is to be handled.

RX_TYPE	Legend	Description
000	**SPEECH_GOOD**	Speech frame with CRC OK, soft values in the RX part of AN also OK
001	**SPARE**	Spare
010	**SPEECH_BAD**	(likely) speech frame, bad CRC (or estimated to be very bad by the RX part of the AN)
011	**SPARE**	Spare
100	**SPARE**	Spare
101	**SID_UPDATE**	Correct SID update frame
110	**SID_BAD**	Corrupt SID update frame (bad CRC ; applicable only for SID_UPDATE frames)
111	**NO_DATA**	Nothing useable was received. The synthesis mode of the previous frame type is used.

Table 4: RX_TYPE identifiers for GSM-EFR

SID Information format

When the TX SCR handler is ordered by the network to operate in ETSI GSM-EFR mode with SCR operation turned on the SID_UPDATE frame format is according to ETSI GSM 06.62. (38 +5 bits).

Annex C: TIA IS-641 SCR Handler

The interworking operation of AMR (only 7.4 kbit/s mode) with IS-641 (TDMA-EFR) is F.F.S. This annex provides the basis for the AMR operation under this condition.

TX-side

F.F.S

RX-side

F.F.S

SID Information format

When the TX SCR handler is ordered by the network to operate in IS-641 mode with SCR operation turned on the SID_UPDATE frame format is according to IS-641-A Chapter 6 The SID_UPDATE timing is according to IS 641-A.

Speech coding internetworking aspects with IS-641 are described in [TBD].

Annex D: TIA TDMA-US1 SCR Handler

The interworking operation of AMR (only 12.2 kbit/s mode) with TDMA-US1 is F.F.S. This annex provides the basis for the AMR operation under this condition.

TX-side

F.F.S

RX-side

F.F.S

SID Information format

F.F.S.

Annex E: ARIB PDC-EFR SCR Handler

The interworking operation of AMR (only 6.7 kbit/s mode) with PDC-EFR-6.7 is F.F.S. This annex provides the basis for the AMR operation under this condition.

Transmit (TX) side

General operation

The TX SCR handler passes traffic frames, individually marked by TX_TYPE, to the TX Framing Unit part of the Access Network. Each frame passed to the AN consists of bit fields containing the information bits, the codec mode indication, and the TX_TYPE. TX_TYPE shall be used to specify the contents of the frame. The table below provides an overview of the different TX_TYPEs used and explains the required contents in the information bit and the mode indication bit fields.

TX_TYPE	Legend	Information Bits	Mode Indication
000	SPEECH	speech frame, size depending on codec mode	current code mode
010	POST1	comfort noise information , information bits	the codec mode that would have been used if TX_TYPE had been 00 (SPEECH)
011	NO_DATA	no useful information	no useful information
001	POST0 (end of speech marker, start of CN generation)	no useful information,	the codec mode that would have been used if TX_TYPE had been 00 (SPEECH)
100	PRE (end of CN marker, start of speech generation)	no useful information,	the codec mode that would have been used if TX_TYPE had been 00 (SPEECH)

Table 5: TX TYPE identifiers for PDC_EFR

TX_TYPE = “NO_DATA” indicates that the Information Bit and Codec Mode fields do not contain any useful data (and should not be transmitted over AN). The purpose of this TX_TYPE is to provide the option to save network transmission between the transcoder and AN.
The scheduling of the frames for transmission on the Access Network is controlled by the TX SCR handler by the use of the TX_TYPE field and the given SCR operation mode.

Speech coding internetworking aspects with PDC-EFR are described in XX.YY.D.[TBD]

Functions of the TX SCR handler

PDC_EFR SCR Timing procedures

To allow an exact verification of the TX SCR handler functions, all frames before the reset of the system are treated as if there were speech frames of an infinitely long time.
SID-frame timing is according to ARIB XXX.

The TX part of the AN

Demands on the TX part of the Access Network

The TX part of the AN operates in the following way regarding SCR:
- frames marked with TX_TYPE = (SPEECH) are scheduled for transmission.
- frames marked with TX_TYPE = (POST0) are scheduled for transmission.
- frames marked with TX_TYPE = (POST1) are scheduled for transmission
- frames marked with TX_TYPE = (PRE) are scheduled for transmission

- for frames marked with TX_TYPE = (NO_DATA) no processing or transmission is carried out.

SPEECH frames shall override other frames in these exceptional cases.

Receive (RX) side

Whatever their context (speech, SID, or none), the deframing unit after AN continuously passes the received traffic frames to the RX SCR handler, individually marked by various pre-processing functions with a 3 bit type indicator RX_TYPE described in Table 6, which serve to classify the traffic frame. This classification allows the RX SCR handler to determine in a simple way how the received frame is to be handled.

RX_TYPE	Legend	Description
000	**SPEECH_GOOD**	Speech frame with CRC OK, soft values in the RX part of AN also OK
001	**PRE**	Next frame should be a Speech frame
010	**SPEECH_BAD**	(likely) speech frame, bad CRC (or estimated to be very bad by the RX part of the AN)
011	**SPARE**	Spare
100	**POST0**	This SID-frame marks the beginning of a comfort noise period.
101	**POST1**	Correct SID update frame
110	**POST1_BAD**	Corrupt SID update frame (bad CRC; applicable only for POST1 frames)
111	**NO_DATA**	Nothing useable was received. The synthesis mode of the previous frame type is used.

Table 6: RX_TYPE identifiers for PDC_EFR

SID information format

When the TX SCR Handler is ordered by the network to operate in PDC-EFR mode with SCR turned on the SID_UPDATE frame format is according to ARIB TBD (134 bits).

History

Document history		
V. 0.1.0	April 1999	First Draft based on GSM 06.93 2.0.0, radio removed SID_FIRST used, additional SID formats referenced. Handover reference removed, Framing introduced, Action upon SPARE RX-type defined.
V 1.0.0	April 22 1999	Editorial changes
V. 2.0.0	June 1999	Presented at S#4 Plenary for approval