



EUROPEAN
TELECOMMUNICATION
STANDARD

ETS 300 960

May 1997

Source: ETSI TC-SMG

Reference: DE/SMG-110601Q

ICS: 33.020

Key words: Digital cellular telecommunications system, Global System for Mobile communications (GSM)



**Digital cellular telecommunications system;
Full rate speech;
Processing functions
(GSM 06.01 version 5.0.1)**

ETSI

European Telecommunications Standards Institute

ETSI Secretariat

Postal address: F-06921 Sophia Antipolis CEDEX - FRANCE

Office address: 650 Route des Lucioles - Sophia Antipolis - Valbonne - FRANCE

X.400: c=fr, a=atlas, p=etsi, s=secretariat - **Internet:** secretariat@etsi.fr

Tel.: +33 4 92 94 42 00 - Fax: +33 4 93 65 47 16

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.

Contents

Foreword	5
1 Scope	7
2 Normative references	7
3 Abbreviations.....	7
4 Introduction.....	8
5 Full rate speech transcoding (GSM 06.10)	8
6 Full rate Discontinuous Transmission (DTX) (GSM 06.31).....	8
7 Full rate Voice Activity Detection (VAD) (GSM 06.32)	9
8 Full rate comfort noise insertion (GSM 06.12)	9
9 Full rate lost speech frame substitution and muting (GSM 06.11).....	9
10 Speech transmission functions of the half rate traffic channel.....	9
History.....	11

Blank page

Foreword

This European Telecommunication Standard (ETS) has been produced by the Special Mobile Group (SMG) Technical Committee (TC) of the European Telecommunications Standards Institute (ETSI).

This ETS describes in general terms the speech processing functions within the digital cellular telecommunications system.

This ETS is a GSM technical specification version 5. The ETS from which this ETS has evolved is GSM Phase 2 ETS 300 580-1 (GSM 06.01 version 4.0.5).

The specification from which this ETS has been derived was originally based on CEPT documentation, hence the presentation of this ETS may not be entirely in accordance with the ETSI/PNE Rules.

Transposition dates	
Date of adoption:	18 April 1997
Date of latest announcement of this ETS (doa):	31 August 1997
Date of latest publication of new National Standard or endorsement of this ETS (dop/e):	28 February 1998
Date of withdrawal of any conflicting National Standard (dow):	28 February 1998

Blank page

1 Scope

This European Telecommunication Standard (ETS) is an introduction to the 06 series of the GSM technical specifications dealing with the Full rate speech processing functions in the GSM system. A general overview of the speech processing parts is given with reference to the technical specifications where each part is specified in detail.

2 Normative references

This ETS 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 ETS only when incorporated in it by amendment or revision. For undated references, the latest edition of the publication referred to applies.

- [1] GSM 01.04 (ETR 350): "Digital cellular telecommunications system (Phase 2+) Abbreviations and acronyms".
- [2] GSM 03.50 (ETS 300 903): "Digital cellular telecommunications system (Phase 2+); Transmission planning aspects of the speech service in the GSM Public Land Mobile Network (PLMN) system".
- [3] GSM 06.10 (ETS 300 961): "Digital cellular telecommunications system; Full rate speech; Transcoding".
- [4] GSM 06.11 (ETS 300 962): "Digital cellular telecommunications system; Full rate speech; Substitution and muting of lost frames for full rate speech channels".
- [5] GSM 06.12 (ETS 300 963): "Digital cellular telecommunications system; Full rate speech; Comfort noise aspect for full rate speech traffic channels".
- [6] GSM 06.31 (ETS 300 964): "Digital cellular telecommunications system; Full rate speech; Discontinuous Transmission (DTX) for full rate speech traffic channels".
- [7] GSM 06.32 (ETS 300 965): "Digital cellular telecommunications system; Voice Activity Detector (VAD)".
- [8] GSM 06.02 (ETS 300 966): "Digital cellular telecommunications system; Half rate speech; Half rate speech processing functions".
- [9] GSM 06.20 (ETS 300 969): "Digital cellular telecommunications system; Half rate speech; Half rate speech transcoding".
- [10] GSM 06.21 (ETS 300 970): "Digital cellular telecommunications system; Half rate speech; Substitution and muting of lost frames for half rate speech traffic channels".
- [11] GSM 06.22 (ETS 300 971): "Digital cellular telecommunications system; Half rate speech; Comfort noise aspects for half rate speech traffic channels".
- [12] GSM 06.41 (ETS 300 972): "Digital cellular telecommunications system; Half rate speech; Discontinuous Transmission (DTX) for half rate speech traffic channels".
- [13] GSM 06.42 (ETS 300 973): "Digital cellular telecommunications system; Half rate speech; Voice Activity Detector (VAD) for half rate speech traffic channels".

3 Abbreviations

Abbreviations used in this ETS are listed in GSM 01.04 [1].

4 Introduction

Figure 1 presents a reference configuration where the various speech processing functions are identified. In this figure, the relevant technical specifications 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 to the PSTN. The specification of the audio parts are contained in GSM 03.50 [2] (Transmission plan aspects of the speech service in the GSM PLMN system). These aspects are only considered in the GSM 06-series to the extent that the performance of the audio parts affect the performance of the speech transcoder.

5 Full rate speech transcoding (GSM 06.10)

As shown in figure 1, the speech encoder takes its input as a 13 bit uniform PCM signal either from the audio part of the mobile station or on the network side, from the PSTN via an 8 bit/A-law 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 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 06.10 [3] describes the detailed mapping between input blocks of 160 speech samples in 13 bit uniform PCM format to encoded blocks of 260 bits and from encoded blocks of 260 bits to output blocks of 160 reconstructed speech samples. The sampling rate is 8 000 sample/s leading to an average bit rate for the encoded bit stream of 13 kbit/s. The coding scheme is the so-called Regular Pulse Excitation - Long Term prediction - Linear Predictive Coder, here-after referred to as RPE-LTP.

The technical specification describes the codec down to the bit level, thus enabling the verification of compliance to the technical specification to a high degree of confidence by use of a set of digital test sequences. These test sequences are also described and are available on floppy disks.

6 Full rate Discontinuous Transmission (DTX) (GSM 06.31)

During a normal 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 on the "air" is reduced, leading to better spectrum efficiency.

The overall DTX mechanism is implemented in the DTX handlers (TX and RX) described in GSM 06.31 [6] and requires the following functions which are described in separate technical specifications:

- A Voice Activity Detector on the transmit side.
- Evaluation of the background acoustic noise on the transmit side, in order to transmit characteristic parameters to the receive side.
- Generation on the receive side of a similar noise, called comfort noise, during periods where the radio transmission is cut.

The transmission of comfort noise information to the receive side is achieved by means of a special frame (Silence descriptor = SID). This frame is transmitted at the end of speech bursts and serves as an end of speech marker for the receive side. In order to update the comfort noise characteristics at the receive 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.

For the overall DTX functionality, the DTX handlers interwork via various flags with the Radio Subsystem, which is in control of the actual transmitter keying on the TX side and which performs various pre-processing functions on the RX side. This is also described in GSM 06.31 [6].

A common terminology used throughout the GSM 06-series of technical specifications is also defined in this ETS.

7 Full rate Voice Activity Detection (VAD) (GSM 06.32)

The input to the VAD is a set of parameters computed by the full-rate speech encoder defined in GSM 06.10 [3]. 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.

The ETS describes the VAD algorithm down to the bit level. The conventions used in the bit-exact specification are the same as those used in GSM 06.10 [3]. The verification of compliance to the technical specification is achieved by use of digital test sequences applied to the same interface as the test sequences for the speech codec. These test sequences are also described and are available on floppy disks.

8 Full rate comfort noise insertion (GSM 06.12)

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 is perceived as very annoying and may reduce the intelligibility of speech if presented to a listener without modification.

This so-called "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.

GSM 06.12 [5] deals with the detailed aspects of the comfort noise process:

- 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 algorithm for updating the noise parameters during speech pauses is also defined.

The comfort noise mechanism is based on the full rate speech codec defined in GSM 06.10 [3].

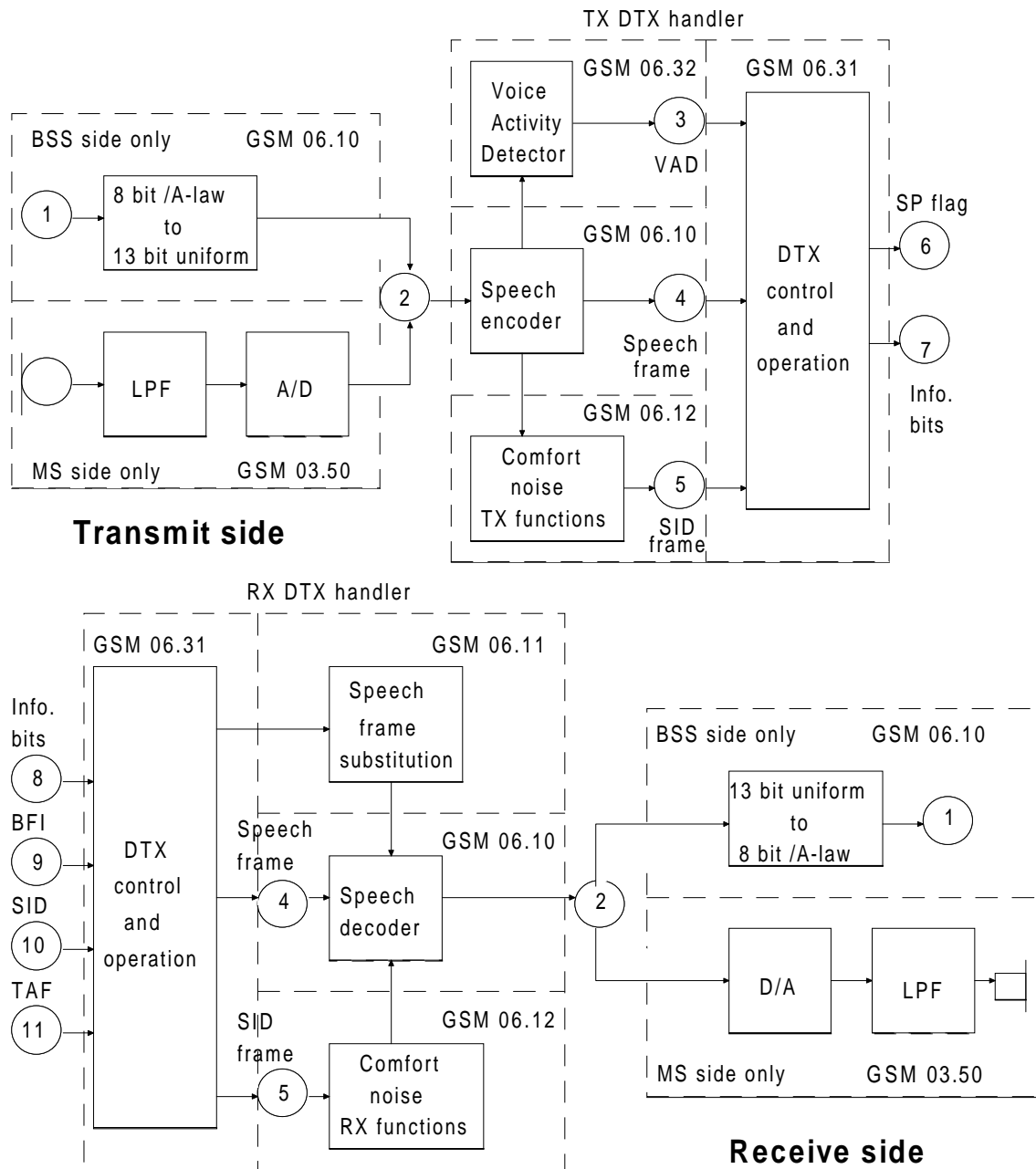
9 Full rate lost speech frame substitution and muting (GSM 06.11)

In the receiver, frames may be lost due to transmission errors or frame stealing. GSM 06.11 [4] 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, a scheme is used, where the lost speech frame is substituted by a predicted frame based on previous frames. Insertion of silence frames is not allowed. For several lost frames in a row, some muting technique must be used to indicate to the customer that transmission is interrupted.

10 Speech transmission functions of the half rate traffic channel

The gross bit rate of the half rate speech traffic channel is 11,4 kbit/s. The algorithm for the half rate speech transcoder as well as the associated DTX functions are defined in GSM 06.02 [8], GSM 06.20 [9], GSM 06.21 [10], GSM 06.22 [11], GSM 06.41 [12] and GSM 06.42 [13].



- (1) 8 bit /A-law PCM (ITU-T recommendation G.711), 8 000 samples/s.
- (2) 13 bit uniform PCM, 8 000 samples/s.
- (3) Voice activity flag.
- (4) Encoded speech frame, 50 samples/s, 260 bits/frame.
- (5) Silence Descriptor (SID) frame, 260 bits/frame.
- (6) Speech 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 (BFI) flag.
- (10) Silence Descriptor (SID) flag.
- (11) Time Alignment Flag (TAF), marks the position of the SID frame within the SACCH multiframe.

Figure 1: Overview of audio processing functions

History

Document history	
December 1996	Unified Approval Procedure UAP 61: 1996-12-16 to 1997-04-11
May 1997	First Edition