

3GPP TS 46.051 V11.0.0 (2012-09)

Technical Specification

3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Enhanced Full Rate (EFR) speech processing functions; General description (Release 11)



The present document has been developed within the 3rd Generation Partnership Project (3GPP™) and may be further elaborated for the purposes of 3GPP.

The present document has not been subject to any approval process by the 3GPP Organizational Partners and shall not be implemented. This Specification is provided for future development work within 3GPP only. The Organizational Partners accept no liability for any use of this Specification. Specifications and reports for implementation of the 3GPP™ system should be obtained via the 3GPP Organizational Partners' Publications Offices.

Keywords

GSM, speech, codec

3GPP

Postal address

3GPP support office address

650 Route des Lucioles - Sophia Antipolis
Valbonne - FRANCE
Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Internet

<http://www.3gpp.org>

Copyright Notification

No part may be reproduced except as authorized by written permission.
The copyright and the foregoing restriction extend to reproduction in all media.

© 2012, 3GPP Organizational Partners (ARIB, ATIS, CCSA, ETSI, TTA, TTC).
All rights reserved.

UMTS™ is a Trade Mark of ETSI registered for the benefit of its members
3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners
LTE™ is a Trade Mark of ETSI currently being registered for the benefit of its Members and of the 3GPP Organizational Partners
GSM® and the GSM logo are registered and owned by the GSM Association

Contents

Foreword	4
1 Scope	5
2 References.....	5
3 Definitions and abbreviations	5
3.1 Definitions	5
3.2 Abbreviations	6
4 General.....	6
5 Enhanced Full Rate speech channel transcoding	7
6 Enhanced Full Rate speech channel discontinuous transmission (DTX)	7
7 Enhanced Full Rate speech channel Voice Activity Detection (VAD).....	8
8 Enhanced Full Rate speech channel comfort noise insertion	8
9 Enhanced Full Rate speech channel lost speech frame substitution and muting	9
10 Enhanced Full Rate codec homing.....	9
11 Alternative Enhanced Full Rate implementation using the Adaptive Multi Rate 12.2 kbit/s mode	9
Annex A (informative): Change History.....	11

Foreword

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

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

Version x.y.z

where:

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

1 Scope

The present document is an introduction to GSM 06.60 [6], GSM 06.61 [7], GSM 06.62 [8], GSM 06.81 [9] and GSM 06.82 [10] ENs dealing with the speech processing functions in the Enhanced Full Rate channel of the GSM system. A general overview of the speech processing functions is given, with reference to the ENs where each function is specified in detail.

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] GSM 01.04: "Digital cellular telecommunications system (Phase 2+); Abbreviations and acronyms".
- [2] GSM 03.50: "Digital cellular telecommunications system (Phase 2+); Transmission planning aspects of the speech service in the GSM Public Land Mobile Network (PLMN) system".
- [3] GSM 05.03: "Digital cellular telecommunications system (Phase 2+); Channel coding".
- [4] GSM 06.53: "Digital cellular telecommunications system (Phase 2+); ANSI-C code for the GSM Enhanced Full Rate (EFR) speech codec".
- [5] GSM 06.54: "Digital cellular telecommunications system (Phase 2+); Test vectors for the GSM Enhanced Full Rate (EFR) speech codec".
- [6] GSM 06.60: "Digital cellular telecommunications system (Phase 2+); Enhanced Full Rate (EFR) speech transcoding".
- [7] GSM 06.61: "Digital cellular telecommunications system (Phase 2+); Substitution and muting of lost frame for Enhanced Full Rate (EFR) speech traffic channels".
- [8] GSM 06.62: "Digital cellular telecommunications system (Phase 2+); Comfort noise aspects for Enhanced Full Rate (EFR) speech traffic channels".
- [9] GSM 06.81: "Digital cellular telecommunications system (Phase 2+); Discontinuous transmission (DTX) for Enhanced Full Rate (EFR) speech traffic channels".
- [10] GSM 06.82: "Digital cellular telecommunications system (Phase 2+); Voice Activity Detector (VAD) for Enhanced Full Rate (EFR) speech traffic channels".

3 Definitions and abbreviations

3.1 Definitions

Definition of terms used in the present document can be found in GSM 06.60 [6], GSM 06.61 [7], GSM 06.62 [8], GSM 06.81 [9] and GSM 06.82 [10].

3.2 Abbreviations

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

ACELP	Algebraic Code Excited Linear Prediction
BFI	Bad Frame Indication
BSS	Base Station System
CCITT	Comité Consultatif International Télégraphique et Téléphonique
DTX	Discontinuous Transmission
ETS	European Telecommunication Standard
GSM	Global System for Mobile communications
MS	Mobile Station
PCM	Pulse Code Modulated
PLMN	Public Land Mobile Network
PSTN	Public Switched Telephone Network
RF	Radio Frequency
RSS	Radio SubSystem
RX	Receive
SACCH	Slow Associated Control CHannel
SID	SIlence Descriptor
SP flag	SPeech flag
TAF	Time Alignment Flag
TX	Transmit

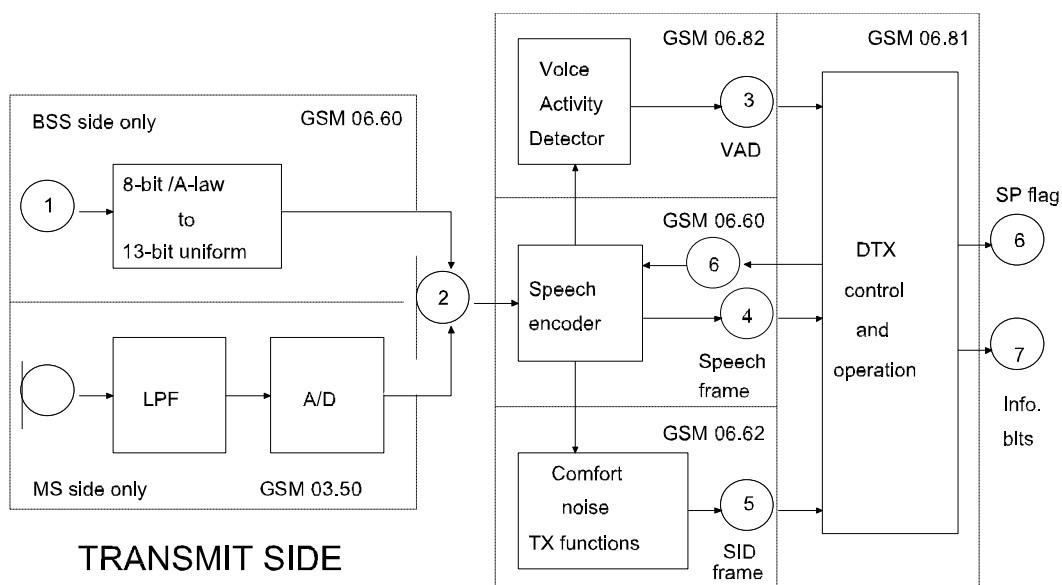
For abbreviations not given in this subclause, see GSM 01.04 [1].

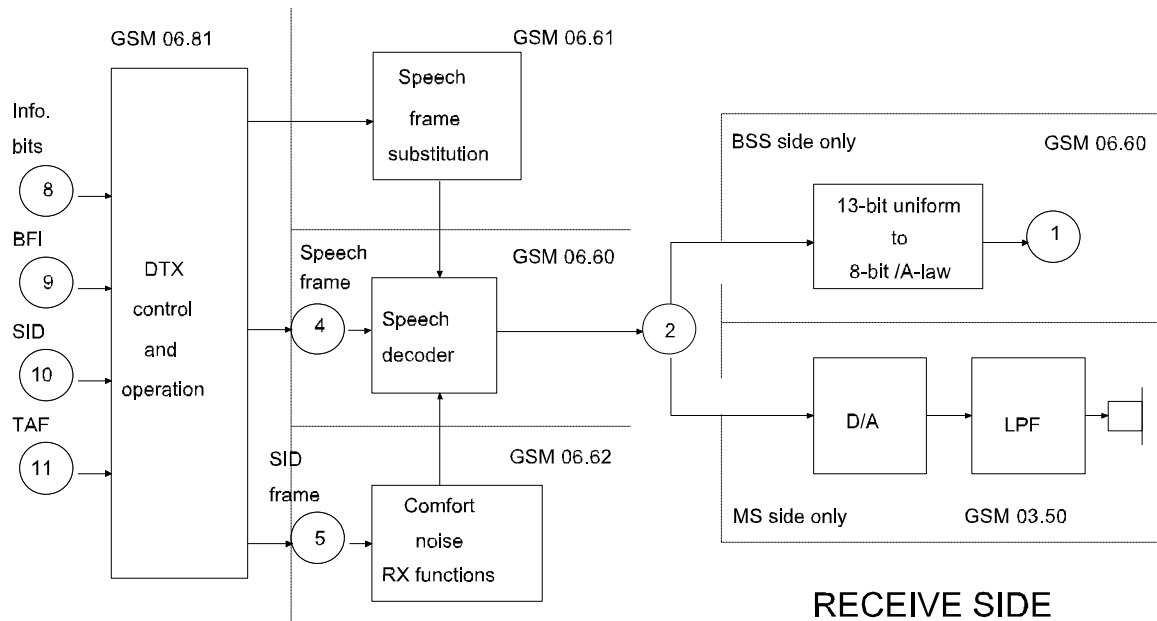
4 General

Figure 1 presents a reference configuration where the various speech processing functions are identified. In this figure, the relevant Standards for each function are also indicated.

In figure 1, the audio parts including analogue to digital and digital to analogue conversion are included, to show the complete speech path between the audio input/output in the Mobile Station (MS) and the digital interface of the PSTN. The detailed specification of the audio parts are contained in GSM 03.50 [2]. These aspects are only considered to the extent that the performance of the audio parts affect the performance of the speech transcoder.

An alternative and fully interoperable implementation using as a basis the 12.2 kbit/s mode of the Adaptive Multi Rate speech coder is described in section 11.





- 1) 8-bit /A-law or μ -law (PCS 1900) PCM (CCITT recommendation G.711), 8 000 samples/s.
- 2) 13-bit uniform PCM, 8 000 samples/s.
- 3) Voice Activity Detector (VAD) flag.
- 4) Encoded speech frame, 50 frames/s, 244 bits/frame.
- 5) Silence Descriptor (SID) frame, 244 bits/frame.
- 6) SPeech (SP) flag, indicates whether information bits are speech or SID information.
- 7) Information bits delivered to the radio subsystem.
- 8) Information bits received from the radio subsystem.
- 9) Bad Frame Indication.
- 10) Silence Descriptor (SID) flag.
- 11) Time Alignment Flag (TAF), marks the position of the SID frame within the Slow Associated Control Channel (SACCH) multiframe.

Figure 1: Overview of audio processing functions

5 Enhanced Full Rate speech channel transcoding

As shown in figure 1, the speech encoder takes its input as a 13-bit uniform Pulse Code Modulated (PCM) signal either from the audio part of the Mobile Station or on the network side, from the Public Switched Telephone Network (PSTN) via an 8-bit/A-law or μ -law (PCS 1900) to 13-bit uniform PCM conversion. The encoded speech at the output of the speech encoder is delivered to the channel coding function defined in GSM 05.03 [3] to produce an encoded block consisting of 456 bits leading to a gross bit rate of 22,8 kbit/s.

In the receive direction, the inverse operations take place. GSM 06.60 [6] describes the detailed mapping between input blocks of 160 speech samples in 13-bit uniform PCM format to encoded blocks of 244 bits and from encoded blocks of 244 bits to output blocks of 160 reconstructed speech samples. The sampling rate is 8 000 sample/s leading to a bit rate for the encoded bit stream of 12,2 kbit/s. The coding scheme is the so-called Algebraic Code Excited Linear Prediction, hereafter referred to as ACELP.

GSM 06.60 [6] describes the codec and GSM 06.53 [4] defines the C code, thus enabling the verification of compliance to GSM 06.60 [6] to a high degree of confidence by use of a set of digital test sequences given in GSM 06.54 [5].

6 Enhanced Full Rate speech channel discontinuous transmission (DTX)

During a normal phone conversation, the participants alternate so that, on the average, each direction of transmission is occupied about 50 % of the time. Discontinuous transmission (DTX) is a mode of operation where the transmitters are switched on only for those frames which contain useful information. This may be done for the following two purposes:

- 1) In the MS, battery life will be prolonged or a smaller battery could be used for a given operational duration.
- 2) The average interference level over the air interface is reduced, leading to better Radio Frequency (RF) spectrum efficiency.

The overall DTX mechanism is implemented in the DTX handlers (Transmit (TX) and Receive (RX)) described in GSM 06.81 [9] and requires the following functions:

- a Voice Activity Detector (VAD) on the TX side, see GSM 06.82 [10];
- evaluation of the background acoustic noise on the TX side, in order to transmit characteristic parameters to the RX side, see GSM 06.62 [8];
- generation of comfort noise on the RX side during periods where the radio transmission is turned off, see GSM 06.62 [8].

The transmission of comfort noise information to the RX side is achieved by means of a Silence Descriptor (SID) frame. The SID frame is transmitted at the end of speech bursts and serves as an end of speech marker for the RX side. In order to update the comfort noise characteristics at the RX side, SID frames are transmitted at regular intervals also during speech pauses. This also serves the purpose of improving the measurement of the radio link quality by the Radio SubSystem (RSS).

The DTX handlers interwork with the RSS using flags. The RSS is in control of the actual transmitter keying on the TX side, and performs various pre-processing functions on the RX side. This is described in GSM 06.81 [9].

The speech flag (SP) indicates whether information bits are speech or SID information. The SP flag is calculated from the VAD flag by the TX DTX handler. When SID information is transmitted (SP="0") the operation of the speech encoder is modified to reduce the remaining computation for that frame. This is described in GSM 06.62 [8].

7 Enhanced Full Rate speech channel Voice Activity Detection (VAD)

The Enhanced Full Rate VAD function is described in GSM 06.82 [10].

The input to the VAD is a set of parameters computed by the Enhanced Full Rate speech encoder defined in GSM 06.60 [6]. The VAD uses this information to decide whether each 20 ms speech coder frame contains speech or not. Note that the VAD flag is an input to TX DTX handler and does not control the transmitter keying directly.

GSM 06.82 [10] describes the VAD algorithm and GSM 06.53 [4] defines the C code. The verification of compliance to GSM 06.82 [10] is achieved by use of digital test sequences (see GSM 06.54 [5]) applied to the same interface as the test sequences for the speech codec.

8 Enhanced Full Rate speech channel comfort noise insertion

The Enhanced Full Rate noise comfort insertion function is described in GSM 06.62 [8].

When switching the transmission on and off during DTX operation, the effect would be a modulation of the background noise at the receiving end, if no precautions were taken. When transmission is on, the background noise is transmitted together with the speech to the receiving end. As the speech burst ends, the connection is off and the perceived noise would drop to a very low level. This step modulation of noise may be perceived as annoying and reduce the intelligibility of speech, if presented to a listener without modification.

This "noise contrast effect" is reduced in the GSM system by inserting an artificial noise, termed comfort noise, at the receiving end when speech is absent.

The comfort noise processes are as follows:

- the evaluation of the acoustic background noise in the transmitter;

- the noise parameter encoding (SID frames) and decoding;
- and the generation of comfort noise in the receiver.

The comfort noise processes and the algorithm for updating the noise parameters during speech pauses are defined in detail in GSM 06.62 [8].

The comfort noise mechanism is based on the Enhanced Full Rate speech codec defined in GSM 06.60 [6].

9 Enhanced Full Rate speech channel lost speech frame substitution and muting

The Enhanced Full Rate speech frame substitution and muting function is described in GSM 06.61 [7].

In the receiver, frames may be lost due to transmission errors or frame stealing. GSM 06.61 [7] describes the actions to be taken in these cases, both for lost speech frames and for lost SID frames in DTX operation.

In order to mask the effect of an isolated lost frame, the lost speech frame is substituted by a predicted frame based on previous frames. Insertion of silence frames is not allowed. For several subsequent lost frames, a muting technique shall be used to indicate to the listener that transmission has been interrupted.

10 Enhanced Full Rate codec homing

The GSM Enhanced Full Rate speech transcoder, VAD, DTX system and comfort noise parts of the audio processing functions (see figure 1) are defined in bit exact arithmetic. Consequently, they shall react on a given input sequence always with the corresponding bit exact output sequence, provided that the internal state variables are also always exactly in the same state at the beginning of the experiment.

The input test sequences provided in GSM 06.54 [5] shall force the corresponding output test sequences, provided that the tested modules are in their home-state when starting.

The modules may be set into their home states by provoking the appropriate homing-functions.

NOTE: This is normally done during reset (initialization of the codec).

Special inband signalling frames (encoder-homing-frame and decoder-homing-frame) described in GSM 06.60 [6] have been defined to provoke these homing-functions also in remotely placed modules.

This mechanism is specified to support three main areas:

- type approval of mobile terminal equipment;
- type approval of infrastructure equipment;
- remote control and testing for operation and maintenance.

At the end of the first received homing frame, the audio functions that are defined in a bit exact way shall go into their predefined home states. The output corresponding to the first homing frame is dependent on the codec state when the frame was received. Any consecutive homing frames shall produce corresponding homing frames at the output.

11 Alternative Enhanced Full Rate implementation using the Adaptive Multi Rate 12.2 kbit/s mode

The 12.2 kbit/s mode of the Adaptive Multi Rate speech coder described in TS 26.071 is functionally equivalent to the GSM Enhanced Full Rate speech coder. An alternative implementation of the Enhanced Full Rate speech service based on the 12.2 kbit/s mode of the Adaptive Multi Rate coder is allowed. Alternative implementations shall implement the functionality specified in TS 26.071 for the 12.2 kbit/s mode, with the difference that the DTX transmission format

from GSM 06.81, the comfort noise generation from GSM 06.62 and the decoder-homing-frame from GSM 06.60 shall be used.

Verification of compliance using the alternative implementation is achieved by use of a set of digital test sequences given in GSM 06.54.

NOTE: The alternative implementation of the GSM Enhanced Full Rate speech coder incurs an additional 5ms look-ahead delay, and the implementation in this specification (GSM 06.51) is the preferred option.

Annex A (informative): Change History

Change history					
SMG No.	TDoc. No.	CR. No.	Section affected	New version	Subject/Comments
SMG#21				4.0.1	ETSI Publication
SMG#20				5.1.2	Release 1996 version
SMG#27				6.0.0	Release 1997 version
SMG#28	P-99-139	A003	Figure 1, Sect.5	7.0.0	Addition of mu-law (PCS 1900)
				7.0.2	Update to Version 7.0.2 for Publication
SMG#31				8.0.0	Release 1999 version
SMG#32	P-00-273	A008	4 and 11	8.1.0	Alternative EFR implementation using the AMR 12.2 mode

Change history							
Date	TSG SA#	TSG Doc.	CR	Rev	Subject/Comment	Old	New
12-2000	10	SP-000572	A013		Definition of the homing frame for the alternative EFR implementation	8.1.0	8.2.0
03-2001	11				Version for Release 4		4.0.0
06-2002	16				Version for Release 5	4.0.0	5.0.0
12-2004	26				Version for Release 6	5.0.0	6.0.0
06-2007	36				Version for Release 7	6.0.0	7.0.0
12-2008	42				Version for Release 8	7.0.0	8.0.0
12-2009	46				Version for Release 9	8.0.0	9.0.0
03-2011	51				Version for Release 10	9.0.0	10.0.0
09-2012	57				Version for Release 11	10.0.0	11.0.0