NORTH ATLANTIC TREATY ORGANIZATION
ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD

NATO STANDARDIZATION AGENCY (NSA)
AGENCE OTAN DE NORMALISATION (AON)

1110 BRUSSELS

NSA/0188-C3/4479

25 February 2002

STANAG 4479 C3 (EDITION 1) - PARAMETERS AND CODING CHARACTERISTICS THAT MUST BE COMMON TO ASSURE INTEROPERABILITY OF 800 BPS DIGITAL SPEECH ENCODER/DECODER AND THE ASSOCIATED ERROR PROTECTION AND INTERLEAVING SCHEMES LEADING TO THE 2400 BPS INTERFACE OF THE SLOW FREQUENCY HOPPING HF-EPM SYSTEM

Reference: AC/322(SC/6)N/103 dated 29 September 1998

1. The enclosed NATO Standardization Agreement which has been ratified by nations as reflected in page (iii) is promulgated herewith.

2. The reference listed above is to be destroyed in accordance with local document destruction procedures.

3. AAP-4 should be amended to reflect the latest status of the STANAG.

ACTION BY NATIONAL STAFFS

4. National staffs are requested to examine page (iii) of the STANAG and, if they have not already done so, advise the NQHC3S through their national delegation as appropriate of their intention regarding its ratification and implementation.

Jan H ERIKSEN
Rear Admiral, NONA
Director, NSA

Enclosure:
STANAG 4479 (Edition 1)

ORGANISATION DU TRAITE DE L’ATLANTIQUE NORD
(OTAN)

AGENCE OTAN DE NORMALISATION
(AON)

ACCORD DE NORMALISATION
(STANAG)

OBJET: PARAMETRES ET CARACTERISTIQUES DE CODAGE QUI DOIVENT ETRE COMMUNS POUR ASSURER L’INTEROPERABILITE DES EQUIPEMENTS DE PHONIE NUMERIQUE A 800 BITS/S ET SCHEMA DU CODE CORRECTEUR ET DE L’ENTRELACEUR ASSOCIES QUI PERMETTENT D’ATTEINDRE UN DEBIT DE 2400 BITS/S CORRESPONDANT AU SYSTEME A EVASION DE FREQUENCE HF-MPE

Promulgué le 25 février 2002

Jan H ERIKSEN
Vice-amiral, NONA
Directeur de l'AON

RECORD OF AMENDMENTS

No.	Reference/date of amendment	Date entered	Signature

EXPLANATORY NOTES

AGREEMENT

1. This NATO Standardization Agreement (STANAG) is promulgated by the Director, NSA under the authority vested in him by the NATO Military Committee.

2. No departure may be made from the agreement without consultation with the tasking authority. Nations may propose changes at any time to the tasking authority where they will be processed in the same manner as the original agreement.

3. Ratifying nations have agreed that national orders, manuals and instructions implementing this STANAG will include a reference to the STANAG number for purposes of identification.

DEFINITIONS

4. Ratification is "In NATO Standardization, the fulfilment by which a member nation formally accepts, with or without reservation, the content of a Standardization Agreement" (AAP-6).

5. Implementation is "In NATO Standardization, the fulfilment by a member nation of its obligations as specified in a Standardization Agreement" (AAP-6).

6. Reservation is "In NATO Standardization, the stated qualification by a member nation that describes the part of a Standardization Agreement that it will not implement or will implement only with limitations" (AAP-6).

RATIFICATION, IMPLEMENTATION AND RESERVATIONS

7. Page iii gives the details of ratification and implementation of this agreement. If no details are shown it signifies that the nation has not yet notified the tasking authority of its intentions. Page iv (and subsequent) gives details of reservations and proprietary rights that have been stated.

FEEDBACK

8. Any comments concerning this publication should be directed to NATO Standardization Agency (NSA) - NATO HQ - Bvd Leopold III - 1110 Brussels - BE

AIR/ARMY/NAVY

NATO STANDARDIZATION AGREEMENT
(STANAG)

PARAMETERS AND CODING CHARACTERISTICS THAT MUST BE COMMON TO ASSURE INTEROPERABILITY OF 800 BPS DIGITAL SPEECH ENCODER/DECODER AND THE ASSOCIATED ERROR PROTECTION AND INTERLEAVING SCHEMES LEADING TO THE 2400 BPS INTERFACE OF THE SLOW FREQUENCY HOPPING HF-EPM SYSTEM

ANNEXES

A. Description of the 800 bps speech coder.
B. Voice digitizer characteristics.
C. Coding and decoding tables.
D. Framing at 800 bps.
E. Example of speech performances (for information only).
F. Error detection & correction and interleaver for slow frequency hopping HF-EPM.
G. End Of Transmission for slow frequency hopping HF-EPM.
H. Late Traffic Entry capability for slow frequency hopping HF-EPM.
I. Functional interface with the Modem for slow frequency hopping HF-EPM.
J. Example of a Speech Smoothing Process (for information only).

RELATED DOCUMENT

STANAG 4198: Parameters and Coding Characteristics that must be common to assure Interoperability of 2400 bps Linear Predictive Encoded Digital Speech.

AIM

1. The aim of this agreement is to define:

   a. The voice digitizer characteristics, the coding tables and the bit format requirements at 800 bps;

   b. The error correction scheme and the interleaver to ensure compatibility of digital voice produced using an overall bit rate of 2400 bps for slow frequency hopping HF-EPM transmission; and

   c. The End Of Transmission sequence for slow frequency hopping HF-EPM transmission.

2. The point 1.a. and the related annexes A,B,C,D, describe a 800 bps coder useable as a stand alone voice coder in any transmission system needing a low bit rate coder.

3. The points 1.b. and 1.c. (and the related annexes F,G,H,I) are specifically aimed at transmission of the aforementionned coder with the Slow frequency hopping HF-EPM waveform.

AGREEMENT

4. Participating nations agree to use the characteristics contained in this STANAG for their equipment used to provide narrow-band digital speech at 800 bps for any channel and narrow-band digital speech at 800 bps and Error correcting code and interleaver for Slow frequency hopping HF-EPM waveform.

GENERAL

5. A description of the speech coding technique is in annex A. 800 bps voice digitizer characteristics are in annex B. The speech encoding and decoding tables are in annex C. Framing at 800 bps is in annex D. For information examples of speech performances are in annex E.

6. The error detection and correction and the characteristics of the interleaver for Slow frequency hopping HF-EPM waveform are in annex F. Description of the End Of Transmission for Slow frequency hopping HF-EPM waveform is at annex G. For information capability of Late Traffic Entry for Slow frequency hopping HF-EPM waveform is presented in annex H. The functional interface with the modem for Slow frequency hopping HF-EPM waveform is in annex I. An example of interpolation strategy when a super-frame is erased is in annex J.

IMPLEMENTATION OF THE AGREEMENT

7. This STANAG is considered implemented by a nation when that nation has issued instructions that all such equipment to be procured for its forces will be manufactured in accordance with the specification detailed in this agreement.

BLANK PAGE BLANCHE

ANNEX A
DESCRIPTION OF 800 BPS SPEECH CODER

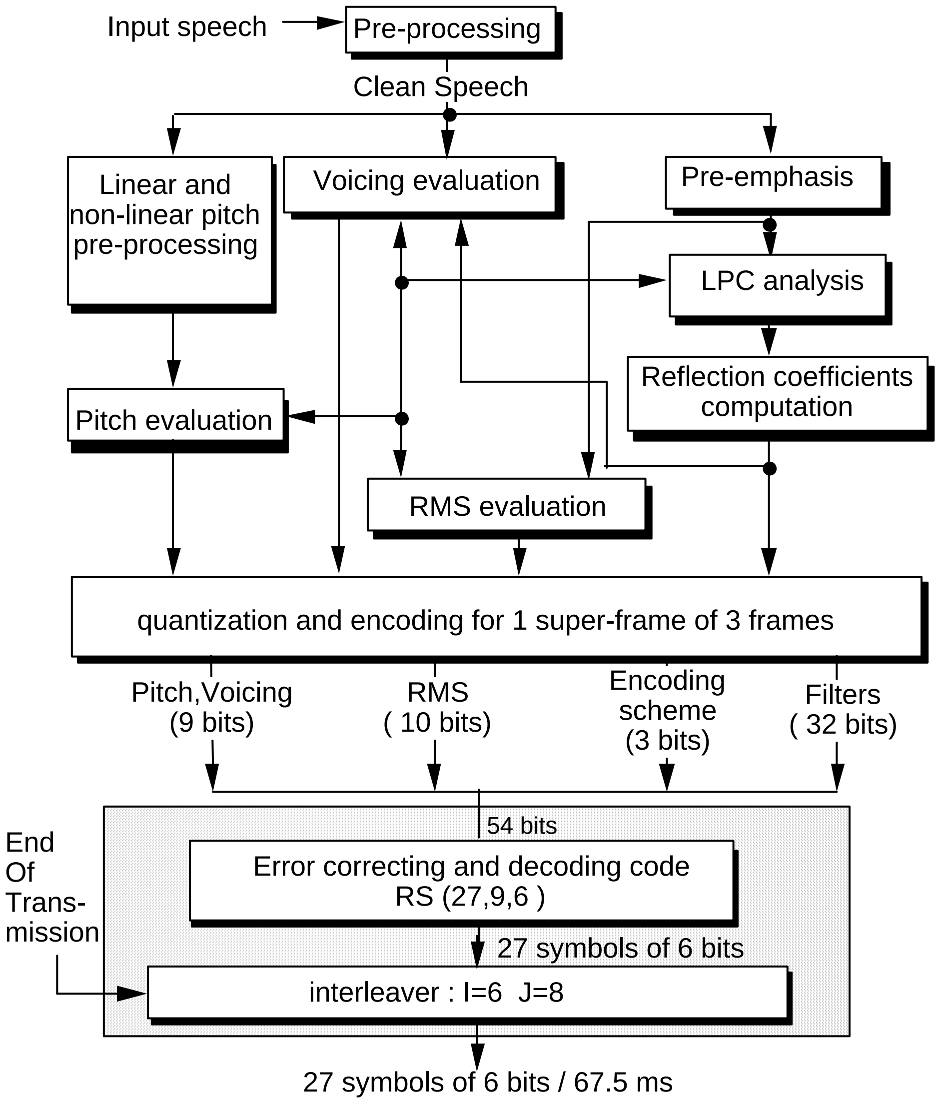


Figure 1: Analysis and Encoding

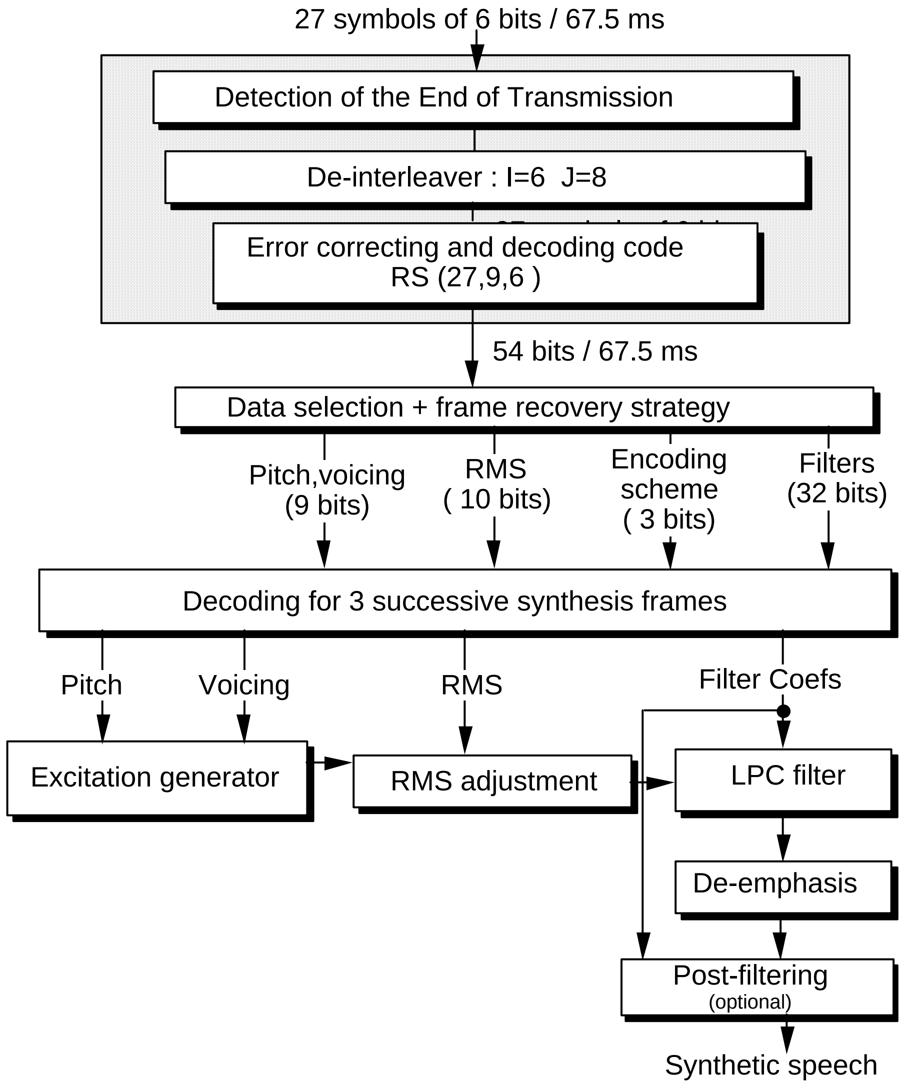


Figure 2: Decoding and Synthesis

ANNEX B
VOICE DIGITIZER CHARACTERISTICS

Sampling rate.. 8000 Hz

Prediction order... 10

Transmission data rate... 800 bps

Frame length for basic speech coder................. 22.5 ms

Super frame length, for coding will be N=3 frames of basic speech coder, i.e. 67.5 ms.

1. Speech pre-processing

1.1 The input speech is bandwidth-limited according to document **(1)**.

1.2 Bias is systematically removed before any futher processing.

1.3 Energy encoding tables are related to an input speech signal digitized on 12 bits [-2048,2047]. For any other case, it would be necessary to modify the energy coding tables and to multiply the parameters of these tables by a factor.

1.4 The result of pre-processing will be called "clean speech" hereafter.

2. Excitation analysis

2.1 Pitch extraction is done once for each analysis frame, and voicing decision is done twice a frame.

2.2 One value of RMS is computed for each analysis frame, using clean speech after removal of its mean value. The width of the energy window is a multiple of the pitch period.

3. Spectrum analysis

3.1 Spectrum analysis is done using pre-emphasized clean speech, after removal of its mean value.

3.2 The pre-emphasis filter is defined by this equation:

$$(1 - 1.2\, z^{-1} + 0.4\, z^{-2})$$

3.3 Autocorrelation or covariance methods can be used to determine 10 Reflection coefficients (convention : $K_1 = R_1/R_0$ where R_i autocorrelation coefficient, K_i Reflection coefficients and i=1,...,10). In order to avoid strong resonances, a bandwidth expansion of 20 Hz is processed on the spectrum coefficients (Prediction coefficients PC_i are multiply by α^i where $\alpha = 0.99217678$ and i=1,...,10). These coefficients are then translated into Log Area Ratio coefficients (LAR) according to :

$$LAR_i = \mathrm{Sign}(K_i) * 4096 * LOG_2\Big((1+|K_i|) / (1-|K_i|) \Big)$$

3.4 An approximation of this conversion is given in Table 1.

4. Encoding

4.1 A total of N=3 successive analysis frames are block-encoded before transmission.

5. PITCH/VOICING ENCODING

5.1 9 bits (PV_{8-0}) are used to encode the values of voicing decision and pitch of the 3 frames.

5.2 If the 3 frames are not all voiced the value of PV8-5 is between 0000 and 0011 according to the voicing configuration (see Table 2). The last 5 bits (PV_{4-0}) represent the value of the pitch (31 differents values, see Table 3-2) on the voiced frames. If there is no voiced frame the 5 last bits are all zero.

5.3 If the 3 frames are all voiced, the value of PV8-6 is between 010 and 111 according to the pitch variation on the 3 frames. There are 6 different cases (see Table 2). The last 6 bits (PV_{5-0}) represent a reference value of the pitch (largest period among the 3 frames) encoded according to Table 3-1.

6. RMS ENCODING

6.1 10 bits (RI_{9-0}) are used to encode the values of RMS of the 3 frames.

6.2 The first 5 bits (RI_{9-5}) indicate an index in a code-book of 32 different contours of energy on 3 successive frames. The 32 different cases are represented in Table 4. The encoding process chooses the contour which best correspond to the signal.

6.3 The last 5 bits (RI_{4-0}) are used to encode a reference value (code in Table 5). This reference is index dependant (see Table 4).

7. PREDICTION FILTER ENCODING

7.1 There are 8 different coding schemes according to the stability of the sound on the 3 frames. For each Super Frame the vocoder chooses the one which gives the lowest average spectral distortion between 8 schemes.

7.2 3 bits (ID_{2-0}) indicate the index of the chosen schemes (between 000 and 111)

7.3 32 bits (F_{31-0}) are used to encode the 3 prediction filters according to the value of ID. The 10 Log Area Ratio (LAR) are scalar quantized after a projection on the eigen vectors of the LARs distribution.

7.4 Table 6 describes the 8 different encoding schemes, Table 7, the decoding table according to the 8 schemes.

7.5 Table 8-1 introduces the rate according to each projection of the LAR coefficients on the eigen vectors and according to scheme 0.

7.6 Table 8-2 introduces the rate according to each projection of the LAR coefficients on the eigen vectors and according to schemes 1,2,4,5.

7.7 Table 8-3 introduces the rate according to each projection of the LAR coefficients on the eigen vectors and according to schemes 3,6,7.

7.8 Table 9 gives the coefficients of the eigen vectors of the LAR

7.9 Table10 introduces the quantized value of the vector projection on the eigen vectors of the LAR when the bit rate is 32 bits by filter. The quantizer choose the nearest quantized values of the original projections

7.10 Table 11 introduces the quantized value of the vector projection on the eigen vectors of the LAR when the bit rate is 24 bits by filter.

7.11 Table 12 introduces the quantized value of the vector projection on the eigen vectors of the LAR when the bit rate is 16 bits by filter.

8. Decoding

8.1 The decoding is the inverse operation of coding, given scaling factors

9. Synthesis

9.1 The synthesis principles are similar to those used for standard LPC10 (see document **(1)**).

9.2 The de-emphasis filter is defined by this equation :

$$\frac{(1 - 1.9998\, z^{-1} + z^{-2})}{(1 - 1.2\, z^{-1} + 0.4\, z^{-2}) * (1 - 1.75\, z^{-1} + 0.78\, z^{-2})}$$

10. Delay

The transmission delay of the vocoder is 9 frames = 202.5 ms. 6 frames are used for the analysis and the coding (with correction of the current pitch with the pitch values of the two next frames) and 3 frames are used for the decoding and the synthesis. This delay is the delay of the vocoder alone without propagation, error correcting code, and interleaver.

BLANK PAGE BLANCHE

ANNEX C

CODING AND DECODING TABLES

Reflection coefficients ==> Log Area Ratio coefficients	
0 Š $\|k_i\|$ < 0.1707	Lar_i = sign(k_i)• 11876• $\|k_i\|$
0.1707 Š $\|k_i\|$ < 0.3573	Lar_i = sign(k_i)•[2028 + 12688• ($\|k_i\|$ - 0.1707)]
0.3573 Š $\|k_i\|$ < 0.5234	Lar_i = sign(k_i)•[4396 + 14672• ($\|k_i\|$ - 0.3573)]
0.5234 Š $\|k_i\|$ < 0.6618	Lar_i = sign(k_i)•[6834 + 18272• ($\|k_i\|$ - 0.5234)]
0.6618 Š $\|k_i\|$ < 0.7703	Lar_i = sign(k_i)•[9363 + 24400• ($\|k_i\|$ - 0.6618)]
0.7703 Š $\|k_i\|$ < 0.8505	Lar_i = sign(k_i)•[12008 + 34736• ($\|k_i\|$- 0.7703)]
0.85055 Š $\|k_i\|$ < 0.9068	Lar_i = sign(k_i)•[14795 + 52496• ($\|k_i\|$ - 0.85055)]
0.9068 Š $\|k_i\|$ < 0.94424	Lar_i = sign(k_i)•[17748 + 83728• ($\|k_i\|$- 0.9068)]
0.94424 Š $\|k_i\|$ < 0.96799	Lar_i = sign(k_i)•[20883 + 140416• ($\|k_i\|$- 0.94424)]
0.96799 Š $\|k_i\|$ < 0.98236	Lar_i = sign(k_i)•[24216 + 246784• ($\|k_i\|$- 0.96799)]
0.98236 Š $\|k_i\|$ < 0.99066	Lar_i = sign(k_i)•[27764 + 453456• ($\|k_i\|$- 0.98236)]
0.99066 Š $\|k_i\|$	Lar_i = sign(k_i)•31528

Log Area Ratio coefficients ==> Reflection coefficients	
0 Š $\|Lar_i\|$ < 2028	k_i = sign(Lar_i) • 8.45 10^{-5}• $\|Lar_i\|$
2028 Š $\|Lar_i\|$ < 4396	k_i = sign(Lar_i)•[0.1707 + 7.87 10^{-5}• ($\|Lar_i\|$ - 2028)]
4396 Š $\|Lar_i\|$ < 6834	k_i = sign(Lar_i)•[0.3573 + 6.81 10^{-5}• ($\|Lar_i\|$ - 4396)]
6834 Š $\|Lar_i\|$ < 9363	k_i = sign(Lar_i)•[0.5234 + 5.47 10^{-5}• ($\|Lar_i\|$ - 6834)]
9363 Š $\|Lar_i\|$ < 12008	k_i = sign(Lar_i)•[0.6618 + 4.09 10^{-5}• ($\|Lar_i\|$ - 9363)]
12008 Š $\|Lar_i\|$ < 14795	k_i = sign(Lar_i)•[0.7703 + 2.87 10^{-5}• ($\|Lar_i\|$ - 12008)]
14795 Š $\|Lar_i\|$ < 17748	k_i = sign(Lar_i)•[0.8505 + 1.90 10^{-5}• ($\|Lar_i\|$ - 14795)]
17748 Š $\|Lar_i\|$ < 20883	k_i = sign(Lar_i)•[0.9067 + 1.19 10^{-5}• ($\|Lar_i\|$ - 17748)]
20883 Š $\|Lar_i\|$ < 24216	k_i = sign(Lar_i)•[0.9442 + 7.12 10^{-6}• ($\|Lar_i\|$ - 20883)]
24216 Š $\|Lar_i\|$ < 27764	k_i = sign(Lar_i)•[0.9679 + 4.05 10^{-6}• ($\|Lar_i\|$ - 24216)]
27764 Š $\|Lar_i\|$ < 31528	k_i = sign(Lar_i)•[0.9823 + 2.20 10^{-6}• ($\|Lar_i\|$ - 27764)]
31528 Š $\|Lar_i\|$	k_i = sign(Lar_i)• 0.9906

Table 1 : Translating table between Reflection coefficients K_i and Log Area Ratio coefficients Lar_i

frame 1 Voicing decision	frame 2 Voicing decision	frame 3 Voicing decision	frame 1 Pitch index	frame 2 Pitch index	frame 3 Pitch index	Code PV_{8-5}
0	0	0				0000
0	0	1			Index	0000
0	1	1		Index	Index	0001
1	0	0	Index			0010
1	1	0	Index	Index		0011
1	1	1	Index	Index	Index	010x
1	1	1	Index	Index - 1	Index - 1	011x
1	1	1	Index	Index - 2	Index - 3	100x
1	1	1	Index -1	Index - 1	Index	101x
1	1	1	Index-3	Index - 2	Index	110x
1	1	1	Index - 1	Index	Index - 1	111x

Index : index of the reference pitch (see Tables 3-1 and 3-2)

If one half frame is voiced, the frame is considered as all voiced

Table 2 : Encoding table of the voicing configurations for the 3 successive analysis frames and pitch period variations during the 3 frames when the frames are all voiced (4 bits) PV_{8-5}

Pitch	PV_{5-0}	Pitch	PV_{5-0}	Pitch	PV_{5-0}
20	000001	40	010101	80	101001
21	000010	42	010110	84	101010
22	000011	44	010111	88	101011
23	000100	46	011000	92	101100
24	000101	48	011001	96	101101
25	000110	50	011010	100	101110
26	000111	52	011011	104	101111
27	001000	54	011100	108	110000
28	001001	56	011101	112	110001
29	001010	58	011110	116	110010
30	001011	60	011111	120	110011
31	001100	62	100000	124	110100
32	001101	64	100001	128	110101
33	001110	66	100010	132	110110
34	001111	68	100011	136	110111
35	010000	70	100100	140	111000
36	010001	72	100101	144	111001
37	010010	74	100110	148	111010
38	010011	76	100111	152	111011
39	010100	78	101000	156	111100

Table 3-1: Encoding table of a reference value of the pitch period (6 bits) PV5-0 when the frames are all voiced

Pitch	PV 4-0	decoded pitch	Pitch	PV 4-0	decoded pitch	Pitch	PV 4-0	decoded pitch
20	00001	20	37-38	01011	37	66-68	10101	67
21	00010	21	39-40	01100	39	70-72	10110	71
22	00011	22	42	01101	42	74-76	10111	75
23-24	00100	23	44	01110	44	78-80	11000	79
25-26	00101	25	46-48	01111	47	84	11001	84
27-28	00110	27	50	10000	50	88-92	11010	90
29-30	00111	29	52-54	10001	53	96	11011	96
31-32	01000	31	56-58	10010	57	100 -104	11100	102
33-34	01001	33	60	10011	60	108	11101	108
35-36	01010	35	62-64	10100	63	112 -116	11110	114
						120 -156	11111	120

One particular case : first half-frame of the super-frame voiced AND last half-frame of the super-frame voiced AND all the other half-frame unvoiced (10/00/01)

THEN **$PV_{8\text{-}5}$ = 0000** & **$PV_{4\text{-}0}$ = 11111**

Table 3-2 : Encoding table of a reference value of the pitch period (5bits) PV4-0 when the frames are not all voiced

β_1	β_2	β_3	RI_{9-5}	ref frame	β_1	β_2	β_3	RI_{9-5}	ref frame
1.00	0.72	0.39	00000	1	1.00	0.80	0.80	10000	1
0.25	0.25	1.00	00001	3	1.00	0.56	0.56	10001	1
0.42	0.99	1.00	00010	3	1.00	0.83	0.63	10010	1
0.09	1.00	0.20	00011	2	0.04	0.06	1.00	10011	3
0.70	1.00	0.95	00100	2	1.00	0.98	0.50	10100	1
1.00	1.00	1.00	00101	2	0.47	0.60	1.00	10101	3
1.00	0.05	0.02	00110	1	0.58	0.24	1.00	10110	3
1.00	0.18	0.92	00111	1	0.83	1.00	0.82	10111	2
0.20	0.72	1.00	01000	3	1.00	0.16	0.13	11000	1
0.93	0.95	1.00	01001	3	1.00	0.60	0.92	11001	1
1.00	0.75	0.20	01010	1	0.68	0.78	1.00	11010	3
0.15	1.00	0.90	01011	2	1.00	0.29	0.21	11011	1
0.07	0.80	1.00	01100	3	1.00	0.48	0.20	11100	1
1.00	0.56	0.05	01101	1	1.00	0.94	0.80	11101	1
1.00	0.30	0.05	01110	1	1.00	0.40	0.40	11110	1
1.00	1.00	0.12	01111	2	0.59	1.00	0.54	11111	2

$$\beta_i = \frac{\text{RMS frame } i}{\text{Max (RMS frames } 1,2,3)}$$

Table 4 : Encoding table of the RMS contour for 3 successive frames RI_{9-5}

RMS	RI_{4-0}	RMS	RI_{4-0}	RMS	RI_{4-0}	RMS	RI_{4-0}
0	00000	8	01000	32	10000	135	11000
1	00001	9	01001	39	10001	164	11001
2	00010	11	01010	46	10010	192	11010
3	00011	13	01011	55	10011	230	11011
4	00100	16	01100	66	10100	275	11100
5	00101	19	01101	79	10101	328	11101
6	00110	23	01110	94	10110	392	11110
7	00111	27	01111	113	10111	468	11111

Table 5 : Encoding table of the RMS reference RI4-0 when the input signal is quantized between -2048 and 2047

ID_{2-0}		Total
000	- A = Average Spectrum Frame 1,2,3 ==> cA=Code (A,32)	32
001	- A = Average Spectrum Frame 1,2 ==> cA =Code (A,24) - Difference between Code (A,24) and Code (frame3,24) on 8 bits ==> cDiff (8 bits)	24+8=32
010	- A = Average Spectrum Frame 2,3 ==> cA = Code (A,24) - Difference between Code (A,24) and Code (frame1,24) on 8 bits ==> cDiff (8 bits)	24+8=32
011	- cFr1 = Code (frame 1,16) - cFr3 = Code (frame 3,16)	16+16=32
100	- cFr1 = Code (frame 1,24) - Difference between Code (frame 1,24) and Code (frame3,24) on 8 bits ==> cDiff (8 bits)	24+8=32
101	- cFr3 = Code (frame 3,24) - Difference between Code (frame 1,24) and Code (frame 3,24) on 8 bits ==> cDiff (8 bits)	24+8=32
110	- A = Average Spectrum Frame 1,2 ==> cA = Code (A,16) - cFr3 = Code (frame 3,16)	16+16=32
111	- cFr1 = Code (frame 1,16) - A = Average Spectrum Frame 2,3 ==> cA= Code (A,16)	16+16=32

Code (X,y) = Encoding of the spectrum X on y bits

Table 6 : Encoding table according to the 8 different spectrum encoding schemes of 3 successive frames ID_{2-0}

ID_{2-0}	Decoded frame 1	Decoded frame 2	Decoded frame 3
000	Decode (cA,32)	Decode (cA,32)	Decode (cA,32)
001	Decode (cA,24)	Decode (cA,24)	Decode (cA+cDiff, 24)
010	Decode (cA+cDiff, 24)	Decode (A,24)	Decode (A,24)
011	Decode (cFr1,16)	$\frac{(Fr1 + Fr3)}{2}$	Decode (cFr3,16)
100	Decode (cFr1,24)	$\frac{(Fr1 + Fr3)}{2}$	Decode (cFr1+cDiff, 24)
101	Decode (cFr3+cDiff, 24)	$\frac{(Fr1 + Fr3)}{2}$	Decode (cFr3,24)
110	Decode (cA,16)	Decode (cA,16)	Decode (cFr3,16)
111	Decode (cFr1,16)	Decode (cA,16)	Decode (cA,16)

Decode (cX,y) = Decoding with the code cX on y bits

Table 7 : Decoding table according to the 8 different spectrum encoding schemes of 3 successive frames ID_{2-0}

Projection on eigen vectors

number of bit for the Quantization

1	4	F_{3-0}
2	4	F_{7-4}
3	4	F_{11-8}
4	4	F_{15-12}
5	3	F_{18-16}
6	3	F_{21-19}
7	3	F_{24-22}
8	3	F_{27-25}
9	2	F_{29-28}
10	2	F_{31-30}

Total number = 32 Bits

Table 8-1 : Encoding table of the projection of the LAR coefficients of a Spectrum on the eigen vectors according to scheme 0

Projection on eigen vectors

number of bit for the Quantization

1	4	F_{3-0}
2	4	F_{7-4}
3	3	F_{10-8}
4	3	F_{13-11}
5	3	F_{16-14}
6	3	F_{19-17}
7	2	F_{21-20}
8	2	F_{23-22}
9	0	
10	0	

Frame quantized on 24 bits

number = 24 Bits

1	2 bit = -2 to 1	F_{25-24}
2	2 bit = -2 to 1	F_{27-26}
3	1 bit = -1 to 0	F_{28}
4	1 bit = -1 to 0	F_{29}
5	1 bit = -1 to 0	F_{30}
6	1 bit = -1 to 0	F_{31}
7	0 bit	
8	0 bit	
9	0 bit	
	0 bit	

8 bits are used for encoding the **difference** between the **code** of 24 bits quantization of two frames

Difference on 8 bits of two 24 bits spectrum encoding

Total number= 24 + 8 = 32 bits

Table 8-2 : Encoding tables of the projection of the LAR coefficients of a Spectrum on the eigen vectors according to schemes 1,2,4,5

Projection on eigen vectors

number of bit for the Quantization

		first fr	second fr
1	**4**	F_{3-0}	F_{19-16}
2	**3**	F_{6-4}	F_{22-20}
3	**3**	F_{9-7}	F_{25-23}
4	**2**	F_{11-10}	F_{27-26}
5	**2**	F_{13-12}	F_{29-28}
6	**2**	F_{15-14}	F_{31-30}
7	**0**		
8	**0**		
9	**0**		
10	**0**		

Total number = 16 Bits • 2 = 32 bits

Table 8-3 : Encoding table of the projection of the LAR coefficients of a Spectrum on the eigen vectors according to schemes 3,6,7

V(1,J)	V(2,J)	V(3,J)	V(4,J)	V(5,J)	V(6,J)	V(7,J)	V(8,J)	V(9,J)	V(10,J)
7713	-1352	1980	611	547	-757	658	292	318	-12
1939	4553	-5116	3085	-2219	63	425	1320	137	244
90	5595	2164	-2254	3882	566	2673	-272	1826	-307
-1086	2065	4337	2977	-917	-2827	-3193	3729	471	916
-1227	-961	1608	1925	-3543	-1804	5475	-892	3217	-1210
-662	-2184	-262	2634	2565	4433	2024	4245	1478	2088
-320	-1654	-2902	804	3865	-3865	-1560	45	4531	-1538
257	-546	-1452	-5096	-1628	-2579	1358	4077	517	3260
649	249	385	-2452	-2722	3387	-2516	1879	3705	-4241
363	442	436	22	-1359	1501	-2115	-3512	4046	5414

$$\mathrm{Proj}(j) = \left(\sum_{j=l}^{10} \mathrm{Lar}\ j \ .\ \mathrm{Vect}\ (i,j) \right) / 32768$$

Note : This table has been normalized on 8192. So to decode, the following formula has to be used :

$$\mathrm{Lar}\ i = \left(\sum_{j=l}^{10} \mathrm{Proj}(j) \ .\ \mathrm{Vect}\ (j,i) \right) / 2048$$

Table 9 : Coefficients of the eigen vectors of the LAR on which LAR coefficients are projected

Code	32 b frame								
		0000 1000	0001 1001	0010 1010	0011 1011	0100 1100	0101 1101	0110 1110	0111 1111
Proj 1	F_{3-0}	-5430 492	-4590 1180	-3791 1878	-3027 2591	-2291 3327	-1578 4091	-880 4890	-192 5730
Proj 2	F_{7-4}	-4050 -17	-3478 452	-2934 927	-2413 1412	-1912 1913	-1427 2434	-952 2978	-483 3550
Proj 3	F_{11-8}	-2530 951	-2036 1356	-1566 1766	-1117 2185	-685 2617	-266 3066	144 3536	549 4030
Proj 4	F_{15-12}	-2920 -86	-2518 243	-2136 577	-1770 918	-1418 1270	-1077 1636	-743 2018	-414 2420
Proj 5	F_{18-16}	-1790	-1119	-507	68	632	1207	1819	2490
Proj 6	F_{21-19}	-2025	-1421	-871	-353	153	671	1221	1825
Proj 7	F_{24-22}	-1630	-1040	-503	2	498	1003	1540	2130
Proj 8	F_{27-25}	-2290	-1830	-1412	-1018	-632	-238	180	640
Proj 9	F_{29-28}	-1000	-465	15	550				
Proj 10	F_{31-30}	-1500	-1017	-583	-100				

Note : This table has been calculated with a table of eigen vectors normalized on 8192 (see Table 9).

Table 10 : Quantized values of the vector projection on the eigen vectors of the LAR when the bit rate is 32 bits by filter

Code	24 b frame	0000	0001	0010	0011	0100	0101	0110	0111
		1000	1001	1010	1011	1100	1101	1110	1111
Proj 1	F_{3-0}	-5430	-4590	-3791	-3027	-2291	-1578	-880	-192
		492	1180	1878	2591	3327	4091	4890	5730
Proj 2	F_{7-4}	-4050	-3478	-2934	-2413	-1912	-1427	-952	-483
		-17	452	927	1412	1913	2434	2978	3550
Proj 3	F_{10-8}	-2310	-1350	-476	347	1153	1976	2850	3810
Proj 4	F_{13-11}	-2740	-1959	-1247	-578	78	747	1459	2240
Proj 5	F_{16-14}	-1790	-1119	-507	68	632	1207	1819	2490
Proj 6	F_{19-17}	-2025	-1421	-871	-353	153	671	1221	1825
Proj 7	F_{21-20}	-1060	-169	629	1520				
Proj 8	F_{23-22}	-1880	-1152	-498	230				
Proj 9		-232							
Proj 10		-720							

Note : This table has been calculated with a table of eigen vectors normalized on 8192 (see Table 9).

Table 11 : Quantized values of the vector projection on the eigen vectors of the LAR when the bit rate is 24 bits by filter.

Code	first 16 b frame	second 16 b frame	0000	0001	0010	0011	0100	0101	0110	0111
			1000	1001	1010	1011	1100	1101	1110	1111
Proj 1	F_{3-0}	F_{19-16}	-5430	-4590	-3791	-3027	-2291	-1578	-880	-192
			492	1180	1878	2591	3327	4091	4890	5730
Proj 2	F_{6-4}	F_{22-20}	-3790	-2679	-1668	-716	216	1168	2179	3290
Proj 3	F_{9-7}	F_{25-23}	-2310	-1350	-476	347	1153	1976	2850	3810
Proj 4	F_{11-10}	F_{27-26}	-2180	-841	361	1700				
Proj 5	F_{13-12}	F_{29-28}	-1280	-155	855	1980				
Proj 6	F_{15-14}	F_{31-30}	-1500	-533	333	1300				
Proj 7			208							
Proj 8			-818							
Proj 9			-232							
Proj 10			-720							

Note : This table has been calculated with a table of eigen vectors normalized on 8192 (see Table 9).

Table 12 : Quantized values of the vector projection on the eigen vectors of the LAR when the bit rate is 16 bits by filter

BLANK PAGE BLANCHE

ANNEX D

FRAMING AT 800 BPS

1 - Ri_0	19 - PV_8	37 - F_{14}
2 - Ri_1	20 - ID_0	38 - F_{15}
3 - Ri_2	21 - ID_1	39 - F_{16}
4 - Ri_3	22 - ID_2	40 - F_{17}
5 - Ri_4	23 - F_0	41 - F_{18}
6 - Ri_5	24 - F_1	42 - F_{19}
7 - Ri_6	25 - F_2	43 - F_{20}
8 - Ri_7	26 - F_3	44 - F_{21}
9 - Ri_8	27 - F_4	45 - F_{22}
10 - Ri_9	28 - F_5	46 - F_{23}
11 - PV_0	29 - F_6	47 - F_{24}
12 - PV_1	30 - F_7	48 - F_{25}
13 - PV_2	31 - F_8	49 - F_{26}
14 - PV_3	32 - F_9	50 - F_{27}
15 - PV_4	33 - F_{10}	51 - F_{28}
16 - PV_5	34 - F_{11}	53 - F_{30}
17 - PV_6	35 - F_{12}	52 - F_{29}
18 - PV_7	36 - F_{13}	54 - F_{31}

Framing Table at 800 bps
54 bits

BLANK PAGE BLANCHE

ANNEX E

EXAMPLE OF SPEECH PERFORMANCES
(for information only)

1. The performance of 800 bps speech processor shall be measured in terms of intelligibility test.

2. The voice intelligibility of the voice processor shall be measured using the Diagnostic Rhyme test (DRT). For the DRT, English/American and French versions are to be used, and the talkers and listeners are to be familiar with the language in each case. The input analog tapes to be used for the English/American DRT and the minimum acceptable scores, which should be obtained from an independant contractor are given below (Standard deviation on a DRT of about 1.0) :

Acoustic Environment	Talkers	Microphone	Score
Quiet	3M	Dynamic	86
Quiet filtered	3M	Dynamic	82
Jeep	3M	H250	78
Tank	3M	EV985	74

3. The Spectrum distortion measure used for theses tests, is the ITAKURA-SAITO distance Measure (see Annex B : Choice of the spectrum encoding scheme).

BLANK PAGE BLANCHE

ANNEX F

ERROR DETECTION & CORRECTION AND INTERLEAVER FOR SLOW FREQUENCY HOPPING HF-EPM

1. Error Detection & Correction code :

1.1 Each 54 bits Superframe (SF) is encoded using a Reed-Solomon (RS) code using 6 bits-symbols.

1.2 In order to achieve a 2400 bps interface with the modem, only one **RS(27,9)** code is used . All the 54 bits of the SF have the same protection and there is only one RS code for each SF.

1.3 The 9 original symbols of 6 bits of the RS code are constitued according to the framing at 800 bps (see Annex D)

Symbol 1	C_1 :	6 (MSB) to 1 (LSB)	==>	RI_5	RI_0
Symbol 2	C_2 :	12 (MSB) to 7 (LSB)	==>	PV_1	RI_6
Symbol 3	C_3 :	18 (MSB) to 13 (LSB)	==>	PV_7	PV_2
Symbol 4	C_4 :	24 (MSB) to 19 (LSB)	==>	F_1	PV_8
Symbol 5	C_5 :	30 (MSB) to 25 (LSB)	==>	F_7	F_2
Symbol 6	C_6 :	36 (MSB) to 31 (LSB)	==>	F_{13}	F_8
Symbol 7	C_7 :	42 (MSB) to 37 (LSB)	==>	F_{19}	F_{14}
Symbol 8	C_8 :	48 (MSB) to 43 (LSB)	==>	F_{25} -	F_{20}
Symbol 9	C_9 :	54 (MSB) to 49 (LSB)	==>	F_{31}	F_{26}

1.4 The 27 symbols are framed according to the polynome C(x) (C_i : symbol number i of the RS word) :

$$C(x) = \sum_{i=1}^{27} C_i . x^{27-i}$$

with $C(\alpha) = C(\alpha^2) = C(\alpha^3) = \dots = C(\alpha^{18}) = 0$
and α root of $G(x) = x^6 + x + 1$ in the Galois field GF(64)

1.5 The redundancy symbols are C_{10} to C_{27}. C_1 is sent first, (LSB first) the others symbols following in increasing index order.

1.6 The elements of GF(64) are defined as follows :

	α^5 (MSB)	α^4	α^3	α^2	α	**1** (LSB)
0	0	0	0	0	0	0
1	0	0	0	0	0	1
α	0	0	0	0	1	0
α^2	0	0	0	1	0	0
α^3	0	0	1	0	0	0
α^4	0	1	0	0	0	0
α^5	1	0	0	0	0	0
α^6	0	0	0	0	1	1
α^7	0	0	0	1	1	0
α^8	0	0	1	1	0	0
α^9	0	1	1	0	0	0
α^{10}	1	1	0	0	0	0
α^{11}	1	0	0	0	1	1
α^{12}	0	0	0	1	0	1
α^{13}	0	0	1	0	1	0
α^{14}	0	1	0	1	0	0
α^{15}	1	0	1	0	0	0
α^{16}	0	1	0	0	1	1
α^{17}	1	0	0	1	1	0
α^{18}	0	0	1	1	1	1
α^{19}	0	1	1	1	1	0
α^{20}	1	1	1	1	0	0
α^{21}	1	1	1	0	1	1
α^{22}	1	1	0	1	0	1
α^{23}	1	0	1	0	0	1
α^{24}	0	1	0	0	0	1
α^{25}	1	0	0	0	1	0
α^{26}	0	0	0	1	1	1
α^{27}	0	0	1	1	1	0
α^{28}	0	1	1	1	0	0
α^{29}	1	1	1	0	0	0
α^{30}	1	1	0	0	1	1

	α^5 (MSB)	α^4	α^3	α^2	α	**1** (LSB)
α^{31}	1	0	0	1	0	1
α^{32}	0	0	1	0	0	1
α^{33}	0	1	0	0	1	0
α^{34}	1	0	0	1	0	0
α^{35}	0	0	1	0	1	1
α^{36}	0	1	0	1	1	0
α^{37}	1	0	1	1	0	0
α^{38}	0	1	1	0	1	1
α^{39}	1	1	0	1	1	0
α^{40}	1	0	1	1	1	1
α^{41}	0	1	1	1	0	1
α^{42}	1	1	1	0	1	0
α^{43}	1	1	0	1	1	1
α^{44}	1	0	1	1	0	1
α^{45}	0	1	1	0	0	1
α^{46}	1	1	0	0	1	0
α^{47}	1	0	0	1	1	1
α^{48}	0	0	1	1	0	1
α^{49}	0	1	1	0	1	0
α^{50}	1	1	0	1	0	0
α^{51}	1	0	1	0	1	1
α^{52}	0	1	0	1	0	1
α^{53}	1	0	1	0	1	0
α^{54}	0	1	0	1	1	1
α^{55}	1	0	1	1	1	0
α^{56}	0	1	1	1	1	1
α^{57}	1	1	1	1	1	0
α^{58}	1	1	1	1	1	1
α^{59}	1	1	1	1	0	1
α^{60}	1	1	1	0	0	1
α^{61}	1	1	0	0	0	1
α^{62}	1	0	0	0	0	1

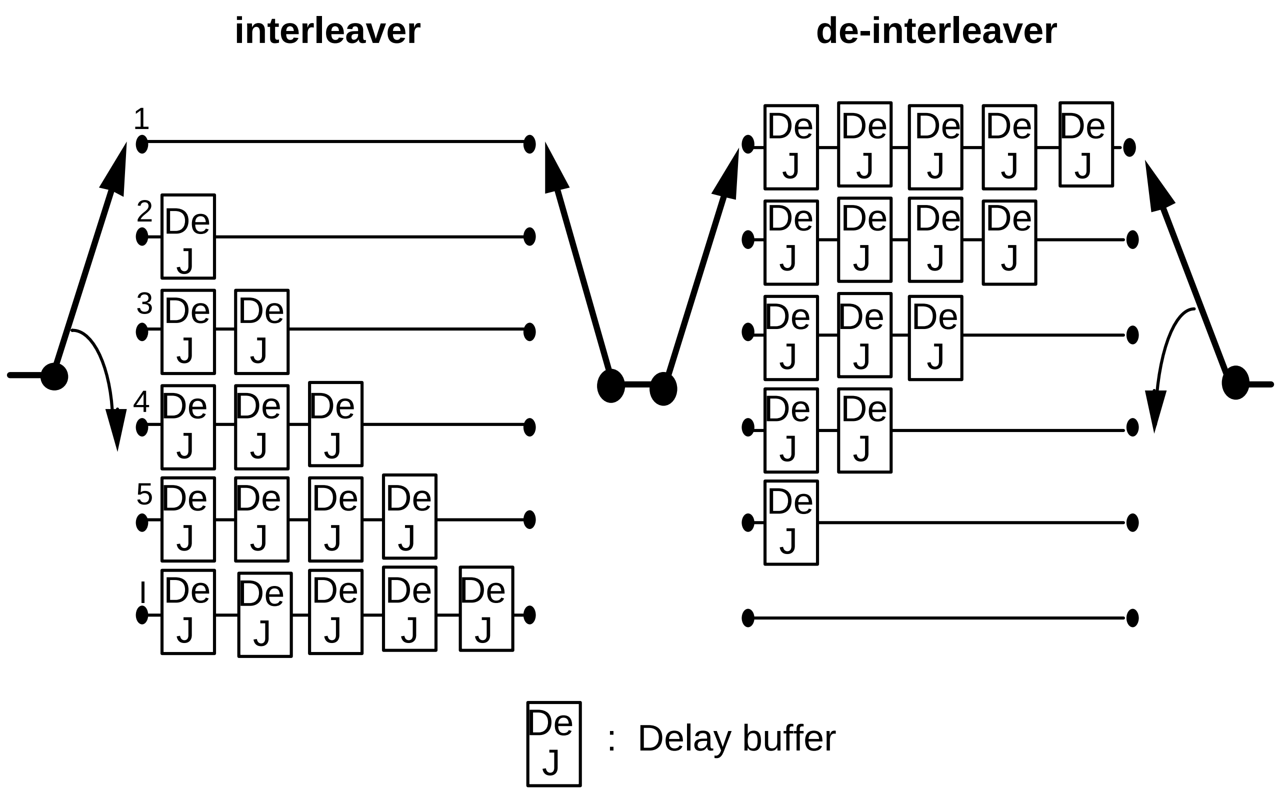
1.7 Addition is performed using bitwise exclusive OR. Multiplication is performed using usual definition of exponential with $\alpha^{63} = \alpha^{0}$ = {000001} = '1'

Example : '23'•'29' = {010111} •{011101} = $\alpha^{54} \cdot \alpha^{41}$
= $\alpha^{(54+41)}$ = α^{32} = {001001}

2. Interleaver :

2.1 In order to minimize the TAT (Turn Around Time) a convolutional interleaver has been choosen. The interleaver is fully described by the number of rows I and by the number of incremental buffer storage per row J. The interleaver works with 6 bits symbols and the clocking of the commutator swiches is done by the Reed-Solomon rate.

I = 6 and J = 8



2.2 The total delay of the interleaver-deinterleaver is equal to **I*(I-1)*J**

==> 6*5*8 = 240 symbols

2.3 There are two cases : the first symbol C_1 of a RS code (= one Superframe) can appear when the commutators are connected to the first or fourth row of the interleaver. At the beginning of the transmission, the first row is used.

ANNEX G

END OF TRANSMISSION FOR SLOW FREQUENCY HOPPING HF-EPM

1. The End Of Transmision (EOT) is a sequence which is led into the transmitter interleaver immediately after the end of the last generated Superframe. According to the case (even or odd Superframe) this sequence will begin on the first rail or on the fourth row . The End Of Transmission sequence at the input of the interleaver will last 240 symboles (= delay of the interleaver in order to be sure all the symbols of the last generated superframe have been transmitted). For the interleaver described above (see Annex F) and with **E** representing a EOT symbol and **I** representing a RS codeword symbol, the channel output of the interleaver should be as follows (Odd Superframe) :

EIIIII	EIIIII	EIIIII	EIIIII	EIIIII	EIIIII	EIIIII	EIIIII
EEIIII	EEIIII	EEIIII	EEIIII	EEIIII	EEIIII	EEIIII	EEIIII
EEEIII	EEEIII	EEEIII	EEEIII	EEEIII	EEEIII	EEEIII	EEEIII
EEEEII	EEEEII	EEEEII	EEEEII	EEEEII	EEEEII	EEEEII	EEEEII
EEEEEI	EEEEEI	EEEEEI	EEEEEI	EEEEEI	EEEEEI	EEEEEI	EEEEEI

2. This yields 8 + (2*8) + (3*8) + (4*8) + (5*8) = 120 symbols for EOT pattern use on the channel side. Correlation search for the EOT can be done on the channel side of the interleaver. Thus saving time because of no waiting for the EOT to appear at the output of the deinterleaver.

3. The sequence of EOT is:

EOT(1) =1

and EOT(i) = 5*EOT(i-1) +1 [64] i= 2 to 240

4. Example of detection of EOT (for information only)

4.1 The detection of the EOT sequence is done by a correlation function between the EOT symbols and the 6 bits symbols provided by the demodulator :

Score = 1 if the two symbols are equals, else score is 0
correlation = Sum on the 120 EOT symbol transmitted

4.2 The correlation function gives a total score between 0 and 120. If this total is superior to a threshold, there is detection of EOT.

4.3 A value for the threshold of 10 gives a good detection rate even with high symbol error rate and with practically no false detection. Experiments have shown that it is better to compute by comparaison of symbols rather than bits.

Warning : As the detection of the EOT sequence is done before de-interleaving, the correlation have to be processed on the "interleaved" symbols.

ANNEX H

LATE TRAFFIC ENTRY CAPABILITY
FOR SLOW FREQUENCY HOPPING HF-EPM

1. Initial Synchronisation:

1.1 The initial sync of the voice coder system is provided by the slow frequency hopping HF-EPM Modem.

2. Late traffic entry:

2.1 Since the parameter **I** of the interleaver (see Annex F) divides the number **H** of 6 bits symbol available on one hop, the interleaver (or the deinterleaver) is synchronized with the hop (the switches are in the same position at the beginning of each hop).

Let's define:

$$\mathbf{n} = \mathrm{MMCQ}(\mathbf{H},27)/27$$
$$\mathbf{m} = \mathrm{MMCQ}(\mathbf{H},27)/\mathbf{H}$$

2.2 Thus there are only **m** positions of synchronisation of the RS codeword at the output of the deinterleaver, or every **n** hops, the RS codeword is synchronized, relatively to the interleaver, at the same position.

2.3 The RS decoding algorithm can be used to detect synchronisation so there is no need to insert a special Late Traffic Entry sequence.

BLANK PAGE BLANCHE

ANNEX I

FUNCTIONAL INTERFACE WITH THE MODEM FOR SLOW FREQUENCY HOPPING HF-EPM

Technical Interface Document Between the TSGCE SG/11-WG/5 Ad Hoc Working Party (AHWP) for the HF EPM STANAG and SG11/WG2 for the HF EPM Low Rate Voice Processor STANAG.

1. BACKGROUND

The 8 February 1988 ATCA/ANCA Military Operational Requirements (MOR) document, which established the need for a High Frequency (HF) Electronics Protective Measures (EPM) NATO Standardization Agreement (STANAG), states that the STANAG shall support voice communications. WG/2 has been given the responsibility for developing and standardizing a digital voice coder (vocoder) system that will be used with the HF EPM STANAG system. Because of apparently overlapping responsibilities, the AHWP and WG/2 representatives met several times to discuss the integration of the vocoder with the system and to determine what characteristics of the functional interface would be developed by each working group. The following paragraphs contain the agreements made in those meetings.

2. AGREEMENT

2.1 Interface

The HF EPM system will provide a transparent channel between the vocoder and the HF channel ; that is, the vocoder transmitter will provide a digital data stream as an input to the HF EPM system at a rate of 2400 bits per second. The HF EPM receiver will return data to the voice processor at a rate of 2400 bps. Each modulation symbol will have a quality information to permit erasure and soft decoding of the Reed Solomon Codes. A functional interface and system block diagram to manage the data transfer between the vocoder system and the HF EPM system have been defined and are attached.

2.2 Communications Security (COMSEC)

COMSEC for the vocoder data will be provided by the HF EPM system. Therefore, the digital stream that is input by the vocoder into the HF EPM system will be COMSEC covered in a non-error extending mode by the HF EPM system as it is provided to that system. The HF EPM receiver will remove the COMSEC cover prior to providing the digital stream to the vocoder system.

2.3 Coding and Interleaving

Error correction coding and interleaving of the digital voice data will be provided by the vocoder. The data stream will be coded and interleaved prior to being provided to the HF EPM system. At the receiver, the voice data and quality information is sent to the vocoder where the deinterleaving and decoding is done.

2.4 Buffering and Service Windows (SWs)

Buffering of voice data, to allow for monitoring of the SW, will take place in the HF EPM system.

2.5 Turn Around Time (TAT)

In order to minimize TAT, the vocoder system will optimize the processing/procedures between CTP (Call To Process) and EOT (End OF Transmission), and the HF EPM system will optimize the processing/procedures involving SW (Service Window) buffer management and LSU (Link Set Up).

2.6 End of Transmission (EOT)

At the receiver, communication is normally terminated by the HF EPM system, when the vocoder has detected an EOT in the voice data stream, and signals this through the interface. In addition, communication is terminated whenever the HF EPM system fails to detect an IT (Initial Training) sequence for a predetermined number of consecutive hops.

2.7 Initial Synchronization

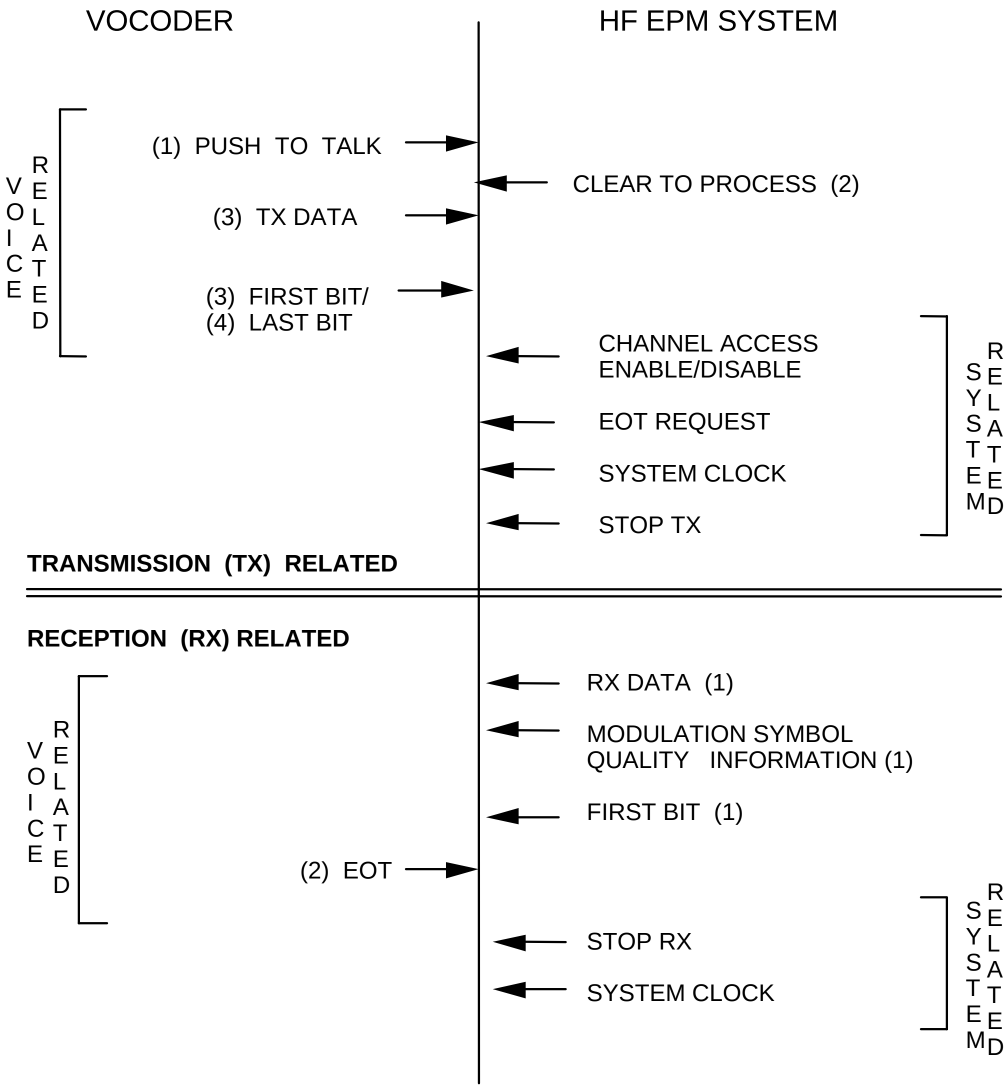
The HF EPM system will provide the initial sync indication for the vocoder as shown in the attached functional interface diagram.

28. Late Traffic Entry

No stated requirement, at this time.

3. STANAG References

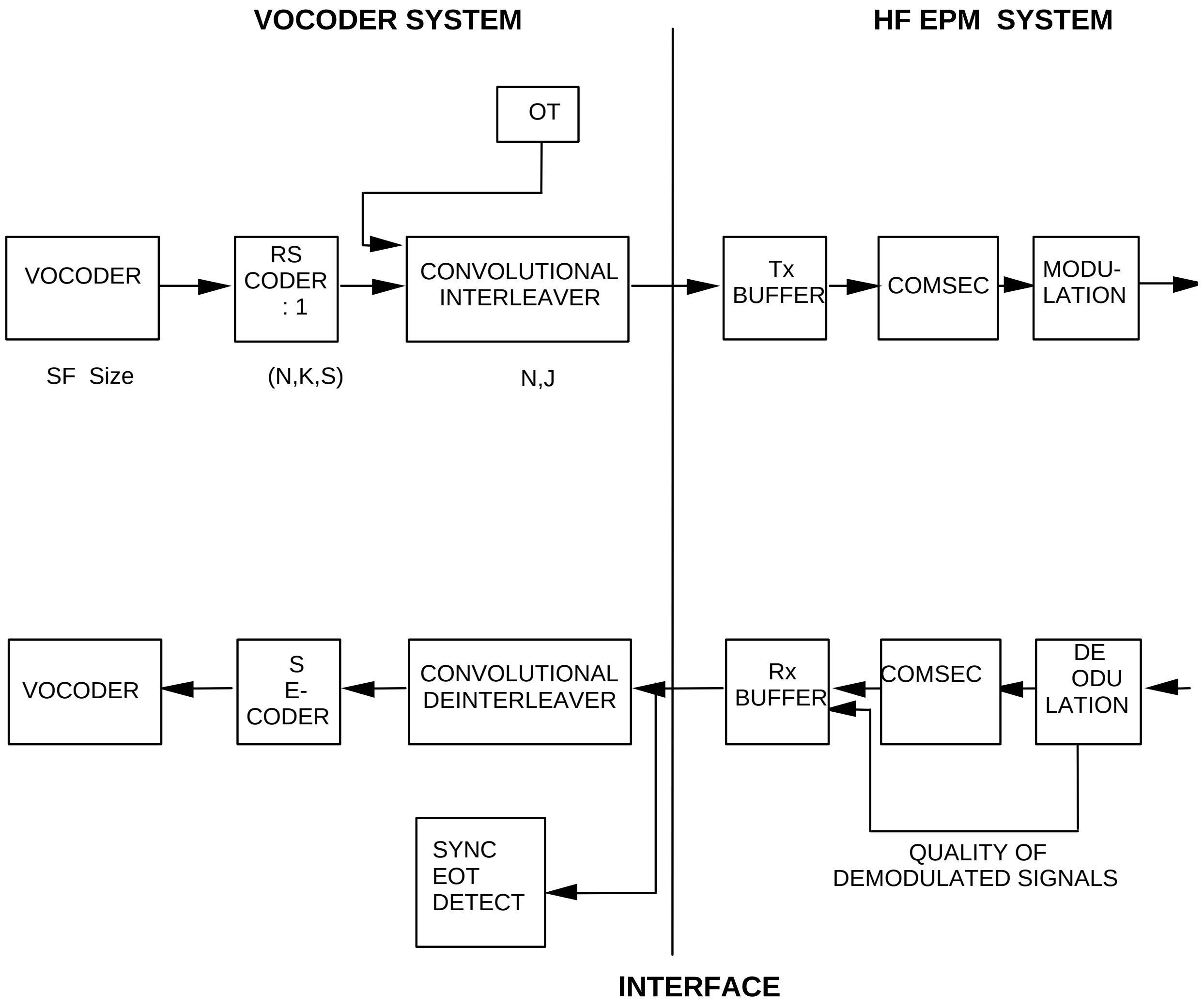
The description of the vocoder interleaving, initial sync, end of transmission, and coding schemes implemented in the vocoder will be included in the vocoder STANAG, which will be referenced in the HF EPM STANAG. Description of the vocoder interleaving, initial sync, end of transmission, and coding schemes will not be repeated in the HF EPM STANAG.



FUNCTIONAL INTERFACE

VOCODER SYSTEM AND HF EPM SYSTEM

FUNCTIONAL BLOCK DIAGRAM OF VOCODER SYSTEM AND HF EPM SYSTEM



BLANK PAGE BLANCHE

ANNEX J
EXAMPLE OF A SPEECH SMOOTHING PROCESS
(for information only)

1. In order to increase the quality of the vocoded speech when there are many transmission errors, one has to take into account of the RS word erasure information given by the decoder. As every RS word is connected to one and only one speech Super-frame, the decoder can say if a given Super-frame is erased or not.

2. It can be replaced by a copy of the last valid Super-frame. However a more sophisticated solution is proposed.

3. **IF super-frame(N) is erased THEN :**

Spectrum (1,N) = Spectrum (3,N-1)
Spectrum (2,N) = Spectrum (3,N-1)
Spectrum (3,N) = Spectrum (3,N-1)

Power (1,N) = Power (3,N-1) * Gamma
Power (2,N) = Power (3,N-1) * $(\text{Gamma})^2$
Power (3,N) = Power (3,N-1) * $(\text{Gamma})^3$

Pitch (1,N) = Pitch (3,N-1)
Pitch (2,N) = Pitch (3,N-1)
Pitch (3,N) = Pitch (3,N-1)

Voicing decision (1,N) = Voicing decision (3,N-1)
Voicing decision (2,N) = Voicing decision (3,N-1)
Voicing decision (3,N) = Voicing decision (3,N-1)

WITH : $0 < \text{Gamma} < 1$
Funct (**i,j**) : Function of frame **i** of super-frame **j**

4. It is also possible to use an interpolation strategy at the cost of a longer delay.

BLANK PAGE BLANCHE