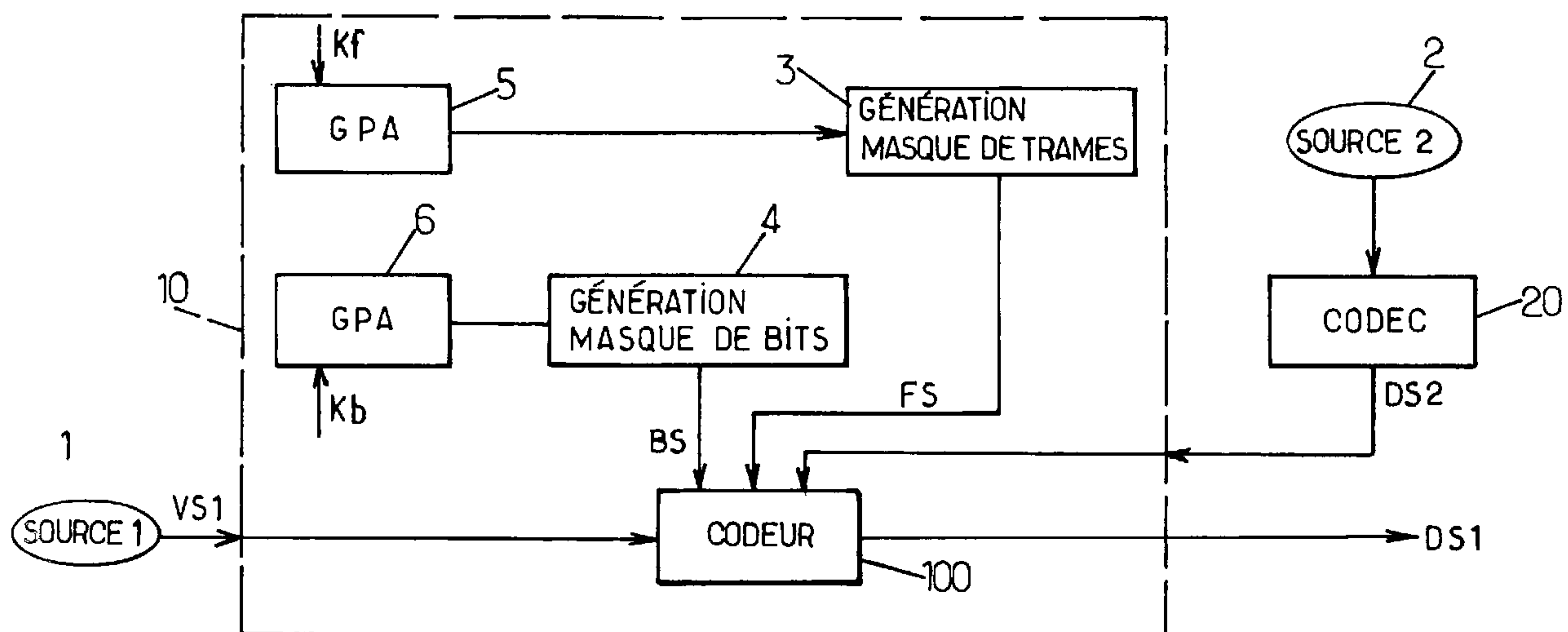




(86) Date de dépôt PCT/PCT Filing Date: 2004/09/06
(87) Date publication PCT/PCT Publication Date: 2005/03/17
(85) Entrée phase nationale/National Entry: 2006/03/02
(86) N° demande PCT/PCT Application No.: FR 2004/002259
(87) N° publication PCT/PCT Publication No.: 2005/024786
(30) Priorité/Priority: 2003/09/05 (FR03/10546)

(51) Cl.Int./Int.Cl. *G10L 19/14* (2006.01),
H04L 9/12 (2006.01)
(71) Demandeur/Applicant:
EADS SECURE NETWORKS, FR
(72) Inventeur/Inventor:
ROUSSEAU, FREDERIC, FR
(74) Agent: OGILVY RENAULT LLP/S.E.N.C.R.L.,S.R.L.

(54) Titre : PROCEDE DE TRANSMISSION D'UN FLUX D'INFORMATION PAR INSERTION A L'INTERIEUR D'UN FLUX DE DONNEES DE PAROLE, ET CODEC PARAMETRIQUE POUR SA MISE EN OEUVRE
(54) Title: INFORMATION FLOW TRANSMISSION METHOD WHEREBY SAID FLOW IS INSERTED INTO A SPEECH DATA FLOW, AND PARAMETRIC CODEC USED TO IMPLEMENT SAME



(57) **Abrégé/Abstract:**

Un procédé de transmission d'un flux d'information secondaire (DS2) entre un émetteur et un récepteur, comprenant l'insertion dudit flux d'information secondaire au niveau d'un vocodeur paramétrique (10) de l'émetteur générant un flux d'information principal (DS1) qui est un flux de données de parole codant un signal de parole (VS1) et qui est transmis de l'émetteur vers le récepteur, prévoit que des bits du flux d'information secondaire sont insérés dans certaines seulement des trames du flux d'information principal, sélectionnées par un masque de trames (FS) connu de l'émetteur et du récepteur, et/ou à l'intérieur d'une trame déterminée du flux d'information principal, en imposant une contrainte à certains seulement des bits de la trame, sélectionnés par un masque de bits (BS) connu de l'émetteur et du récepteur.



(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international(43) Date de la publication internationale
17 mars 2005 (17.03.2005)

PCT

(10) Numéro de publication internationale
WO 2005/024786 A1(51) Classification internationale des brevets⁷ :
G10L 19/14, H04L 9/12

Jean-Pierre Timbaud, F-78180 Montigny le Bretonneux (FR).

(21) Numéro de la demande internationale :
PCT/FR2004/002259

(72) Inventeur; et

(75) Inventeur/Déposant (pour US seulement) : **ROUSSEAU, Frédéric** [FR/FR]; 22, rue Victor Hugo, F-78180 Montigny-le-Bretonneux (FR).(22) Date de dépôt international :
6 septembre 2004 (06.09.2004)(74) Mandataires : **VERDURE, Stéphane** etc.; Cabinet Plasseraud, 65/67, rue de la Victoire, F-75440 Paris Cedex 09 (FR).

(25) Langue de dépôt : français

(26) Langue de publication : français

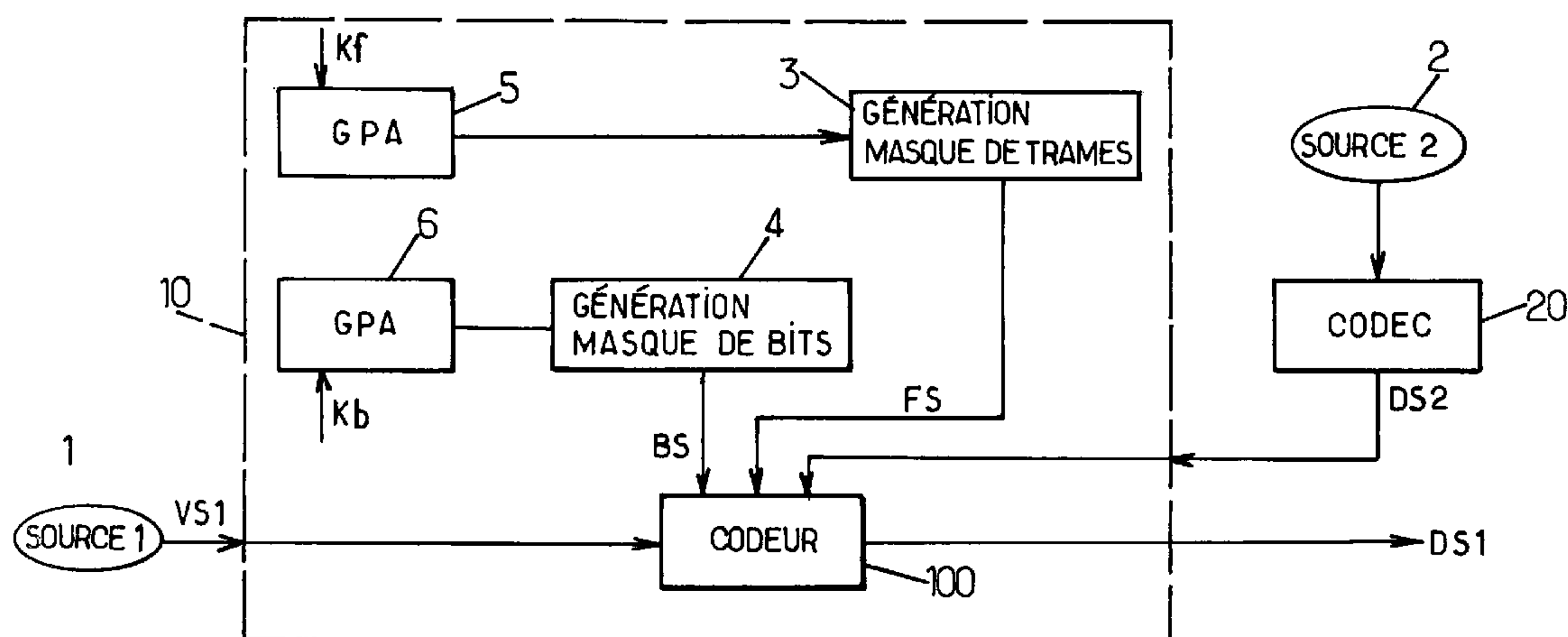
(81) États désignés (sauf indication contraire, pour tout titre de protection nationale disponible) : AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH,

(30) Données relatives à la priorité :
03/10546 5 septembre 2003 (05.09.2003) FR(71) Déposant (pour tous les États désignés sauf US) : **EADS TELECOM** [FR/FR]; Rue Jean-Pierre Timbaud, Batiment

[Suite sur la page suivante]

(54) Title: INFORMATION FLOW TRANSMISSION METHOD WHEREBY SAID FLOW IS INSERTED INTO A SPEECH DATA FLOW, AND PARAMETRIC CODEC USED TO IMPLEMENT SAME

(54) Titre : PROCEDE DE TRANSMISSION D'UN FLUX D'INFORMATION PAR INSERTION A L'INTERIEUR D'UN FLUX DE DONNEES DE PAROLE, ET CODEC PARAMETRIQUE POUR SA MISE EN OEUVRE



3 FRAME MASK GENERATION
4 BIT MASK GENERATION
100 ENCODER

(57) Abstract: The invention relates to a method for the transmission of a secondary information flow (DS2) between a transmitter and a receiver, whereby said secondary information flow is inserted at a parametric vocoder (10) of the transmitter which generates a main information flow (DS1). The aforementioned main information flow is a speech data flow encoding a speech signal (VS1) and is transmitted from the transmitter to the receiver. According to the invention, bits from the secondary information flow are inserted into: (i) only some of the frames of the main information flow, said frames being selected by a frame mask (FS) which is known to the transmitter and the receiver; and/or (ii) into a determined frame of the main information flow, by imposing a constraint on only some of the bits of the frame, said bits being selected by a bit mask (BS) known to the emitter and the receiver.

[Suite sur la page suivante]

WO 2005/024786 A1

WO 2005/024786 A1

PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

Déclaration en vertu de la règle 4.17 :

— *relative à la qualité d'inventeur (règle 4.17.iv)) pour US
seulement*

(84) États désignés (*sauf indication contraire, pour tout titre
de protection régionale disponible*) : ARIPO (BW, GH,
GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), eurasién (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
européen (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI,
SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

Publiée :

— *avec rapport de recherche internationale*

*En ce qui concerne les codes à deux lettres et autres abrégia-
tions, se référer aux "Notes explicatives relatives aux codes et
abréviations" figurant au début de chaque numéro ordinaire de
la Gazette du PCT.*

(57) Abrégé : Un procédé de transmission d'un flux d'information secondaire (DS2) entre un émetteur et un récepteur, comprenant l'insertion dudit flux d'information secondaire au niveau d'un vocodeur paramétrique (10) de l'émetteur générant un flux d'information principal (DS1) qui est un flux de données de parole codant un signal de parole (VS1) et qui est transmis de l'émetteur vers le récepteur, prévoit que des bits du flux d'information secondaire sont insérés dans certaines seulement des trames du flux d'information principal, sélectionnées par un masque de trames (FS) connu de l'émetteur et du récepteur, et/ou à l'intérieur d'une trame déterminée du flux d'information principal, en imposant une contrainte à certains seulement des bits de la trame, sélectionnés par un masque de bits (BS) connu de l'émetteur et du récepteur.

**PROCEDE DE TRANSMISSION D'UN FLUX D'INFORMATION PAR
INSERTION A L'INTERIEUR D'UN FLUX DE DONNEES DE PAROLE, ET
CODEC PARAMETRIQUE POUR SA MISE EN OEUVRE**

La présente invention se rapporte de façon générale au domaine du codage de la parole, et en particulier à un procédé d'insertion d'un flux d'information à l'intérieur d'un flux de données de parole, le flux d'information inséré pouvant être un flux de données de parole à plus faible débit ou un flux de données transparentes.

L'invention trouve des applications, en particulier, dans les systèmes de radiocommunication mobile publics ou professionnels (systèmes PMR, de l'anglais "Professional Mobile Radiocommunication").

On appelle signal de parole un signal acoustique émis par un appareil vocal humain.

On appelle codec une unité matérielle et/ou logicielle de codage et de décodage d'un flux numérique. Sa fonction de codage permet de transcoder un flux numérique d'échantillons quantifiés dans le domaine temporel d'un signal source (par exemple un signal de parole) en un flux numérique comprimé. Sa fonction de décodage permet d'effectuer une opération pseudo-inverse dans l'objectif de restituer des attributs représentatifs du signal source, par exemple des attributs perceptibles dans un récepteur tel que l'oreille humaine.

Un flux de données de parole est un flux de données généré par un codec de parole, à partir du codage d'un signal de parole. Un flux de données transparentes est une suite numérique binaire dont le type de contenu est non spécifié, qu'il soit effectivement un flux de données informatiques ou un flux de données de parole. Les données sont dites transparentes en ce sens que, d'un point de vue externe, tous les bits ont une égale importance vis-à-vis, par exemple de la correction des erreurs de transmission en sorte qu'un codage correcteur d'erreurs doit donc être uniforme sur l'ensemble des bits. A l'inverse, si le flux est un flux de données de parole, certains bits sont plus importants à protéger que d'autres.

Un codec de parole, aussi appelé vocodeur (en anglais "Vocoder" "Speech Codec" ou "Voice Codec") est un codec spécialisé qui est adapté au codage d'un signal de parole quantifié et au décodage d'un flux de trames de paroles. En particulier, il présente pour sa fonction codage une sensibilité qui dépend des caractéristiques de la parole du locuteur et un bas débit binaire

associé à une bande de fréquences plus limitée que la bande de fréquences audio générale (20 Hz-20 kHz).

Il existe plusieurs familles de techniques de codage de la parole, notamment des techniques de codage de la forme d'onde du signal de parole (par exemple le codage ITU-T G.711 MIC loi A/mu), des techniques de codage à modèle de source (le plus connu étant le codage CELP, de l'anglais "Code-Excited Linear Prediction"), des codages perceptuels, et des techniques hybrides fondées sur la combinaison de techniques appartenant à au moins deux des familles ci-dessus.

10 L'invention vise l'application à des techniques de codage "à modèle de source". Ces techniques sont aussi appelées techniques de codage paramétrique, car elles sont basées sur la représentation de paramètres d'excitation de la source de parole et/ou de paramètres décrivant l'enveloppe spectrale du signal émis par le locuteur (par exemple selon un modèle de
15 codage par prédiction linéaire exploitant la corrélation entre les valeurs consécutives des paramètres associés à un filtre de synthèse, ou encore selon un modèle cepstral) et/ou de paramètres acoustiques dépendant de la source, par exemple l'amplitude et la fréquence centrale fondamentale perçue ("Pitch" en anglais), la période ("Pitch period" en anglais) et l'amplitude des pics
20 d'énergie des premières harmoniques d'une fréquence de pitch à différents intervalles, son degré de voisement ("voicing rate" en anglais), sa mélodie et ses enchaînements.

On appelle vocodeur paramétrique un vocodeur mettant en œuvre un codage numérique de la parole utilisant un modèle paramétrique de la source
25 de parole. En pratique, un tel vocodeur associe plusieurs paramètres à chaque trame du flux de parole. Premièrement des paramètres spectraux de prédiction linéaire aussi appelés, par exemple, coefficients LP (de l'anglais "Linear Prediction") ou coefficients LPC (de l'anglais "Linear Prediction Coding"), qui définissent le filtre de prédiction linéaire du vocodeur (filtre à court terme).
30 Deuxièmement des paramètres d'excitation adaptative associés à un (ou plusieurs) vecteur(s) d'excitation adaptative, aussi appelés paramètres LTP (de l'anglais "Long Term Predictor") ou encore coefficients de prédiction adaptative, qui définissent un filtre à long terme sous la forme d'un premier

vecteur d'excitation et d'un gain associé à appliquer en entrée du filtre de synthèse.. Et, troisièmement, des paramètres d'excitation fixe associés à (ou plusieurs) vecteur(s) d'excitation fixe, aussi appelés paramètres algébriques ou paramètres stochastiques qui définissent un second vecteur d'excitation et un gain associé à appliquer en entrée du filtre de synthèse.

Du document EP-A-1 020 848, on connaît un procédé pour transmettre de l'information auxiliaire dans un flux d'information principale correspondant à un signal de parole, ladite information auxiliaire étant insérée au niveau du vocodeur CELP qui code le signal de parole, en remplacement de l'index du vecteur d'excitation adaptative et/ou de l'index du vecteur d'excitation fixe. Plus spécifiquement, les bits d'information auxiliaire sont insérés dans le vocodeur de l'émetteur en lieu et place des bits codant normalement l'index correspondant, et la valeur du gain est mise à zéro afin d'en informer le vocodeur du récepteur.

Selon un inconvénient, l'insertion d'un flux d'information auxiliaire dans le flux n'est pas discrète, en ce sens qu'il suffit de constater la valeur nulle du gain pour savoir que les bits normalement alloués au codage de l'index associé contiennent en fait l'information auxiliaire. Ceci est considéré comme un inconvénient pour la mise en œuvre de la méthode dans un système dans lequel la confidentialité des transmissions est importante.

L'invention a pour principal objet de permettre l'insertion discrète d'un flux secondaire dans un flux principal correspondant à un flux de parole. D'autres objets de l'invention visent à maximiser le débit du flux secondaire pouvant être inséré, tout en préservant au mieux la performance du codage du flux principal vis à vis d'attributs de la source (i.e. en préservant la qualité perçue à l'audition lors de la synthèse du flux de parole). Un autre l'objet de l'invention est aussi de préserver simultanément la performance du codage du flux secondaire vis à vis d'attributs de la source du flux secondaire, notamment lorsqu'il s'agit également d'un flux de parole.

Certains ou la totalité de ces objets sont atteints, selon un premier aspect de l'invention grâce à un procédé de transmission d'un flux d'information secondaire entre un émetteur et un récepteur, comprenant l'insertion dudit flux d'information secondaire au niveau d'un vocodeur paramétrique de l'émetteur

généralant un flux d'information principal qui est un flux de données de parole codant un signal de parole et qui est transmis de l'émetteur vers le récepteur. Des bits du flux d'information secondaire sont insérés :

5 - dans certaines seulement des trames du flux d'information principal, sélectionnées par un masque de trames connu de l'émetteur et du récepteur ; et/ou,

 - à l'intérieur d'une trame déterminée du flux d'information principal, en imposant une contrainte à certains seulement des bits de la trame, sélectionnés par un masque de bits connu de l'émetteur et du récepteur.

10 L'émetteur et le récepteur, de même que la transmission, doivent être interprétés dans leur acception la plus large. Dans un exemple d'application à un système de radiocommunication, l'émetteur et le récepteur sont des équipements terminaux du système, et la transmission est une transmission radio.

15 L'insertion est réalisée au niveau d'un vocodeur paramétrique de l'émetteur qui produit ledit flux d'information principal, sans modification du débit binaire de ce dernier par rapport à ce qu'il serait sans insertion. Dit autrement, le flux d'information secondaire est interprété comme une suite de contraintes sur la suite de valeurs de certains paramètres du modèle de
20 codage paramétrique du flux d'information principal. Par rapport au procédé d'insertion connu dans l'art antérieur, le procédé selon l'invention présente l'avantage que rien dans le flux d'information principal qui est transmis ne trahit la présence du flux d'information secondaire inséré. De plus, en limitant l'insertion à certaines trames et/ou à certaines bits dans une trame seulement,
25 on préserve l'intelligibilité du signal de parole codé dans le flux d'information principal, ce qui n'est nullement le cas avec le procédé d'insertion connu précité.

 Afin de renforcer la discrétion de l'insertion, et donc la robustesse vis-à-vis des tentatives de piratage de la transmission, le masque de trames peut
30 être variable. Il est alors généré selon un algorithme commun parallèlement dans l'émetteur et dans le récepteur, afin d'assurer la synchronisation du codage et du décodage du flux d'information principal, respectivement dans l'émetteur et dans le récepteur.

Le masque de trames peut avantageusement définir une sous-suite de groupes de trames consécutives dans chacune desquelles des bits du flux d'information secondaire sont insérés, afin de profiter de l'effet de glissement du codage qui résulte de la mémorisation des trames dans le vocodeur paramétrique. Ceci contribue à préserver la fidélité du flux d'information principal au signal de parole.

De préférence, la longueur en nombre de trames d'un groupe de trames consécutives est alors sensiblement égale à la profondeur de mémorisation des trames dans le vocodeur paramétrique.

Lorsque le modèle de source du vocodeur paramétrique prévoit, pour certaines au moins des trames du flux d'information principal, différentes classes de bits en fonction de leur sensibilité vis-à-vis de la qualité du codage du signal de parole, le masque de bits peut être tel que des bits du flux d'information secondaire sont insérés dans ces trames en imposant une contrainte en priorité aux bits appartenant à la classe de bits la moins sensible. Ceci contribue aussi à préserver la fidélité du flux d'information principal au signal de parole.

Le flux d'information secondaire peut-être un flux de données de parole ayant un débit plus faible que le débit d'information principal. Ceci est le cas lorsque le flux d'information secondaire sort d'un autre vocodeur ayant un débit plus faible que le débit du vocodeur paramétrique.

Bien entendu, le flux d'information secondaire peut aussi être un flux de données transparentes.

Lorsque le débit du flux d'information secondaire à insérer est trop élevé par rapport au débit du vocodeur paramétrique, on peut être amené à supprimer des bits du flux d'information secondaire, si cela est compatible avec l'application. Inversement, en cas de débit trop faible du flux d'information secondaire, on peut répéter certains bits ou introduire des bits de bourrage.

Le flux d'information secondaire est soumis à un codage correcteur d'erreurs avant insertion dans le flux d'information principal. Ceci permet de pallier le fait que, dans le contexte des vocodeurs paramétriques, certains bits des trames du flux d'information principal sont faiblement voire non soumis à

un codage correcteur d'erreurs (formant codage de canal) avant la transmission.

Dans un mode de mise en oeuvre possible, des bits du flux d'information secondaire sont insérés en imposant des valeurs à des bits qui appartiennent à des paramètres d'excitation d'un filtre du modèle de source du vocodeur paramétrique, par exemple des paramètres d'excitation adaptative et/ou des paramètres d'excitation fixe du filtre de prédiction linéaire d'un vocodeur CELP. Le fait de ne pas imposer de contrainte sur les bits des paramètres de prédiction linéaire préserve l'intelligibilité du flux d'information principal. A cet effet également, on préfère imposer des contraintes bits formant les paramètres d'excitation fixe plutôt que sur ceux formant les paramètres d'excitation fixe.

Dans un mode de mise en oeuvre, des bits du flux d'information secondaire peuvent également être insérés dans des trames de silence du flux d'information principal, à la place ou en plus de l'insertion dans des trames de parole.

Dans un autre mode de mise en oeuvre, des bits du flux d'information secondaire peuvent être insérés en imposant des contraintes à des bits non chiffrés au titre d'un chiffrement de bout en bout du flux d'information principal. Cela permet à un équipement récepteur de pouvoir, après extraction, décoder le flux d'information secondaire bien que n'ayant pas la capacité de déchiffrement à ce titre. Bien entendu, les bits concernés peuvent néanmoins subir une ou plusieurs opérations de chiffrement/déchiffrement à un autre titre, par exemple des chiffrements de lien ou d'interface radio.

Par exemple, la contrainte d'insertion peut être une contrainte d'égalité des bits de la trame du flux d'information principal avec les bits du flux d'information secondaire insérés.

Un deuxième aspect de l'invention se rapporte à un vocodeur paramétrique adapté pour la mise en oeuvre du procédé selon le premier aspect. En ce qui concerne sa fonction de codage, un tel vocodeur paramétrique comprend des moyens d'insertion pour l'insertion d'un flux d'information secondaire dans un flux d'information principal qui est généré par le vocodeur

paramétrique à partir d'un signal de parole. Ces moyens d'insertion sont adaptés pour insérer des bits du flux d'information secondaire :

- dans certaines seulement des trames du flux d'information principal, sélectionnées par un masque de trames déterminé ; et/ou,
- 5 - à l'intérieur d'une trame déterminée du flux d'information principal, en imposant une contrainte à certains seulement des bits de la trame, sélectionnés par un masque de bits déterminé.

Pour sa fonction décodage, le vocodeur comprend des moyens d'extraction du flux d'information secondaire à partir du flux d'information
10 principal.

Un troisième aspect de l'invention se rapporte encore à un équipement terminal d'un système de radiocommunications comprenant un vocodeur paramétrique selon le deuxième aspect.

D'autres caractéristiques et avantages de l'invention apparaîtront
15 encore à la lecture de la description qui va suivre. Celle-ci est purement illustrative et doit être lue en regard des dessins annexés sur lesquels :

- la figure 1 est un diagramme illustrant un exemple de flux de données de paroles codées (flux de parole) organisé en trames et sous-trames ;
- la figure 2 est un schéma synoptique partiel d'un exemple
20 d'équipement émetteur selon l'invention ;
- la figure 3 est un schéma synoptique partiel d'un exemple d'un vocodeur selon l'invention ; et
- la figure 4 est un schéma synoptique partiel d'un exemple de vocodeur utilisé dans équipement récepteur selon l'invention.

25 La figure 1 est un schéma illustrant le principe général de l'insertion d'un flux de données secondaire DS2 dans un flux de données principal DS1 codant un signal de parole VS1. Cette insertion est réalisée au niveau d'un émetteur qui, après multiplexage et codage de canal, émet le flux DS1, et donc le flux DS2 qu'il contient, vers un récepteur distant. Un tel émetteur et un tel
30 récepteur sont par exemple des terminaux mobiles d'un système de radiocommunications public tel que le GSM ou l'UMTS, ou un système de radiocommunications professionnel tel que TETRA ou TETRAPOL.

Le flux DS1 est généré par un vocodeur 10 à partir du signal de parole VS1, lequel est produit par une source de parole 1 telle que l'appareil vocal d'un individu. A cet effet, le signal de parole VS1 est numérisé selon un codage MIC (codage par modulation d'impulsion) linéaire, et segmenté en trames
5 appelées trames de parole. De plus, chaque trame est en général segmentée au niveau du vocodeur 10 en un nombre M fixé de segments appelés sous-trames dans le domaine temporel (modèle CELP) ou dans le domaine fréquentiel (modèle MBE, de l'anglais "Multi-Band Excitation"). Typiquement M est compris entre 2 et 6, selon les vocodeurs). Chaque trame comprend un
10 nombre déterminé N de bits.

La figure 2 illustre un signal de parole numérisé et segmenté en trames $F[i]$ successives, pour i compris entre 0 et l'infini. De plus, au moins pour certains paramètres, chaque trame $F[i]$ peut être segmentée en M sous-trames notées $SF[m]$, pour m compris entre 1 et M. A la figure, on note D la durée
15 d'une trame.

De retour à la figure 1, le vocodeur 10 peut être un vocodeur EFR (de l'anglais "Enhanced Full Rate") du système GSM (voir spécification EN 300 726 GSM 06.60 de l'ETSI), un vocodeur AMR (de l'anglais "Adaptive Multi-Rate") du système UMTS (voir la spécification 3GPP TS26.101 de l'ETSI)
20 pour lequel $D=20$ ms et $M=4$, un vocodeur d'un terminal de radiocommunication TETRA conforme à la spécification EN 300 395-2 de l'ETSI, ou encore un vocodeur TETRAPOL à 6 kbits/s (référéncé dans le rapport ITU-R M.2014) pour lequel le nombre $D=20$ ms, $M=3$ et $N=120$.

Le flux de données secondaire DS2 est par exemple généré par un
25 codec 20, lequel reçoit un flux de données à coder d'une source 2. Dans un exemple d'application de l'invention, la source 2 émet aussi un signal de parole, le codec 20 étant alors un vocodeur de débit inférieur à celui du vocodeur 10. Dans ce cas, le flux DS2 est aussi un flux de trames de parole. Dans cette application, l'invention permet l'insertion discrète d'une
30 communication secondaire dans une communication principale. Le codec 20, plus spécifiquement le vocodeur 20, peut être un vocodeur de type MF-MELP (de l'anglais "Multi-Frame - Mixed Excitation Linear Prediction") à 1200/2400 bits/s décrit dans NATO STANAG 4591.

Eventuellement, le flux DS2 peut être soumis à un codage correcteur d'erreurs, par exemple un codage CRC (de l'anglais "Cyclic Redundancy Code") ou un codage convolutif, qui forme un codage de canal en vue de sa transmission à travers le canal de transmission. En effet, on sait que certains
5 bits des trames du flux de parole DS1 sont peu voire pas protégés par le un codage de canal, en sorte qu'une protection spécifique des bits du flux d'information DS2 peut être requise, selon les applications.

Le vocodeur 10 comprend un codeur 100 qui met en oeuvre un algorithme de codage à modèle de source (ou modèle paramétrique), par exemple de type CELP ou de type MELP. Dans un tel cas, les paramètres correspondant au codage d'une trame de parole côté émetteur incluent, entre autres, des vecteurs d'excitation qui sont soumis, côté récepteur, à un filtre dont la réponse modélise la parole.

Les algorithmes de codage paramétriques utilisent des paramètres
15 calculés soit directement en fonction du flux de trames de paroles entrant et d'un état interne du vocodeur, soit calculés par itérations (sur des trames et/ou des sous-trames successives) en optimisant un critère donné. Typiquement, les premiers paramètres comprennent les paramètres de prédiction linéaire (LP) définissant un filtre court terme, et les seconds paramètres comprennent
20 les paramètres d'excitation adaptative (LTP) définissant un filtre long terme et les paramètres d'excitation fixe. Chaque itération correspond au codage d'une sous-trame dans une trame du flux d'entrée.

Ainsi, par exemple, les paramètres d'excitation adaptative et les paramètres d'excitation fixe sont sélectionnés par itérations successives afin de
25 minimiser l'erreur quadratique entre le signal de parole synthétisé et le signal de parole VS1 original. Dans la littérature anglo-saxonne, cette sélection itérative est parfois appelée "Codebook search" ou "Analysis by Synthesis Search", ou "Error Minimization Loop" ou encore "Closed Loop Pitch Analysis".

En général, les paramètres d'excitation adaptative et/ou les paramètres
30 d'excitation fixe peuvent comprendre chacun, d'une part un indice correspondant à une valeur d'un vecteur dans le dictionnaire adaptatif (dépendant de la sous-trame) ou dans un dictionnaire fixe, respectivement, et d'autre part une valeur de gain associée audit vecteur. Néanmoins, dans

certains vocodeurs tels que le vocodeur TETRAPOL, les paramètres de l'une au moins des excitations adaptative et fixe définissent directement le vecteur d'excitation à appliquer, c'est-à-dire sans adressage d'un dictionnaire par un indice. Dans ce qui suit, il n'est pas fait de distinction entre le mode de
5 définition des vecteurs d'excitation. Les contraintes imposées par les bits du flux DS2 s'appliquant soit à l'indice se rapportant à la valeur du vecteur d'excitation dans le dictionnaire, soit à la valeur de l'excitation elle-même.

En plus du flux de données principal (flux de trames de parole) VS1 et du flux de données secondaire DS2, le vocodeur 10 reçoit, selon l'invention un
10 flux TS de masques de trames, et/ou un flux BS de masques de bits.

Le flux FS est généré par un générateur de masques de trames 3, à partir d'un flux binaire reçu d'un générateur pseudo-aléatoire 5, lequel fonctionne à partir d'une clé secrète K_f connue de l'émetteur et du récepteur. Un masque de trames a pour fonction de sélectionner, parmi un nombre
15 déterminé de trames du flux de trames de parole DS1, celles dans lesquelles, seulement, les bits du flux de données secondaires DS2 sont insérés.

A cet effet, le générateur 3 exécute le processus suivant. Soit la suite des trames $F[i]$ du flux principal DS1, soit h une fonction numérique à valeurs entières, et soit k un nombre entier déterminé, qui est de préférence
20 sensiblement égal à la profondeur de mémorisation de trames successives dans le vocodeur 10 (voir plus loin, nombre P , en référence au schéma de la figure 3), alors les trames $F[h(i)]$, $F[h(i)+1]$, ..., $F[h(i)+k]$ définissent ce qui est appelé ici une sous-suite de groupes de trames de la suite des trames $F[i]$.

Selon une modalité préférée de l'invention, les trames subissant la
25 contrainte d'insertion sont des trames appartenant à une sous-suite de groupes de trames consécutives du flux principal DS1. Ceci permet de profiter de l'effet glissant du codage de parole résultant de la mémorisation de trames prévue dans le vocodeur 10, afin de préserver la qualité du codage du signal de parole VS1 dans le flux principal DS1. C'est pour cela que le nombre k , qui
30 correspond à la longueur en trames d'un groupe de trames, est de préférence égal à, ou du moins proche de la profondeur de mémorisation R du vocodeur 10, ainsi qu'il a été dit plus haut.

Par exemple, en choisissant $h(i) = 10 \times i$ et $k = 5$, alors les trames $F[0]$ à $F[5]$ subissent la contrainte d'insertion, les trames $F[6]$ à $F[9]$ ne subissent pas la contrainte d'insertion, les trames $F[10]$ à $F[15]$ subissent la contrainte d'insertion, les trames $F[16]$ à $F[19]$ ne subissent pas la contrainte d'insertion, etc. Dit autrement, dans cet exemple, 6 trames consécutives sur 10 subissent la contrainte d'insertion.

Le flux BS est quant à lui généré par un générateur de masques de bits 4, à partir d'un flux binaire reçu d'un générateur pseudo-aléatoire 6, lequel fonctionne à partir d'une clé secrète K_b , également connue de l'émetteur et du récepteur. Un masque de bits a pour fonction de sélectionner, parmi les N bits d'une trame du flux de trames de parole DS1 sélectionnée en vertu du masque de trames associée à la trame $F[i]$ courante, ceux qui, seulement, sont contraints par des bits du flux de données secondaire DS2.

A cet effet, le générateur 4 exécute le processus suivant. Il produit un flux d'un nombre fixé S_{max} bits, où S_{max} désigne le nombre maximum de bits d'une trame courante F_i du flux principal DS1 qui peuvent être contraints par des bits du flux secondaire DS2. Un nombre déterminé S de bits parmi ces S_{max} bits, où S est inférieur ou égal à S_{max} ($S \leq S_{max}$), ont la valeur logique 1, les autres ayant la valeur logique 0. Ces S_{max} bits sont insérés dans une chaîne de N bits, à des positions prédéfinies et fixes qui sont prévues dans le logiciel du vocodeur 10, en sorte de former un masque binaire sur la trame. Ce masque, appelé masque de bits, comprend donc S bits égaux à 1. Dans un exemple, lorsqu'un bit du masque de bits est égal à 1, il indique une position d'insertion d'un bit du flux secondaire DS2 dans la trame courante F_i du flux principal DS1.

Le nombre S_{max} est fixé en réalisant un compromis entre le nombre de bits maximum du flux secondaire DS2 qu'on peut insérer dans une trame du flux principal DS1, d'une part, et le souci de préserver la qualité du codage du signal de parole VS1 dans le flux principal DS1, d'autre part. Le nombre S_{max} étant fixé, le nombre S dépend du débit du flux secondaire DS2. Le rapport S/N définit ce qu'on peut appeler le taux d'insertion du flux secondaire DS2 dans le flux principal DS1 pour la trame courante $F[i]$, le rapport S_{max}/N définissant le taux d'insertion maximum.

Dans un exemple où on utilise un vocodeur TETRAPOL (pour lequel $N = 120$) avec $h(i) = 10 \times i$, $k = 5$ et $S = 50$, on obtient pour l'insertion du flux secondaire un canal de débit moyen de 1215 bit/s. Un tel débit permet l'insertion d'un flux de données secondaire généré par un codec de type MF-MELP à 1200 bit/s (nécessitant 81 bits dans 67,5 ms) décrit dans NATO STANAG 4591. Dit autrement, le taux d'insertion obtenu est suffisant pour transmettre discrètement un flux secondaire qui est aussi un flux de parole généré par un vocodeur secondaire 20 de débit inférieur à celui du vocodeur principal 10.

10 Un exemple de contrainte d'insertion consiste à remplacer (i.e., écraser) les bits du flux principal DS1 normalement généré suivant l'algorithme de codage standard mis en œuvre par le vocodeur 10 à partir du signal de parole VS1, par des bits du flux secondaire DS2. Dit autrement, les contraintes appliqués aux paramètres de codage de la parole du flux principal sont des
15 contraintes d'égalité avec les bits du second flux, combinées à des contraintes de sélection par opération ET logique appliquant un masque binaire sur les bits formant le flux principal.

Cet exemple est le plus simple, mais il n'est pas le seul. En effet, des algorithmes sur le flux principal et sur le flux secondaire utilisant toute
20 grammaire contextuelle ou algèbre linéaire ou non-linéaire, y compris l'algèbre de Boole et l'algèbre temporelle de Allen (voir l'article "Maintaining Knowledge about Temporal Intervals", Communications of the ACM, 26/11/1983, pp. 832-84), d'éventuelles mémoires auxiliaires et dépendant de la valeur de paramètres tiers, permettent à l'Homme du métier de définir des contraintes
25 complexes, qui respectent par exemple des propriétés statistiques imposées par le modèle de parole du flux principal.

Notons en particulier que l'ensemble des indices des excitations dans un dictionnaire a généralement une distribution des bits à 0 et à 1 totalement neutre vis à vis d'une analyse statistique d'occurrences. Il est généralement
30 possible de chiffrer le flux secondaire DS2 sous une forme pseudo-aléatoire avant insertion, sans modifier la distribution statistique des 0 et 1 dans les bits modifiés du flux principal. Dans l'hypothèse d'un modèle de codage de parole conduisant à un flux codé dont certaines sous-trames auraient une corrélation

vers 0 ou vers 1, le générateur pseudo-aléatoire susmentionné ou un algorithme de chiffrement du flux secondaire devront aussi avoir ce biais.

Ainsi qu'on l'aura compris, le nombre de bit constraints lors du codage varie d'une trame à l'autre selon une loi d'évolution connue de l'émetteur et du
5 récepteur, qui sont supposés être synchronisés.

La synchronisation de l'émetteur et du récepteur en ce qui concerne l'application des masques de trame et/ou des masques de bits résulte de la synchronisation générale entre ces deux équipements. Typiquement, cette synchronisation est assurée par l'étiquetage des trames à l'aide de valeurs
10 générées par un compteur de trames. De façon connue, la synchronisation générale entre l'émetteur et le récepteur peut aussi provenir, en totalité ou en complément, d'éléments de synchronisation (motifs de bits particuliers) insérés dans le flux principal DS1.

Le codeur 100 de l'émetteur et le décodeur du récepteur partagent une
15 même information initiale permettant de déterminer la sous-suite des groupes trames et des sous-trames où l'insertion du flux secondaire a lieu. Cette information peut comprendre un vecteur d'initialisation des générateurs pseudo-aléatoires 5 et 6. Elle peut être fixe. Elle peut aussi dépendre, par exemple, du débit moyen imposé par le flux secondaire, ou encore dépendre
20 de paramètres non contraints du codec principal 10 calculés lors du codage du flux principal.

Ainsi qu'il est représenté à la figure 3, le codeur 100 comprend un module 11 qui est un module matériel et/ou logiciel de synthèse des paramètres de prédiction linéaire, recevant en entrée le signal de parole VS1 et
25 délivrant en sortie une information LP correspondant aux paramètres de prédiction linéaire (coefficients du filtre de prédiction linéaire court terme). L'information LP est passée en entrée d'une unité logique 12, par exemple un multiplexeur, qui est commandée par le flux de masques de trames FS et le flux de masques de bits BS. L'unité 12 génère en sortie une information LP'
30 correspondant à l'information LP dont certains bits au moins pour certaines trames au moins, ont été altérés par application des contraintes résultant du flux secondaires DS2 via le masque de trame et le masque de bits associés à la trame courante. Une mémorisation de l'information LP', avec une profondeur

de mémorisation correspondant à un nombre déterminé P de trames successives, peut-être prévue pour le module 11.

Le codeur 100 comprend aussi un module 21 qui est un module matériel et/ou logiciel de synthèse des paramètres d'excitation adaptative, recevant en entrée l'information LP' et délivrant en sortie une information LTP correspondant aux paramètres d'excitation adaptative (définissant un premier vecteur de quantification et un gain un associé pour le filtre de synthèse court terme). L'information LTP est passée en entrée d'une unité logique 22, par exemple un multiplexeur, qui est commandée par le flux de masques de trames FS et le flux de masques de bits BS. L'unité 22 génère en sortie une information LTP' correspondant à l'information LTP dont certains bits au moins pour certaines trames et/ou pour certaines sous-trames au moins, ont été altérés par application des contraintes résultant du flux secondaires DS2 via le masque de trame et le masque de bits associés à la trame courante. Une mémorisation de l'information LTP', avec une profondeur de mémorisation correspondant à un nombre déterminé Q de sous-trames successives de la trame courante ($Q \leq M-1$), peut-être prévue pour le module 21.

Le codeur 100 comprend enfin un module 31 qui est un module matériel et/ou logiciel de synthèse des paramètres d'excitation fixe, recevant en entrée l'information LTP' et délivrant en sortie une information FIX correspondant aux paramètres d'excitation fixe (définissant un second vecteur de quantification et un gain un associé pour le filtre de synthèse court terme). L'information FIX est passée en entrée d'une unité logique 32, par exemple un multiplexeur, qui est commandée par le flux de masques de trames FS et le flux de masques de bits BS. L'unité 32 génère en sortie une information FIX' correspondant à l'information FIX dont certains bits au moins pour certaines trames et/ou pour certaines sous-trames au moins, ont été altérés par application des contraintes résultant du flux secondaires DS2 via le masque de trame et le masque de bits associés à la trame courante. Une mémorisation de l'information FIX', avec une profondeur de mémorisation correspondant à un nombre déterminé R de sous-trames successives de la trame courante ($R \leq M-1$), est prévue pour le module 21. De plus, une mémorisation de l'information FIX', avec une profondeur de mémorisation correspondant par exemple à un

nombre déterminé W de sous-frames successives de la trame courante ($W \leq M-1$), peut-être prévue pour le module 21.

Pour chaque trame courante, l'information $LP'(F[i])$ correspondant aux paramètres de prédiction linéaire de la trame, les informations $LTP'(SF[1]), \dots,$
 5 $LTP'(SF[M])$ correspondant aux paramètres d'excitation adaptative respectivement pour chacune des sous-frames $SF[1]$ à $SF[M]$ de la trame, et les informations $FIX'(SF[1]), \dots, FIX'(SF[M])$ correspondant aux paramètres d'excitation fixe respectivement pour chacune des sous-frames $SF[1]$ à $SF[M]$ de la trame, sont transmises en entrée d'un multiplexeur 41 qui les concatène
 10 pour former une trame du flux principal DS1.

Les mémorisations dont il est question ci-dessus permettent ici d'atténuer l'effet des contraintes appliquées aux bits des paramètres de prédiction linéaire, des paramètres d'excitation adaptative et/ou des paramètres d'excitation fixe, vis-à-vis de la fidélité du flux principal DS1 au
 15 signal de parole source VS1. En effet, ces mémorisations permettent un effet de glissement dans le calcul des paramètres, en sorte que, pour une trame déterminée, les contraintes appliquées à des premiers paramètres sont au moins partiellement compensées, du point de vue perceptuel, par le calcul de paramètres calculés ensuite à partir d'une synthèse de parole basée sur lesdits
 20 premiers paramètres.

Plus spécifiquement, on peut écrire les relations suivantes, où f désigne une fonction traduisant l'analyse par synthèse :

- 1°) $LP'(F[i]) = f(LP'(F[i-1]), LP'(F[i-2]), \dots, LP'(F[i-P]))$;
- 2°) $LTP'(SF[i]) = f(LTP'(SF[i-1]), \dots, LTP'(SF[i-R]), FIX'(SF[i-1]), \dots, FIX'(SF[i-W]))$;
- 25 3°) $FIX'(SF[i]) = f(FIX'(SF[i-1]), \dots, FIX'(SF[i-W]))$.

Ces compensations, et aussi le fait que l'insertion des bits du flux secondaire n'est pas aléatoire, permettent d'atteindre en pratique, pour certains vocodeurs, des taux d'insertion de l'ordre de 10 % sans générer de dégradation (du point de vue perceptuel) du signal de parole VS1 supérieure à ce que
 30 génère un taux d'erreur bit résiduel (après codage canal) de l'ordre de quelques %.

On va maintenant décrire les implications du procédé côté récepteur.

Notons tout d'abord que, pour un équipement récepteur ne traitant pas le flux secondaire DS2, le décodage des trames du flux DS1 reçues, est seul effectué selon l'algorithme de synthèse standard du vocodeur 10 de l'équipement émetteur.

5 Pour un équipement récepteur traitant le flux secondaire DS2, la récupération de l'information codée par les bits de ce flux secondaire nécessite une synchronisation de l'équipement avec l'équipement émetteur, des moyens d'extraction du flux secondaire DS2 à partir du flux principal DS1. identique au codec 20 de l'équipement émetteur.

10 On se réfère au schéma de la figure 4, qui montre de manière schématique les moyens d'un vocodeur 10a d'équipement récepteur destiné au traitement du flux secondaire transmis grâce au procédé selon l'invention.

Le vocodeur 10a, le cas échéant après démultiplexage et décodage canal, reçoit le flux principal DS1 en entrée, et délivre un signal de parole VS1' en sortie.

15 Le signal VS1' est moins fidèle au signal de parole source VS1 (figure 3) qu'il ne le serait en absence de mise en œuvre du procédé d'insertion selon l'invention. Ceci traduit la perte de qualité du codage effectué côté émetteur, en raison des contraintes extérieures appliquées au vocodeur 1 de l'équipement émetteur.

20 L'équipement récepteur peut aussi comprendre un moyen de restitution du signal de parole VS1', par exemple un haut-parleur ou similaire.

Ainsi qu'il a déjà été dit plus haut, les protocoles de transmission connus prévoient une synchronisation générale de l'équipement récepteur avec l'équipement émetteur. La mise en œuvre de l'invention ne requiert donc pas

25 de moyens particuliers à cet égard.

Pour l'extraction du flux secondaire, le vocodeur 10a comprend un générateur de masques de trames 3a et un générateur de masques de bits 4a, respectivement associés à un générateur pseudo-aléatoire 5a et à un

30 générateur pseudo-aléatoire 6a, qui sont identiques et agencés de la même façon que les moyens respectivement 3, 4, 5 et 6 du vocodeur 10 de l'équipement émetteur (figure 3). On notera que les générateurs 5a et 6a du récepteur reçoivent la même clé secrète, respectivement Kf et Kb, que les

générateurs 5 et 6 du vocodeur 10 de l'équipement émetteur. Ces clés sont stockées dans une mémoire ad hoc des équipements. Les générateurs 3a et 4a génèrent respectivement un flux de masques de trames FSa et un flux de masques de bits BSa, qui sont fournis en entrée du d'un décodeur 100a du vocodeur 10a.

L'extraction des bits du flux secondaire DS2 se fait par application synchrone (par exemple via des opération ET logique) des masques de trames et des masques de bits en entrée du décodeur 100a (par exemple via des opération ET logique), sans que cela affecte le décodage du flux principal DS1 par ce dernier. A cet effet, le flux DS1 est fourni en entrée du décodeur 100a via une unité logique 7a, qui extrait le flux d'information secondaire DS2 du flux d'information principal DS1 sous la commande du flux de masques de trames FSa et du flux de masques de bits BSa.

L'équipement récepteur peut aussi comprendre un codec secondaire, identique au codec 20 de l'équipement émetteur pour le décodage du flux secondaire DS2. Lorsque ce flux est un flux de parole, le codec secondaire génère un signal de parole qui peut être restitué via un haut parleur ou similaire.

On notera que la fluctuation du taux de transmission des bits du flux secondaire DS2 ne pose pas de problème particulier côté récepteur, dès lors que le flux secondaire DS2 est fourni en entrée d'un codec secondaire à débit variable comme c'est le cas de tous les vocodeurs du marché. En effet, un tel codec comprend une mémoire tampon d'entrée ("Input Buffer" en anglais) dans laquelle les données du flux DS2 sont stockées en vue de leur décodage. Il faut juste s'assurer que la mémoire tampon d'entrée n'est jamais vide. A cet effet, on détermine le taux d'insertion qui convient, en tenant compte en particulier du débit binaire du codeur 100 et du vocodeur secondaire 20 et des objectifs de préservation de la fidélité du flux principal VS1 au signal de parole VS1. Compte tenu des taux d'insertion élevés obtenus en pratique (de l'ordre de 10 %), cette question de l'alimentation du vocodeur secondaire de l'équipement récepteur ne devrait pas poser de problème, avec un vocodeur principal 10 de type AMR dans son mode de codage à 12,2 kbits/s et un vocodeur secondaire 20 de débit environ dix fois moindre.

Par ailleurs, dans le cas où le flux secondaire est un flux de parole et afin de fournir au second décodeur un flux régulier de trames, on peut optionnellement mémoriser les séquences et de ne pas commencer immédiatement le décodage.

5 Dans le cas où le flux secondaire est un flux de données transparentes, il est proposé de les concaténer et de les traiter comme si elles avaient été transmises au moyen d'une messagerie courte de longueur maximale (service SMS en GSM, par exemple), et d'y adjoindre un code convolutif correcteur d'erreurs. Alternativement, le flux de données
10 transparentes peut être envoyé à un module de chiffrement ou à un module de transcodage et de synthèse de type "Text-to-Speech ».

Revenons maintenant à la description générale des modalités de mise en œuvre du procédé de transmission selon l'invention.

Le choix des bits d'une trame déterminée du flux principal qui subissent
15 l'application de la contrainte du flux secondaire est déterminé selon les particularités de chaque application. On donne ci-après plusieurs modes de mise en œuvre possibles à cet égard, ainsi que d'autres particularités et avantages de l'invention.

Dans un mode de mise en œuvre possible, des contraintes sont
20 imposées lors du codage sur la valeur de zéro, plusieurs ou tous les bits de la trame qui sont associés à un vecteur d'excitation de type déterminé, adaptative ou fixe, avant d'effectuer les itérations permettant de calculer les paramètres qui dépendent dudit vecteur d'excitation en vertu des mémorisations réalisées dans le vocodeur. Ces bits de valeur contrainte sont alors les informations du
25 flux secondaire transportées par la trame et constituent le canal du flux d'information secondaire DS2. Autrement dit, le flux secondaire est inséré en imposant des valeurs à des bits formant les paramètres des vecteurs d'excitation adaptative ou fixe. Ceci peut éventuellement être étendu en appliquant des contraintes simultanément aux vecteurs d'excitation de l'autre
30 type, respectivement fixe ou adaptative.

Lorsque la transmission entre l'émetteur et le récepteur prévoit un chiffrement partiel des trames du flux principal (c'est-à-dire un chiffrement de certains bits seulement dans chaque trame), le masque de bits peut

avantageusement coïncider avec un ensemble de bits non chiffrés d'une trame. Ceci permet à l'équipement récepteur jouant le rôle de passerelle d'effectuer l'extraction du flux secondaire inséré dans le flux principal sans disposer des moyens de déchiffrer le flux principal.

5 Ceci est particulièrement utile tout en préservant la confidentialité du flux principal, sous l'hypothèse approximative de linéarité du modèle de parole du vocodeur, c'est-à-dire en considérant que les paramètres résiduels ou d'excitation des cordes vocales sont non corrélés aux coefficients décrivant l'enveloppe spectrale de réponse du conduit vocal.

10 Autrement dit, ce mode de mise en oeuvre du procédé est caractérisé en ce que le flux d'information secondaire est inséré en imposant des contraintes à des bits non chiffrés de paramètres du modèle de parole du flux principal.

Ce mode de mise en oeuvre est illustré par un exemple concernant un
15 vocodeur EFR (voir plus haut) utilisé comme codec principal. On choisit d'utiliser des bits parmi les bits non protégés de chaque trame comme canal pour le flux secondaire, en écrasant leur valeur calculée par l'algorithme de codage source du flux principal par application d'un masque binaire sur les 78 bits non protégés de chaque trame. Ces 78 bits non protégés sont identifiés
20 dans le tableau 6 (intitulé "Ordering of Enhanced Full Rate Speech Parameters for the Channel Encoder" dans la spécification ETSI EN 300 909 V8.5.1 GSM 05.03 "Channel coding") et concernent un sous-ensemble des bits décrivant les vecteurs d'excitation fixes. Avec ces 78 bits de classe 2 par trame de 20 ms, on obtient un canal secondaire de débit nominal 3900 bit/s. On peut utiliser de
25 préférence les bits les moins sensibles du mode de codage 12,2 kbit/s du codec AMR (voir plus haut) identifiés par ordre de sensibilité dans le tableau B.8 (intitulé "Ordering of the Speech Encoder Bits from the 12,2 kbit/s Mode" dans la spécification 3GPP TS26.101 "Adaptive Multi-Rate (AMR) Speech Codec Frame Structure").

30 Il est donc également possible d'introduire, dans le mode de codage à 12,2 kbits/s du codec AMR, le flux d'un codec secondaire, par exemple le codeur MELP 1200/2400 bit/s décrit dans NATO STANAG 4591, nécessitant 81 bits par 67,5 ms à 1200 bits/s (respectivement 54 bits par 22,5 ms 2400

bits/s), enrobé dans son propre codage correcteur d'erreur (taux 2/3 FEC), par exemple, qui protège 100% des bits à 1200 bit/s (respectivement 50% des bits à 2400 bit/s), et/ou enrobé dans des trames de négociation d'interopérabilité de sécurité de type FNBDT ("Future Narrow Band Digital Terminal") définies par
5 l'OTAN, ou d'un type de protocole de sécurité plus léger.

Dans un autre mode de mise en œuvre, applicable aux vocodeurs utilisant un algorithme basé sur la sélection d'excitations quantifiées dans un dictionnaire, la contrainte consiste à imposer une valeur d'excitation déterminée, tirée du dictionnaire. En variante, le dictionnaire est partitionné en
10 plusieurs sous-dictionnaires, et la contrainte consiste à imposer l'un des sous-dictionnaires. Une autre variante comprend la combinaison des deux types de contrainte ci-dessus. Lors du décodage du flux principal côté récepteur, la connaissance de l'excitation reçue permet d'identifier le sous-dictionnaire et/ou l'excitation concernés, et d'en déduire la contrainte qui détermine les bits du
15 flux secondaire. Notons qu'à une permutation près des excitations, la contrainte d'imposition du sous-dictionnaire peut être équivalente à l'application des contraintes sur les bits de poids faible des indices d'excitation dans le dictionnaire.

Dans un autre mode de mise en œuvre, le flux secondaire définit un
20 codage différentiel des indices de vecteurs d'excitation, par exemple de vecteurs d'excitation fixes, dans la sous-suite de trames successives du flux principal.

Dans un autre mode de mise en œuvre, les bits contraints peuvent être les bits de poids faibles des excitations fixes (c'est-à-dire des excitations non
25 adaptatives) pour chaque trame de parole et éventuellement pour chaque sous-trame définie dans la trame de parole au sens de l'algorithme de codage du vocodeur 10.

Dans un autre mode de mise en œuvre, le nombre et la position des bits contraints sont identifiés pour chaque trame successive en fonction d'un
30 algorithme de calcul d'un masque et d'un élément secret connu des de l'émetteur et du récepteur, afin d'accroître les chances de non-détection de l'existence du flux secondaire par un tiers.

Un autre mode de mise en œuvre, applicable à un algorithme de codage nécessitant plusieurs vecteurs d'excitation fixe par trame ou sous-trame, tel que le codec CELP pour la parole d'un flux MPEG-4 (défini dans la spécification ISO/IEC 14496-3 Sub-part 3) pour lequel certaines excitations
5 fixes d'une trame sont choisies à partir de calculs précédents et où d'autres excitations fixes de la même trame sont calculés par analyse par synthèse sur un dictionnaire (voir la spécification ISO/IEC 14496-3 §7.9.3.4 "Multi-Pulse Excitation for the bandwidth extension tool"), consiste à imposer la contrainte sur le choix par dictionnaire de la première excitation fixe et à utiliser ensuite
10 les itérations d'analyse par synthèse sur la seconde excitation fixe pour rattraper l'erreur imposée par la contrainte sur la première excitation fixe.

Dans un autre mode de mise en œuvre, la sous-suite des trames du flux principal qui sont concernées par l'insertion du flux secondaire ne comprend que les trames qui présentent suffisamment d'énergie et de parole
15 au sens du vocodeur. Dans une variante applicable par exemple aux vocodeurs MELP (qui définissent plusieurs niveaux de voisement) ou aux vocodeurs HVXC (de l'anglais "Harmonic Vector eXcitation Codec", qui sont des vocodeurs paramétriques d'un flux de parole MPEG-4 définis dans la spécification ISO/IEC 14496-3 Sub-part 2) la sous-suite ne concerne que les
20 segments peu voisés ou totalement non voisés des trames.

Lorsque la contrainte est appliquée sur les paramètres d'excitation, par exemple sur les indices d'excitation fixes, les paramètres d'une sous-trame du flux principal DS1 restent tout à fait conformes au modèle de codage de parole du vocodeur 10. Néanmoins, la séquence des excitations fixes modifiées est
25 peut-être statistiquement atypique pour une parole humaine ou éventuellement atypique pour le procédé de reconnaissance du locuteur, selon les contraintes appliquées et l'objectif de fidélité souhaité. Pour éviter que la présence du flux secondaire dans ces excitations ne puisse être détectée dans un équipement récepteur, un traitement des paramètres comprenant un lissage des gains des
30 excitations fixes associé à un traitement des impulsions isolées des vecteurs d'excitation suivi d'un post-filtrage après la synthèse de parole, peuvent être appliqués au décodage. Ces traitements permettent d'exclure des séquences acoustiques apparaissant après transmission dans un canal bruité, qui seraient

impossible à prononcer par un appareil vocal humain dans l'ambiance d'un microphone. Il s'agit par exemple de certaines séquences de cliquetis, chuintement, crissements, sifflements ou autres, dans le bruit de fond que le vocodeur standard n'aurait pas suffisamment filtrées lors de la synthèse de parole du fait des contraintes imposées. C'est ainsi que peuvent être rendus imperceptibles des sons non voisés indésirables, qui seraient corrélés aux séquences d'excitation fixe contraintes selon le procédé de l'invention.

Néanmoins, lorsque l'application de contraintes risque de conduire à la perception de sons non-voisés indésirables corrélés à une séquence d'excitation fixe atypique d'une parole humaine et non filtrée par le filtrage du décodeur standard du vocodeur, la sous-suite des trames sur lesquelles sont appliquées les contraintes peut être définie en fonction d'analyses statistiques préalables sur les valeurs des paramètres consécutifs du modèle de parole du vocodeurs, par exemple en tirant parti de la texture des paramètres de la parole, définie par une inertie, une entropie ou une énergie dérivée de la probabilité des séquences de valeurs des paramètres, par exemple dans huit trames consécutives représentatives de la durée d'un phonème.

Pour chaque mode de mise en oeuvre, la performance de la synthèse du flux principal DS1, c'est-à-dire la fidélité au signal VS1, est inversement proportionnelle au débit relatif du flux secondaire DS2. La performance de fidélité subjective à la source 1 du signal de parole VS1 peut toutefois être atteinte lorsque le procédé proposé garde invariants certains attributs subjectifs (par exemple certains critères psycho-acoustiques) de la source 1. Elle peut être mesurée par des mesures statistiques ("Mean Opinion Score", ou MOS) selon une échelle standardisée (voir la recommandation ITU-T P.862 "Perceptual evaluation of speech quality –PESQ").

Dans certains modes de réalisation, la dégradation de la qualité subjective du flux de parole DS1 issu du vocodeur 10, qui est due à l'insertion du flux secondaire DS2, est supposée acceptable pour justifier l'application du procédé proposé. C'est en particulier le cas lorsque le flux secondaire est également un flux de parole et que le contenu auditif du flux principal est bien moins important que le contenu du flux secondaire pour l'auditeur légitime. En effet, la perception psycho-acoustique de la présence éventuelle du flux

secondaire lors de l'écoute du flux principal décodé et restitué ne permet pas d'aider à localiser le flux secondaire dans le flux principal et donc d'apporter une preuve formelle de son existence. Ceci est en particulier le cas pour un vocodeur 10 à bas débit utilisé dans un environnement bruité, car le décodage et la restitution du flux principal DS1 fournissent des séquences de parole conformes au modèle du vocodeur 10. C'est aussi le cas, dans certaines limites psycho-acoustiques, lorsque le débit minimal du flux secondaire doit être assuré au détriment de la qualité de restitution du flux principal.

Afin de préserver au mieux l'intelligibilité de la synthèse du flux principal DS1, on préfère ne pas appliquer de contraintes sur les paramètres spectraux de prédiction linéaire (LP) définissant le filtre court terme, et ne pas trop perturber les paramètres à long terme (LTP) adaptés à chaque sous-trame, afin de conserver des caractéristiques subjectives jugées essentielles dans le signal de parole VS1. En particulier, un mode de mise en oeuvre consiste à appliquer de préférence les contraintes sur des sous-frames différentes des sous-frames sur lesquelles les fenêtres d'analyse à long terme de la trame sont concentrées, à savoir, par exemple, la seconde et la quatrième sous-trame pour le mode de codage 12,2 kbit/s du vocodeur AMR évoqué supra (voir la spécification 3GPP TS 26.090 V5.0.0, § 5.2.1 "Windowing and auto-correlation computation"). En particulier, on évitera de perturber beaucoup de segments voisés, généralement porteurs de la majorité des caractéristiques d'identification du locuteur.

A titre d'exemple élaboré, dans le mode de codage 12,2 kbit/s du vocodeur AMR, il est possible d'imposer une contrainte sur le choix de l'excitation adaptative en imposant des valeurs initiales aux échantillons $u(n)$ $n=0, \dots, 39$, dans l'équation réursive (38) de calcul du vecteur adaptatif décrit au paragraphe 5.6.1 (intitulé "Adaptative Codebook Search") de la spécification 3GPP TS 26.090 évoquée supra, en substituant aux valeurs du résidu LP, calculé dans l'équation (36), 40 valeurs extraites du flux secondaire. L'erreur entre le signal du flux principal et le signal synthétisé par le filtre court terme avec la contribution du vecteur adaptatif contraint est compensée par le choix du vecteur d'excitation fixe qui tente de rattraper l'erreur résiduelle (par exemple l'erreur résiduelle quadratique) de la prédiction long terme sur la

même sous-trame, ainsi que les vecteurs d'excitation des sous-frames successives. Ainsi les vecteurs d'excitation contraints codent le flux secondaire comme résidu adaptatif au dessus de la réponse du filtre de synthèse court terme du flux principal corrigé par le résidu fixe.

5 Dans une autre exemple, pour un modèle de parole du vocodeur paramétrique de type STC (de l'anglais "Sinusoidal Transform Coding") ou de type MBE ("Multi Band Excitation") par exemple selon le standard spécifications ANSI/TIA/EIA 102.BABA ("APCO Project 25 Vocoder Description"), un mode de mis en oeuvre conduit à s'intéresser aux bits de
10 poids faible des paramètres d'amplitude des harmoniques des segments des trames ou aux paramètres d'amplitudes d'échantillons de l'enveloppe spectrale. Dans un codec MBE, les paramètres d'excitation sont la fréquence fondamentale ainsi que la décision voisé/non-voisé pour chaque bande de fréquences.

15 Dans ce qui précède, on a décrit des modes de mise en oeuvre prévoyant l'insertion des bits du flux secondaire dans des trames de parole du flux principal. Néanmoins, on sait que le flux principal DS1 contient aussi des trames de silence, qui sont des trames codées par le vocodeur 10 avec un moindre débit binaire et émise avec une périodicité moindre que les trames de
20 parole, pour synthétiser lorsque les périodes de silences contenues dans le signal de parole VS1. Ces trames de silence synthétise ce qu'on appelle un bruit de confort.

Or, un mode de mise en oeuvre du procédé peut prévoir, en variante ou en complément, l'insertion du flux secondaire via des contraintes numériques
25 sur les valeurs des paramètres descripteurs du bruit de confort à générer au titre du flux principal.

Ce mode de mise en oeuvre est illustré par un exemple concernant un codec EFR ou AMR (voir plus haut) utilisé comme codec principal. Dans les système GSM et UMTS, les trames transportant du bruit de confort (trames de
30 silence) sont nommées trames SID (voir par exemple la spécification 3GPP TS 26.092 "Mandatory Speech Codec Speech Processing Functions ; AMR Speech Codec ; Confort Noise Aspects" de l'ETSI). Plus précisément les

trames considérées ici sont les trames SID-UPDATE qui contiennent 35 bits de paramètres de bruit de confort et un code correcteur d'erreur sur 7 bits.

Dans un système GSM ou UMTS, c'est la source qui contrôle l'émission des trames de silence, c'est-à-dire le codec de l'émetteur (sous
5 réserve des interactions avec le processus de détection d'activité vocale et de transmission discontinue, en particulier sur la voie descendante du relais vers le terminal mobile). Il est donc possible de procéder par insertion du second flux selon un procédé similaire à celui applicable à une trame contenant suffisamment d'énergie de parole (trame de parole).

10 De manière alternative, il est possible de commander l'émission d'une trame de silence particulière à partir de l'entrée analogique numérisée du codec en générant le bruit analogique de confort représentatif des 35 bits du flux secondaire. Dans les systèmes GSM et UMTS, la fréquence des trames de silence est contrôlée par la source ou par le relais et correspond soit à une
15 trame de silence toutes les 20 ms soit à une trame de silence toutes les 160 ms, soit encore à une trame de silence toutes les 480 ms pour le codec EFR du système GSM. Ceci détermine le débit maximal pour le flux secondaire dans cette variante du procédé.

Dans une modalité particulière, il est possible d'utiliser le canal de
20 transmission duplex pour envoyer des trames de silence lorsque le locuteur est un second participant à la communication ou dans les silences dans une première conversation, c'est-à-dire entre les groupes de phonèmes émis selon le flux principal.

On notera que la spécification 3GPP TS 26.090 précise que la taille du
25 champ de codage du bruit de confort du codec EFR, à savoir 35 bits par trame de silence, est identique à la taille du paramètre d'excitation fixe pour ce même codec. Cela signifie qu'on peut appliquer les mêmes contraintes et obtenir un débit d'insertion minimal permanent en utilisant toutes les trames indépendamment de la nature, parole ou silence, du flux principal.

REVENDICATIONS

1. Procédé de transmission d'un flux d'information secondaire entre un émetteur et un récepteur, comprenant l'insertion dudit flux d'information secondaire au niveau d'un vocodeur paramétrique de l'émetteur générant un flux d'information principal qui est un flux de données de parole codant un signal de parole et qui est transmis de l'émetteur vers le récepteur, suivant lequel des bits du flux d'information secondaire sont insérés :
 - dans certaines seulement des trames du flux d'information principal, sélectionnées par un masque de trames connu de l'émetteur et du récepteur ; et/ou,
 - 10 - à l'intérieur d'une trame déterminée du flux d'information principal, en imposant une contrainte à certains seulement des bits de la trame, sélectionnés par un masque de bits connu de l'émetteur et du récepteur.
2. Procédé selon la revendication 1, suivant lequel le masque de trames est variable et est généré selon un algorithme commun parallèlement dans l'émetteur et dans le récepteur.
3. Procédé selon la revendication 1 ou la revendication 2, suivant lequel le masque de trames définit une sous-suite de groupes de trames consécutives dans chacune desquelles des bits du flux d'information secondaire sont insérés.
4. Procédé selon la revendication 3, suivant lequel la longueur en nombre de trames d'un groupe de trames consécutives est sensiblement égale à la profondeur de mémorisation des trames dans le vocodeur paramétrique.
5. Procédé selon l'une quelconque des revendications précédentes, suivant lequel, le modèle de source du vocodeur paramétrique prévoyant, pour certaines au moins des trames du flux d'information principal différentes classes de bits en fonction de leur sensibilité vis à vis de la qualité du codage du signal de parole, le masque de bits est tel que des bits du flux d'information

secondaire sont insérés dans ces trames en imposant une contrainte en priorité aux bits appartenant à la classe de bits la moins sensible.

6. Procédé selon l'une quelconque des revendications 1 à 5, suivant lequel
5 le flux d'information secondaire est un flux de données de parole sortant d'un autre vocodeur ayant un débit plus faible que le débit du vocodeur paramétrique.

7. Procédé selon l'une quelconque des revendications 1 à 5, suivant lequel
10 le flux d'information secondaire est un flux de données transparentes.

8. Procédé selon l'une quelconque des revendications précédentes, suivant lequel le flux d'information secondaire est soumis à un codage correcteur d'erreurs avant insertion dans le flux d'information principal.

15

9. Procédé selon l'une quelconque des revendications précédentes, suivant lequel des bits du flux d'information secondaire sont insérés en imposant des valeurs à des bits qui appartiennent à des paramètres d'excitation d'un filtre du modèle de source du vocodeur paramétrique.

20

10. Procédé selon l'une quelconque des revendications précédentes, suivant lequel des bits du flux d'information secondaire sont insérés dans des trames de silence du flux d'information principal.

25 11. Procédé selon l'une quelconque des revendications précédentes, suivant lequel des bits du flux d'information secondaire sont insérés en imposant des contraintes à des bits non chiffrés au titre d'un chiffrement de bout en bout du flux d'information principal.

30 12. Procédé selon l'une quelconque des revendications précédentes, suivant lequel la contrainte est une contrainte d'égalité des bits de la trame du flux d'information principal avec les bits du flux d'information secondaire insérés.

13. Vocodeur paramétrique comprenant, pour l'insertion d'un flux d'information secondaire dans un flux d'information principal qui est généré par le vocodeur paramétrique à partir d'un signal de parole, des moyens d'insertion
5 adaptés pour insérer des bits du flux d'information secondaire :
- dans certaines seulement des trames du flux d'information principal, sélectionnées par un masque de trames déterminé ; et/ou,
 - à l'intérieur d'une trame déterminée du flux d'information principal, en imposant une contrainte à certains seulement des bits de la trame,
10 sélectionnés par un masque de bits déterminé .
14. Vocodeur paramétrique selon la revendication 13, dans lequel le masque de trames est variable et est généré selon un algorithme à base d'une clé secrète.
- 15
15. Vocodeur paramétrique selon la revendication 13 ou la revendication 14, dans lequel le masque de trames définit une sous-suite de trames consécutives dans chacune desquelles des bits du flux d'information secondaire sont insérés.
- 20
16. Vocodeur paramétrique selon l'une quelconque des revendications 13 à 15, dans lequel la longueur en nombre de trames de la sous-suite de trames consécutives est sensiblement égale à la profondeur de mémorisation des trames dans le codec de parole paramétrique.
- 25
17. Vocodeur paramétrique selon l'une quelconque des revendications 13 à 16, dans lequel, le modèle de source du Vocodeur paramétrique prévoyant dans certaines au moins des trames du flux d'information principal différentes classes de bits en fonction de leur sensibilité vis à vis de la qualité du codage
30 du signal de parole, le masque de bits est tel que des bits du flux d'information secondaire sont insérés dans ces trames en imposant une contrainte en priorité aux bits appartenant à la classe de bits la moins sensible.

18. Vocodeur paramétrique selon l'une quelconque des revendications 13 à 17, comprenant en outre des moyens pour soumettre le flux d'information secondaire à un codage correcteur d'erreurs avant insertion dans le flux d'information principal.

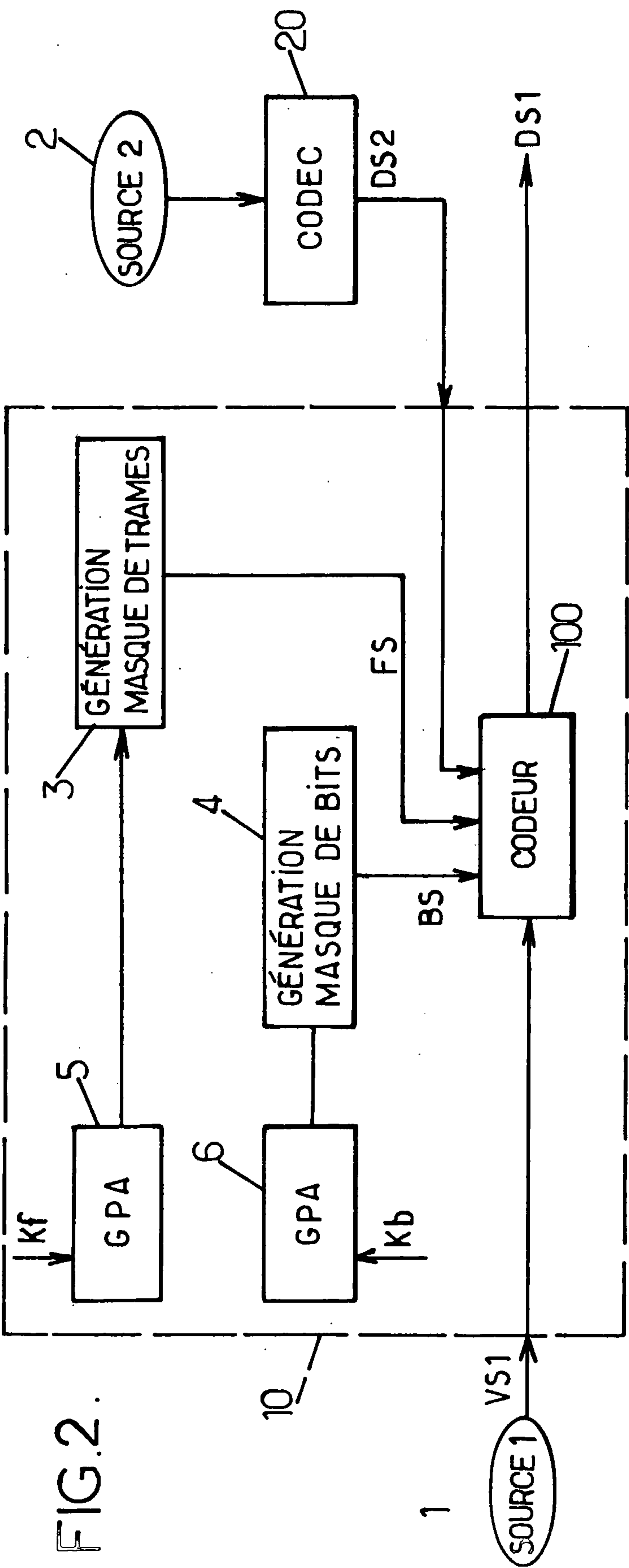
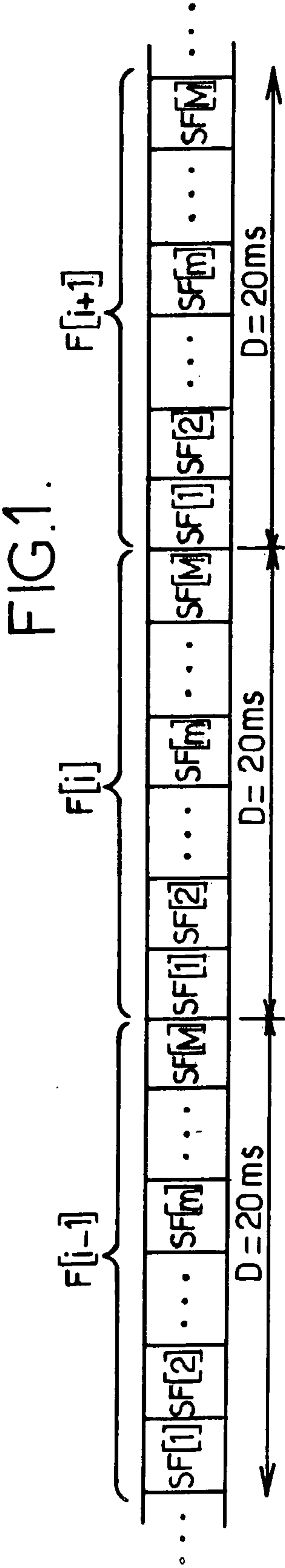
5

19. Vocodeur paramétrique selon l'une quelconque des revendications 13 à 18, dans lequel les moyens d'insertion sont adaptés pour insérer des bits du flux d'information secondaire en imposant des valeurs à des bits qui appartiennent à des paramètres d'excitation d'un filtre du modèle de source du
10 vocodeur paramétrique.

20. Vocodeur paramétrique selon l'une quelconque des revendications 13 à 19, dans lequel les moyens d'insertion sont adaptés pour insérer des bits du flux d'information secondaire dans des trames de silence du flux d'information
15 principal.

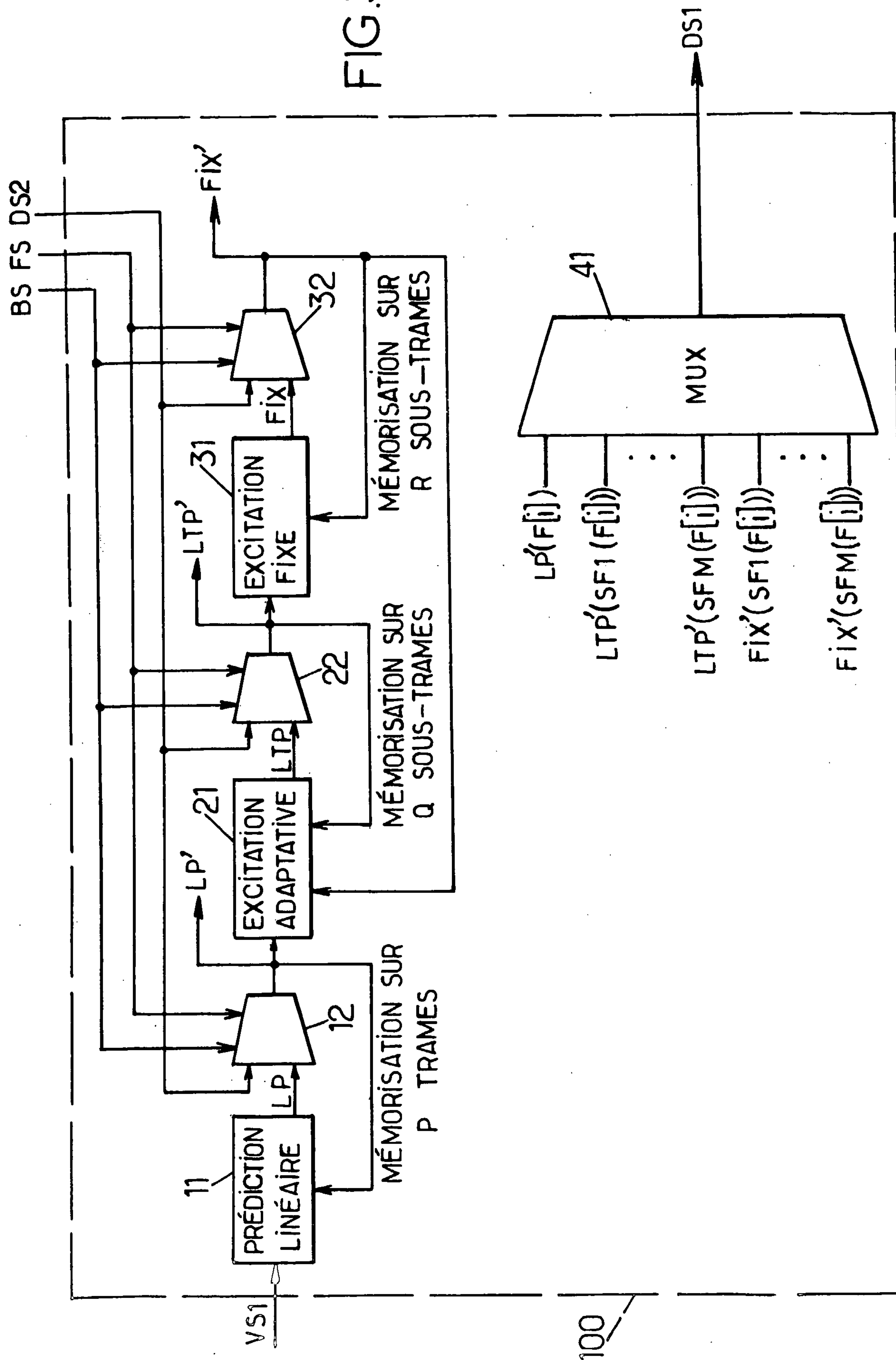
21. Vocodeur paramétrique selon l'une quelconque des revendications 13 à 20, dans lequel les moyens d'insertion sont adaptés pour insérer des bits du flux d'information secondaire en imposant des contraintes à des bits de bout en
20 bout au titre du flux d'information principal.

22. Equipement terminal d'un système de radiocommunications comprenant un vocodeur paramétrique selon l'une quelconque des revendications 13 à 21.



2/3

FIG. 3.



3/3

