



UNION INTERNATIONALE DES TÉLÉCOMMUNICATIONS

UIT-T

SECTEUR DE LA NORMALISATION
DES TÉLÉCOMMUNICATIONS
DE L'UIT

P.83

(03/93)

**QUALITÉ DE LA TRANSMISSION TÉLÉPHONIQUE
ESSAIS SUBJECTIFS D'OPINION**

**ÉVALUATION SUBJECTIVE DE LA QUALITÉ
DES CODECS NUMÉRIQUES À BANDE
TÉLÉPHONIQUE ET À LARGE BANDE**

Recommandation UIT-T P.83

(Antérieurement «Recommandation du CCITT»)

AVANT-PROPOS

L'UIT-T (Secteur de la normalisation des télécommunications) est un organe permanent de l'Union internationale des télécommunications (UIT). Il est chargé de l'étude des questions techniques, d'exploitation et de tarification, et émet à ce sujet des Recommandations en vue de la normalisation des télécommunications à l'échelle mondiale.

La Conférence mondiale de normalisation des télécommunications (CMNT), qui se réunit tous les quatre ans, détermine les thèmes que les Commissions d'études de l'UIT-T doivent examiner et à propos desquels elles doivent émettre des Recommandations.

La Recommandation UIT-T P.83, élaborée par la Commission d'études XII (1988-1993) de l'UIT-T, a été approuvée par la CMNT (Helsinki, 1-12 mars 1993).

NOTES

1 Suite au processus de réforme entrepris au sein de l'Union internationale des télécommunications (UIT), le CCITT n'existe plus depuis le 28 février 1993. Il est remplacé par le Secteur de la normalisation des télécommunications de l'UIT (UIT-T) créé le 1^{er} mars 1993. De même, le CCIR et l'IFRB ont été remplacés par le Secteur des radiocommunications.

Afin de ne pas retarder la publication de la présente Recommandation, aucun changement n'a été apporté aux mentions contenant les sigles CCITT, CCIR et IFRB ou aux entités qui leur sont associées, comme «Assemblée plénière», «Secrétariat», etc. Les futures éditions de la présente Recommandation adopteront la terminologie appropriée reflétant la nouvelle structure de l'UIT.

2 Dans la présente Recommandation, le terme «Administration» désigne indifféremment une administration de télécommunication ou une exploitation reconnue.

© UIT 1994

Droits de reproduction réservés. Aucune partie de cette publication ne peut être reproduite ni utilisée sous quelque forme que ce soit et par aucun procédé, électronique ou mécanique, y compris la photocopie et les microfilms, sans l'accord écrit de l'UIT.

TABLE DES MATIÈRES

	<i>Page</i>
1 Introduction	1
2 Enregistrements de la source	2
2.1 Système d'émission	2
2.2 Echantillons de parole	2
3 Choix des paramètres d'expérience	2
3.1 Conditions requises pour les codecs	2
3.2 Conditions de référence	7
4 Conception des essais	9
5 Procédure des essais d'écoute	9
5.1 Système de réception	10
5.2 Echelles d'appréciation subjective	10
5.3 Bruit électrique	10
6 Analyse des résultats	10
Annexe A – Comparaison des différentes définitions du SNR	12
Annexe B – Ensemble de paramètres permettant de déterminer la qualité de fonctionnement des codecs	13
B.1 Systèmes à bande étroite (300-3400 Hz)	13
B.2 Systèmes à large bande (100-7000 Hz)	13
Références	14
Bibliographie	14

ÉVALUATION SUBJECTIVE DE LA QUALITÉ DES CODECS NUMÉRIQUES À BANDE TÉLÉPHONIQUE ET À LARGE BANDE

(Helsinki, 1993)

1 Introduction

La téléphonie numérique s'intègre maintenant de plus en plus rapidement dans le réseau téléphonique public avec commutation (RTPC). Depuis les années 60, la transmission numérique dans le RTPC a été synonyme de modulation par impulsions et codage (MIC) de loi A à 64 kbit/s (voir la Recommandation G.711). Maintenant, compte tenu d'une meilleure connaissance des techniques de traitement des signaux combinée aux progrès de la technologie, notamment des techniques LSI/VLSI, on s'intéresse davantage à des méthodes de codage plus efficaces que la MIC. Un exemple de cette évolution est l'introduction de la MICDA à 32 kbit/s (voir la Recommandation G.726). Plus récemment, un codec à large bande de 7 kHz et à 64 kbit/s a été normalisé par le CCITT (voir la Recommandation G.722). On accorde actuellement une attention considérable au codage à faible débit binaire (LRE) (*low rate encoding*), c'est-à-dire à des débits de 16 kbit/s et/ou moins.

Lorsque le trajet de transmission est numérique et/ou non linéaire, de simples mesures objectives telles que celles de la Recommandation G.712 sont insuffisantes pour assurer une qualité de transmission adéquate. Le but d'une méthode d'essai subjective est de mesurer la dégradation à laquelle la partie non linéaire du trajet de transmission contribue, afin que la qualité de l'ensemble du système soit satisfaisante. A cet effet, les mesures doivent être:

- a) fiables; et
- b) effectuées de telle sorte qu'elles tiennent compte des principales interactions entre la partie non linéaire et les autres parties du système de transmission.

Cela implique la possibilité d'imputer sans ambiguïté une contribution numérique à chaque processus numérique et d'utiliser celle-ci conjointement avec d'autres dégradations pour estimer la qualité des connexions téléphoniques.

La présente Recommandation, fondée sur l'Annexe B/P.80 et l'expérience acquise à la suite de plusieurs études internationales [2, 3, 4, 5, 6], définit une méthode d'essai spécifique pour évaluer les processus numériques de telle sorte que les effets de distorsion de quantification de ces processus sur la qualité de transmission puissent être pris en considération dans le réseau téléphonique international en cours d'évolution.

Diverses méthodes sont possibles pour caractériser la qualité des processus numériques. Actuellement, ces méthodes comprennent des essais d'écoute seulement qui impliquent:

- 1) des notes d'opinion (par catégories);
- 2) des comparaisons doubles ou multiples; et
- 3) des tests d'articulation.

Pour la plupart des applications, le CCITT recommande d'utiliser la méthode d'évaluation par catégories absolues (ACR) (*absolute category rating*) fondée sur l'échelle de qualité d'écoute. Cependant, il y a des cas où d'autres échelles et méthodes d'évaluation sont plus appropriées et mieux adaptées; celles-ci sont utilisées également dans la présente Recommandation. Seuls seront signalés ci-après les cas où, exceptionnellement, la méthode ACR fondée sur l'échelle de qualité d'écoute n'est pas utilisée.

NOTE – La méthode d'évaluation par catégories de dégradation (DCR) (*degradation category rating*) [1] est décrite en détail dans l'Annexe D/P.80. Cette méthode est supposée convenir lorsque les dégradations numériques sont faibles. Elle peut donc être particulièrement utile pour l'optimisation des systèmes. Il faut noter en outre qu'une méthode de seuil pour comparaison directe, décrite en détail dans l'Annexe E/P.80, est elle aussi applicable pour l'optimisation des systèmes.

La présente Recommandation contient des indications sur la façon d'évaluer la qualité des codecs numériques. Elle doit être lue conjointement avec la Recommandation P.80.

Il est nécessaire en outre de compléter les essais d'écoute seulement par des essais de conversation mais aucune méthode n'a encore été adoptée pour inclusion dans la présente Recommandation.

2 Enregistrements de la source

Cet article est le même que B.1/P.80, avec les exceptions indiquées ci-après.

2.1 Système d'émission

Pour les systèmes à bande étroite (300-3400 Hz), il est recommandé, si une comparaison des résultats est nécessaire entre des laboratoires et des Administrations, que le côté émission du système de référence intermédiaire (IRS) (*intermediate reference system*) soit conforme à la Recommandation P.48 et soit étalonné selon la Recommandation P.64. Cependant, on a constaté que la qualité des codecs vocaux à faible débit binaire pouvait dépendre sensiblement de la caractéristique de fréquence appliquée au signal vocal d'entrée.

NOTE – L'IRS est représentatif des appareils téléphoniques analogiques (années 1970) et la Commission d'études 12 étudie actuellement un «téléphone de référence» représentatif d'appareils téléphoniques modernes (voir la Figure 1/P.31).

Si l'on se propose d'utiliser un système à large bande (100-7000 Hz) pour audioconférence, le côté émission doit être conforme à la Publication 581-7 de la CEI.

2.2 Echantillons de parole

Il est recommandé d'utiliser au minimum deux phrases par échantillon, ce qui permet, par exemple, d'évaluer des taux d'erreur binaire (BER) (*bit-error-rate*) de 0, 1:10000, 1:1000 et 1:100 (voir 3.1.4). Les BER compris entre 0 et 1:10000 exigeront un plus grand nombre de phrases par échantillon.

3 Choix des paramètres d'expérience

3.1 Conditions requises pour les codecs

3.1.1 Niveaux d'entrée de parole

Lors de l'évaluation des codecs, on règle généralement les niveaux d'entrée par rapport au point de surcharge du codec. Il est recommandé d'utiliser des niveaux d'entrée de 14, 26 et 38 dB (équivalant à -8, -20, -32 dBm0) au-dessous du point de surcharge des codecs. Ces niveaux correspondent approximativement à la valeur moyenne mesurée au point de commutation international \pm deux écarts types. D'autres applications peuvent exiger des niveaux d'entrée différents. La Figure 1 montre l'effet des variations du niveau d'entrée vocal.

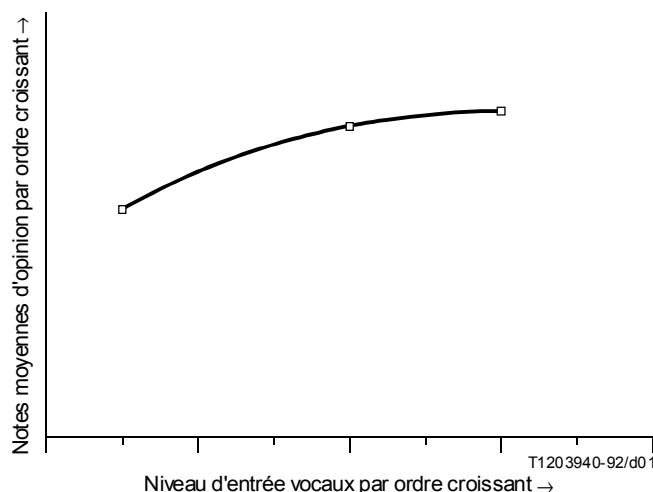


FIGURE 1/P.83

Notes moyennes d'opinion pour les niveaux d'entrée vocaux

3.1.2 Niveaux d'écoute

Il est recommandé d'utiliser au moins trois niveaux d'écoute. Ces niveaux doivent se situer à 10 dB de part et d'autre du niveau d'écoute préféré (voir B.2.1/P.80 pour obtenir des informations sur le niveau d'écoute préféré). La Figure 2 montre l'effet des variations du niveau d'écoute.

3.1.3 Locuteurs

3.1.3.1 Locuteurs différents

Il est recommandé d'utiliser au minimum deux hommes et deux femmes. La Figure 2 montre l'effet de l'utilisation de locuteurs différents. Cependant, si l'on veut tester les variations en fonction des locuteurs, il est recommandé d'utiliser les locuteurs suivants:

- 8 hommes;
- 8 femmes;
- 8 enfants.

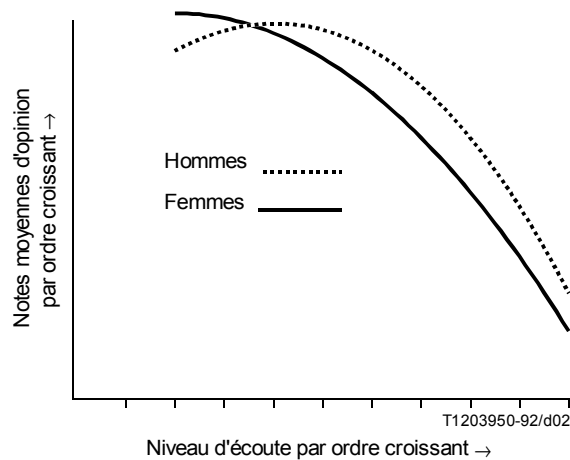


FIGURE 2/P.83

Notes moyennes d'opinion pour les niveaux d'écoute

3.1.3.2 Locuteurs multiples

Il est important de connaître la sensibilité du codec à de multiples signaux d'entrée vocaux et de déterminer s'il se produit éventuellement des effets défavorables (par exemple, «interruption» de transmission, signaux parasites, etc.).

Si l'on ne veut prendre en considération que l'utilisation du combiné, une combinaison de deux locuteurs différents avec des niveaux vocaux d'entrée séparés de 20 dB est généralement suffisante.

Cependant, si l'on veut utiliser le codec dans le mode conférence, c'est-à-dire «mains-libres», avec plusieurs locuteurs pouvant parler en même temps, il faut s'assurer que l'algorithme de codage sera à même de traiter le cas de plusieurs locuteurs dont les différences de niveau vocal pourraient être nulles.

Il est recommandé d'utiliser la méthode d'évaluation par catégories de dégradation (DCR) fondée sur l'échelle à 5 points du paragraphe 5.2.3 (voir également l'Annexe D/P.80) ou la méthode de détection de la réponse discontinue alternative fondée sur l'échelle à 3 points du paragraphe 5.2.4 (voir également l'Annexe C/P.80) pour évaluer les effets de locuteurs multiples.

3.1.4 Erreurs

Elles dépendront de l'application, c'est-à-dire qu'elles varieront si celle-ci est utilisée dans le RTPC ordinaire (systèmes de lignes) ou, par exemple, dans un environnement radioélectrique tel que les radiocommunications mobiles.

Il est recommandé, pour les systèmes de lignes, d'utiliser des erreurs à distribution aléatoire avec des taux d'erreur binaire (BER) compris entre 0 et 1:1000. Naturellement, cela dépendra du nombre de phrases utilisé dans un échantillon (voir 2.2). Dans certaines circonstances, il peut être nécessaire de descendre jusqu'à un taux de 1:100.

Pour d'autres applications telles que les radiocommunications mobiles, les erreurs peuvent très bien être du type à salves d'erreurs et il convient donc d'utiliser des erreurs de ce type.

On utilise généralement la méthode ACR fondée sur l'échelle de qualité d'écoute du paragraphe 5.2.1. Si, cependant, on s'attend à des conditions médiocres, la méthode ACR fondée sur l'échelle d'effort d'écoute du paragraphe 5.2.2 peut être plus appropriée.

3.1.5 Débits binaires

Le codec doit être testé à tous les débits binaires auxquels il est capable de fonctionner (par exemple pour la Recommandation G.722, ces débits sont de 48 kbit/s, 56 kbit/s et 64 kbit/s et pour la Recommandation G.726, ils sont de 16 kbit/s, 24 kbit/s, 32 kbit/s et 40 kbit/s). Cependant, le fonctionnement à certains débits binaires peut dépendre des conditions d'exploitation et la charge du système peut entrer en jeu (voir la Recommandation P.84). La Figure 3 montre l'effet des variations du débit binaire.

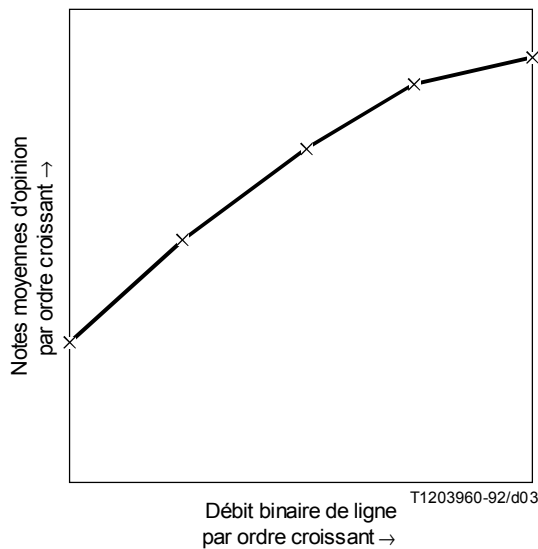


FIGURE 3/P.83

Notes moyennes d'opinion à différents débits binaires de ligne

3.1.6 Transcodages

On évalue généralement un codec en association avec un processus de compression-extension de loi A ou μ . Un seul codec analogique-analogique aura donc la configuration indiquée sur la Figure 4.

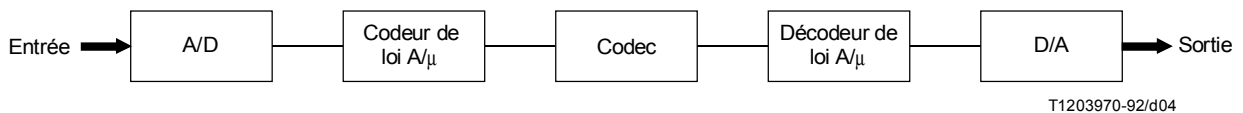


FIGURE 4/P.83

Système de transcodage

3.1.6.1 Mise en cascade asynchrone

Cette méthode reflète l'utilisation du codec dans un réseau fondé sur une architecture à 64 kbit/s. Chaque transcodage comporte donc une distorsion de quantification supplémentaire produite par le processus de compression-extension linéaire/loi A ou μ , une distorsion d'affaiblissement supplémentaire résultant des tolérances dans les filtres anti-distorsion de repliement et de reconstruction, ainsi que le bruit accumulé dans la voie inactive. Cette configuration est illustrée sur la Figure 5.

Le CCITT recommande de tester 1, 2 et, au moins, 3 codecs en cascade.

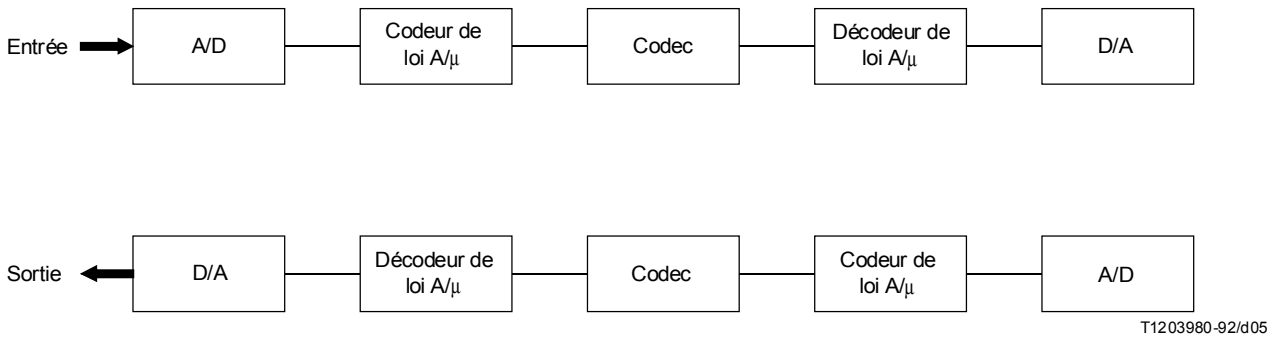


FIGURE 5/P.83
Mise en cascade asynchrone

3.1.6.2 Mise en cascade synchrone

La configuration de circuit est représentée sur la Figure 6; elle a l'avantage de ne nécessiter qu'un processus de compression-extension de loi A ou μ et présente donc une distorsion de quantification réduite comparée à celle du paragraphe 3.1.6.1.

Le CCITT recommande de tester 1, 2 et, au moins, 3 codecs en cascade.

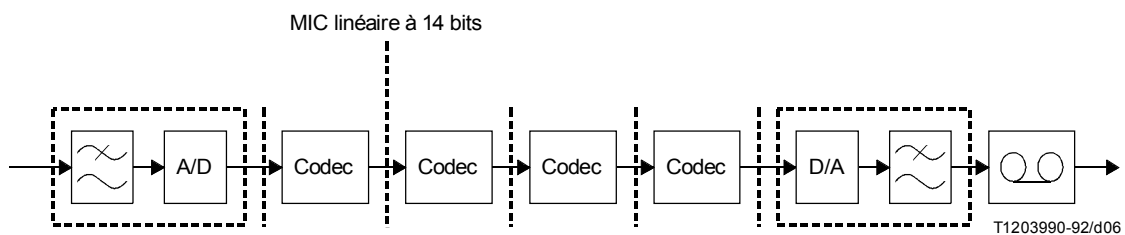
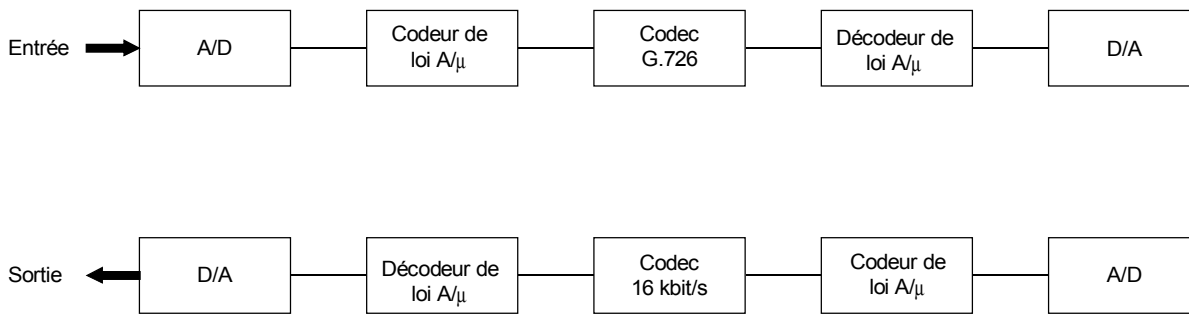


FIGURE 6/P.83
Mise en cascade synchrone

3.1.6.3 Mise en cascade mixte

Il est important de déterminer les effets de systèmes en cascade qui utilisent un codage à différents débits binaires et une configuration de ce type est illustrée sur la Figure 7. Il est essentiel de tester les combinaisons les plus probables.



T1204000-92/d07

FIGURE 7/P.83
Mise en cascade mixte

L'exemple suivant illustre les combinaisons possibles:

8-16 kbit/s →	64 kbit/s →	32 kbit/s →	64 kbit/s →	16-32 kbit/s →
Radiocommunications mobiles, annonces enregistrées et appli- cations aéronautiques	Circuit interurbain	DCME à débit binaire variable	Circuit interurbain	Téléphone sans cordon

Pour les systèmes à large bande (par exemple, G.722), il est nécessaire de tester un système à large bande en cascade avec un système à bande étroite (par exemple, système G.722 en cascade avec système G.726 et vice versa).

3.1.7 Défauts d'adaptation

Il est recommandé, pour les systèmes à large bande (G.722), que les effets dus à un problème de non-adaptation, c'est-à-dire à différents modes de fonctionnement entre l'émetteur et le récepteur, soient étudiés en fonction des conditions suivantes:

<i>Débit binaire de l'émetteur</i>		<i>Débit binaire du récepteur</i>
56 kbit/s	→	64 kbit/s
48 kbit/s	→	56 kbit/s
48 kbit/s	→	64 kbit/s

3.1.8 Bruit d'origine externe (émission)

Comme pour les considérations du paragraphe 3.1.3.2 relatives aux locuteurs multiples, l'effet d'interaction du bruit d'origine externe est un facteur important pour les raisons suivantes:

- les codecs à faible débit binaire peuvent utiliser des processus de synthèse vocale ainsi qu'un codage de forme d'onde classique; et
- l'application d'un codec pourrait consister à l'utiliser dans un environnement bruyant, par exemple un véhicule en mouvement pour les radiocommunications mobiles ou un bureau bruyant.

Un nombre suffisant d'essais doit être effectué, avec le bruit approprié (voir A.1.1.2.2/P.80) pour vérifier cet effet et les rapports signal/bruit suivants (voir 2.2.3 pour les définitions) sont recommandés:

- 30 dB pour le bruit de salle;
- 10 dB et 20 dB pour le bruit à l'intérieur d'un véhicule.

Comme la qualité perçue est généralement un peu inférieure à la normale, le CCITT recommande d'utiliser la méthode ACR avec échelle d'effort d'écoute (voir 5.2.2) fondée sur l'aptitude à comprendre le sens des phrases.

NOTES

1 Le bruit peut être combiné électriquement avec les bandes magnétiques de la source de telle sorte que le niveau du bruit et de la parole soit connu de manière précise. Il n'est pas recommandé que les enregistrements de la source soient effectués dans un environnement bruyant.

2 Il ne faut pas supposer automatiquement que la combinaison de signaux vocaux enregistrés séparément et d'un bruit d'origine élevé produira les mêmes effets qu'un sujet parlant dans un environnement bruyant. Cela tient au fait que les locuteurs adaptent leurs caractéristiques vocales (niveau et spectre) ainsi que leur comportement en matière de parole dans un environnement bruyant. Etant donné que certains codecs peuvent effectuer un traitement différent selon les parties du spectre audiofréquence considérées, cette modification des caractéristiques vocales peut influencer les résultats.

3.1.9 Signaux d'information de réseau

Dans tout réseau national, un grand nombre de signaux ou de tonalités d'information sont adressés à l'utilisateur, d'autres constituant des instructions pour les équipements. Ces signaux peuvent provenir de l'intérieur du RTPC ou de réseaux privés qui lui sont rattachés. Il est important que, lorsque ces signaux passent par un dispositif quelconque de traitement des signaux vocaux, leur dégradation ne les rende pas méconnaissables pour le client ni pour l'équipement conçu pour prendre des mesures à la réception. Il est possible que cette dernière situation (par exemple, tonalités DTMF), tolère moins bien une dégradation que la première, mais elle offre l'avantage que de simples essais objectifs devraient être suffisants pour détecter une limite admissible de cette dégradation.

Il est recommandé de tester les tonalités de signalisation conformes à la Recommandation Q.35 émises par le réseau et au minimum les tonalités suivantes:

- tonalité d'invitation à numéroté;
- tonalité de retour d'appel de l'abonné demandé;
- tonalité d'occupation de l'abonné demandé;
- tonalité d'occupation de l'équipement;
- tonalité de numéro impossible à obtenir.

Il est recommandé d'utiliser la méthode d'évaluation par catégories de dégradation (DCR) fondée sur l'échelle à 5 points du paragraphe 5.2.3 (voir également l'Annexe D/P.80) pour déterminer si les signaux d'information sont appropriés (reconnaissables).

3.1.10 Musique

Certains équipements utilisés dans le RTPC, notamment les commutateurs privés, ont la possibilité de transmettre de la musique au client en cas de «mise en attente».

Le CCITT recommande de n'utiliser que des essais simples pour s'assurer que la musique est d'une qualité raisonnable.

3.2 Conditions de référence

3.2.1 Références directes

Il s'agit des conditions les meilleures que l'on puisse obtenir lors de l'expérience; elles correspondent à un rapport Q_N ou Q_W infini (voir 3.2.2).

3.2.2 Appareil de référence à bruit modulé (MNRU)

Il convient d'utiliser comme système de référence un bruit aléatoire avec une amplitude proportionnelle à l'amplitude du signal instantanée en termes de rapport Q_N ou Q_W , conformément à l'appareil de référence à bruit modulé (MNRU) (*modulated noise reference unit*) spécifié dans la Recommandation P.81; la qualité subjective des processus numériques devra être exprimée en fonction de ce système de référence pour les raisons suivantes:

- a) pour les codecs à forme d'onde et éventuellement certains codecs autres qu'à forme d'onde, le signal traité par le MNRU est, du point de vue perceptif, sensiblement identique aux signaux traités par ces codecs, ce qui facilite, en principe, l'évaluation par les sujets participant aux essais;
- b) l'expérience a montré que le MNRU est un étalon de transfert utile qui permet d'effectuer des comparaisons significatives entre différents laboratoires; et
- c) une expérience et des informations considérables ont été accumulées avec le MNRU.

La gamme de bruit pour les systèmes à bande étroite doit aller de $Q_N = 5$ à 35 dB (de préférence 5 à 7 valeurs différentes) et pour les systèmes à large bande de $Q_W = 10$ à 45.

La Figure 8 montre l'effet des variations de Q_N ou Q_W .

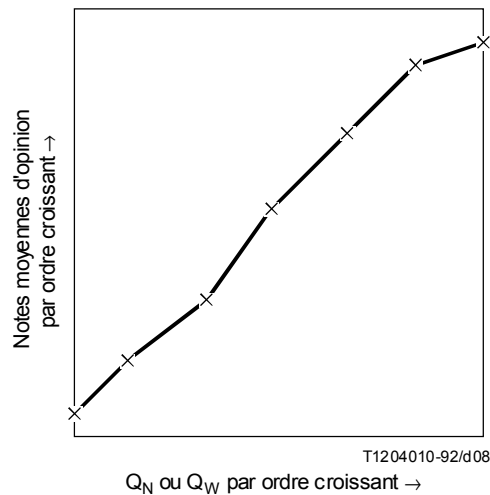


FIGURE 8/P.83

Notes moyennes d'opinion pour Q_N ou Q_W

3.2.3 Rapport signal/bruit (SNR)

Les Administrations et les compagnies d'exploitation ont jugé utile, dans le passé, d'évaluer les effets des dégradations en termes de SNR.

Le CCITT recommande d'utiliser les définitions suivantes du rapport signal/bruit, mesuré sur les connexions, en ce qui concerne le bruit en régime permanent:

- *Mesures téléphoniques (bande étroite) – Valeurs psophométriques* (voir la Recommandation O.41)
SNR(p) = Niveau de parole active (P.56)/mesure de bruit psophométrique pondérée
- *Mesures à large bande – Valeurs pondérées A* (voir la Recommandation P.54)
SNR(A) = Niveau de parole active (voir la Recommandation P.56)/mesure de bruit pondérée A
Si des définitions, autres que celles indiquées ci-dessus, sont utilisées, il convient d'adopter le système de notation suivant:
- *Bande étroite 300-3400 Hz – Valeurs non pondérées*
SNR(N) = Niveau de parole active (voir la Recommandation P.56)/mesure de bruit non pondérée
- *Large bande 100-7000 Hz – Valeurs non pondérées*
SNR(W) = Niveau de parole active (voir la Recommandation P.56)/mesure de bruit non pondérée
- *Pondération pour message C* (voir la Recommandation O.41)
SNR(C) = Niveau de parole active (voir la Recommandation P.56)/mesure de bruit pondérée pour message C
- *Largeur de bande 100-5000 Hz*
SNR(0,1-5) = Niveau de parole active (voir la Recommandation P.56)/mesure de bruit non pondérée
NOTE – Si d'autres largeurs de bande sont utilisées, il convient de remplacer 0,1-5 par la largeur de bande de mesure.

On trouvera en Annexe A une comparaison des SNR établie à l'aide des définitions ci-dessus pour différents types de spectres de bruit et différentes caractéristiques d'efficacité en fonction de la fréquence à l'émission.

Si on utilise un bruit gaussien lors de l'essai, il est suggéré que, pour les systèmes à bande étroite et à large bande, la gamme de bruit soit comprise entre 15 et 45 dB.

La Figure 9 montre l'effet des variations du rapport SNR.

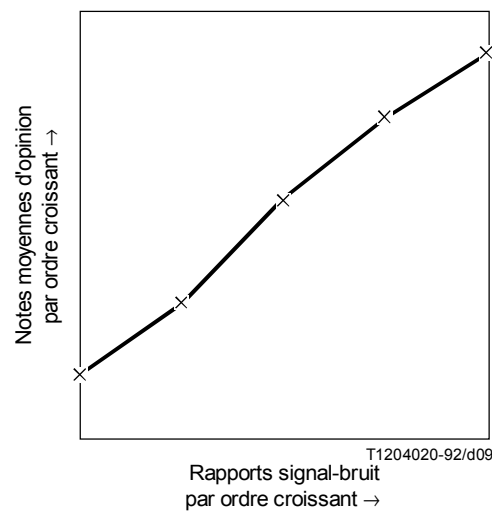


FIGURE 9/P.83

Notes moyennes d'opinion pour le SNR

3.2.4 Codecs de référence

Les codecs de référence peuvent servir à deux usages:

- 1) on peut les utiliser pour déterminer les règles de planification (voir la Recommandation G.113); et
- 2) on peut les utiliser comme norme pour évaluer la qualité globale en termes de paramètres (voir 3.1).

S'ils servent à déterminer les règles de planification, le CCITT recommande d'utiliser la mise en cascade asynchrone (voir 3.1.6.1) de 1, 2, 4, 8 et 16 codecs de loi A ou μ conformes aux Recommandations G.711/G.712.

S'ils servent à déterminer la qualité relative, par exemple par rapport aux Recommandations G.726 ou G.722, les considérations formulées en 3.1.6.1 et 3.1.6.2 s'appliquent.

NOTE – La spécification relative à un codec de référence G.711 est actuellement étudiée.

4 Conception des essais

Les considérations formulées en détail dans B.3/P.80 s'appliquent.

L'utilisation de chaque combinaison des paramètres décrits à l'article 2 conduirait à effectuer une seule expérience qui serait logistiquement trop importante. Il est recommandé de mener un ensemble minimal d'expériences qui, même s'il ne couvre pas chaque combinaison, fournira des données suffisantes pour prendre des décisions en connaissance de cause. L'Annexe B indique des ensembles d'expériences jugés appropriés par le CCITT à la suite des études de codecs à bande étroite et à large bande qui ont conduit à l'élaboration des Recommandations de la série G.700.

Il convient d'être extrêmement prudent lorsqu'on compare, lors d'un même essai, des systèmes présentant des dégradations très différentes (par exemple, codecs numériques, systèmes de multiplexage par répartition en fréquence, vocodeurs, etc.).

5 Procédure des essais d'écoute

Cet article est identique à celui de B.4/P.80, avec les exceptions suivantes.

5.1 Système de réception

Pour les systèmes à bande étroite (300-3400 Hz), il est recommandé que le côté réception du système de référence intermédiaire (IRS) soit conforme à la Recommandation P.48 et soit étalonné selon la Recommandation P.64.

Si l'on se propose d'utiliser un système à large bande (100-7000 Hz) pour audioconférence, le côté réception doit être conforme à la Publication 581-7 de la CEI.

5.2 Echelles d'appréciation subjective

La méthode à utiliser est l'essai de type évaluation par catégories absolues (ACR) fondée sur une échelle à 5 points décrite dans B.4.5/P.80.

Il est recommandé d'utiliser les échelles d'appréciation subjective suivantes pour évaluer les processus numériques.

5.2.1 Echelle de qualité d'écoute

Qualité de la parole:

Excellent	5
Bonne	4
Assez bonne	3
Médiocre	2
Mauvaise	1

5.2.2 Echelle d'effort d'écoute

Effet nécessaire pour comprendre la signification des phrases:

Détente absolue; aucun effort	5
Attention nécessaire; pas d'effort appréciable	4
Effort modéré	3
Effort considérable	2
Signification incompréhensible en dépit de tous les efforts possibles	1

5.2.3 Echelle de catégories de dégradation

Dégradation inaudible	5
Dégradation audible mais pas gênante	4
Dégradation un peu gênante	3
Dégradation gênante	2
Dégradation très gênante	1

5.2.4 Echelle d'appréciation subjective de possibilité de détection

Génant	3
Décelable	2
Indécelable	1

5.3 Bruit électrique

Il convient d'ajouter un bruit gaussien équivalant à -68 dBmp à l'entrée du système de réception pour réduire les effets de contraste de bruit au moment où les premières phrases sont prononcées.

6 Analyse des résultats

Les considérations formulées en détail dans B.4.7/P.80 s'appliquent.

L'un des objectifs de l'analyse est de déterminer une fonction $Q_2 = F(L)$ où Q_2 est la valeur Q pour le codec et L le débit binaire de ligne. Une méthode simple utilise, pour déterminer cette fonction, les valeurs de MOS (note moyenne d'opinion) indiquées sur les Figures 3 et 8 et permet d'établir un graphique de cette fonction, comme indiqué sur la Figure 10. Cette méthode est décrite sur la Figure 11, où on choisit une valeur de débit binaire de ligne, par exemple L_2 , et où on détermine sa valeur MOS correspondante. On transfère ensuite cette valeur MOS dans le graphique de droite

pour trouver la valeur de Q, en l'occurrence Q_2 , correspondant à cette valeur MOS. Les valeurs de Q pour toutes les autres valeurs L sont obtenues d'une manière similaire et l'ensemble de gains (L_i, Q_i) qui en résulte est tracé comme indiqué sur la Figure 10.

NOTE – L'ajustement de courbe appliqué aux résultats MOS en fonction de Q est actuellement étudié.

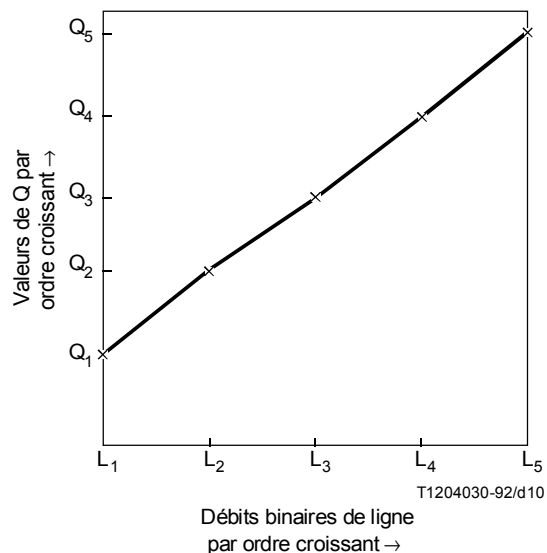


FIGURE 10/P.83

Valeur de Q en fonction du débit binaire de ligne

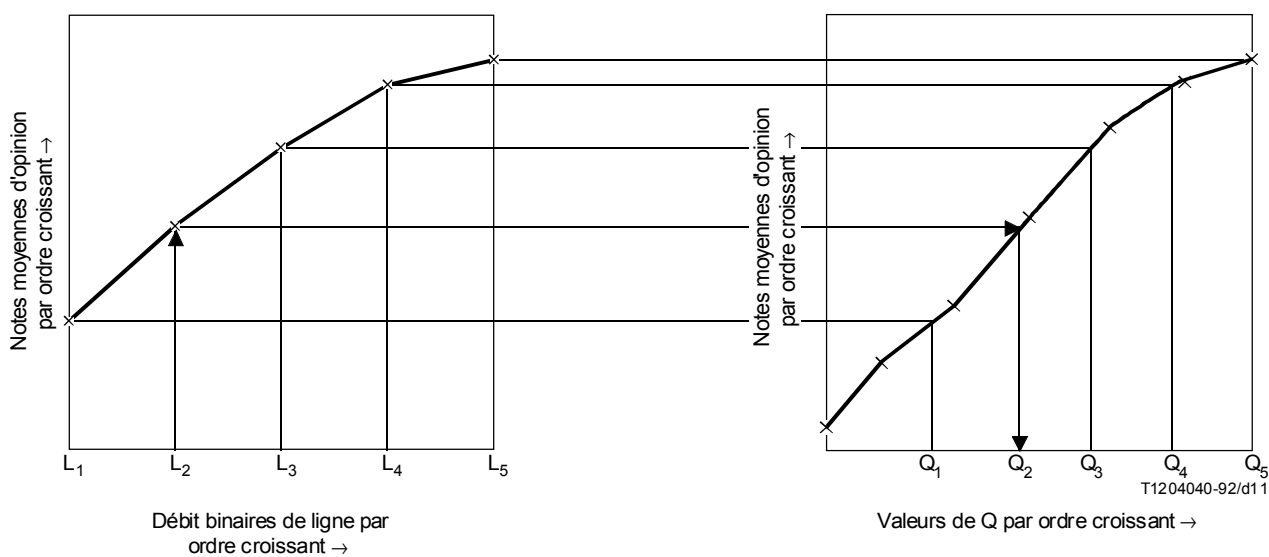


FIGURE 11/P.83

Méthode graphique permettant d'obtenir la Figure 10 à partir des Figures 3 et 8

Annexe A

Comparaison des différentes définitions du SNR

(Cette annexe fait partie intégrante de la présente Recommandation)

Le Tableau A.1 indique quelques calculs utilisant les pondérations décrites dans la Recommandation O.41 et s'appliquant à des spectres connus avec deux réponses en fréquence différentes. Ces spectres ont été choisis de manière à représenter ceux qui sont le plus généralement utilisés lors d'essais subjectifs.

- Bruit blanc (gaussien) compatible avec le bruit de circuit – psophométrique et pour message C, conçu pour mesurer les effets de ce type de bruit.
- Bruit de Hoth (de salle) et bruit à l'intérieur d'un véhicule utilisés comme bruits d'origine externe lors d'essais subjectifs (voir A.1.1.2.2/P.80) et représentant le ou les bruits, captés par le microphone et mesurés aux bornes de la ligne téléphonique.

TABLEAU A.1/P.83

		Bruit blanc	Bruit de Hoth	Véhicule en mouvement	Véhicule à l'arrêt
Réponse uniforme (bande étroite)	Non pondérée	0	0	0	0
	Message C	-1,9	-3,6	-6,4	-4,2
	Psophométrique	-2,6	-2,6	-4,1	-2,8
Réponse IRS (bande étroite)	Non pondérée	0	0	0	0
	Message C	-1,9	-2,0	-3,7	-2,3
	Psophométrique	-3,6	-2,3	-2,4	-2,3
Réponse uniforme (large bande)	Non pondérée	0	0	0	0
	Message C	-5,2	-7,3	-16,5	-12,7
	Psophométrique	-5,9	-6,3	-14,0	-11,3
Réponse IRS (large bande)	Non pondérée	0	0	0	0
	Message C	-1,9	-2,1	-4,4	-2,5
	Psophométrique	-3,6	-2,4	-3,1	-2,5

NOTES

- Toutes les valeurs sont des différences en dB par rapport à la valeur non pondérée.
- Le signe négatif signifie «moins bruyant».
- Pour calculer la valeur non pondérée à bande étroite, on a utilisé les bandes d'1/3 d'octave, centrées sur les fréquences préférées, comme indiqué dans l'ISO R.266, c'est-à-dire de 315 Hz à 3150 Hz inclusivement.
- Pour calculer la valeur non pondérée à large bande, on a utilisé les bandes d'1/3 d'octave, centrées sur les fréquences préférées, comme indiqué dans l'ISO R.266, c'est-à-dire de 100 Hz à 6300 Hz inclusivement.
- Pour calculer les pondérations message C et psophométriques, on a utilisé les bandes d'1/3 d'octave, centrées sur les fréquences préférées, comme indiqué dans l'ISO R.266, c'est-à-dire de 100 Hz à 5000 Hz inclusivement.
- Il convient d'être prudent lorsqu'on suppose que la pondération pour message C équivaut à une pondération psophométrique, ce qui n'est vrai que pour le bruit de type gaussien (voir la Recommandation O.41).

Annexe B

Ensemble de paramètres permettant de déterminer la qualité de fonctionnement des codecs

(Cette annexe fait partie intégrante de la présente Recommandation)

B.1 Systèmes à bande étroite (300-3400 Hz)

- a) *Expérience 1 – Effet des erreurs, du niveau d'entrée et du niveau d'écoute*
- | | |
|-------------------------------------|--|
| Niveau d'entrée vocal: | 3 (14, 26 et 38 dB au-dessous de la surcharge) |
| Niveau d'écoute: | 3 (P + 10, P, P – 10) |
| Erreurs: | 3 (0, 1:10000, 1:1000) |
| Transcodage: | 1 transcodage |
| Bruit d'origine externe (émission): | 1 (< 30 dBA) |
- b) *Expérience 2 – Effet des transcodages, du niveau d'entrée et du niveau d'écoute*
- | | |
|-------------------------------------|--|
| Niveau d'entrée vocal: | 3 (14, 26 et 38 dB au-dessous de la surcharge) |
| Niveau d'écoute: | 3 (P + 10, P, P – 10) |
| Transcodage: | «x» |
| Erreurs: | 1 (1:1000) |
| Bruit d'origine externe (émission): | 1 (< 30 dBA) |
- c) *Expérience 3 – Effet du bruit d'origine externe, du bruit de salle, du niveau d'entrée et du niveau d'écoute*
- | | |
|-------------------------------------|--|
| Niveau d'entrée vocal: | 3 (14, 26 et 38 dB au-dessous de la surcharge) |
| Niveau d'écoute: | 3 (P + 10, P, P – 10) |
| Transcodage: | 1 transcodage |
| Erreurs: | 1 (1:1000) |
| Bruit d'origine externe (émission): | 2 (< 30 dBA et «y») |
| Bruit de salle: | «z» |
- où «P» est le niveau d'écoute préféré;
«x» est le nombre de combinaisons de transcodage à tester;
«y» est le niveau de bruit (émission) à tester;
«z» est le nombre de conditions de bruit de salle (généralement 2).

Les trois expériences doivent toutes inclure également les conditions du MNRU à bande étroite.

Ces trois expériences ne sont pas nécessairement exhaustives et doivent être complétées par d'autres expériences pour mieux caractériser le codec à l'essai.

B.2 Systèmes à large bande (100-7000 Hz)

- a) *Expérience 1 – Effet du débit binaire, du BER, du niveau d'entrée et du niveau d'écoute*
- | | |
|-------------------------------------|--|
| Niveau d'entrée vocal: | 2 (20 et 38 dB au-dessous de la surcharge) |
| Niveau d'écoute: | 3 (P + 10, P, P – 10) |
| Transcodage: | 1 transcodage |
| Bruit d'origine externe (émission): | 1 (< 30 dBA) |
| Débits binaires: | 3 (48, 56 et 64 kbit/s) |
| BER: | 3 (0, 1:10000, 1:1000) |
| Bruit de salle: | 1 (< 30 dBA) |

- b) *Expérience 2 – Effet des transcodages, du niveau d'entrée et du niveau d'écoute*
- | | |
|-------------------------------------|--|
| Niveau d'entrée vocal: | 2 (20 et 38 dB au-dessous de la surcharge) |
| Niveau d'écoute: | 3 (P + 10, P, P – 10) |
| Transcodages: | «x», y compris synchrones et asynchrones |
| Bruit d'origine externe (émission): | 1 (< 30 dBA) |
| Débits binaires: | 3 (48, 56 et 64 kbit/s) |
| BER: | 3 (0, 1:10000 et 1:1000) |
| Bruit de salle: | 1 (< 30 dBA) |
- c) *Expérience 3 – Effet des défauts d'adaptation, du niveau d'entrée et du niveau d'écoute*
- | | |
|-------------------------------------|--|
| Niveau d'entrée vocal: | 2 (20 et 38 dB au-dessous de la surcharge) |
| Niveau d'écoute: | 3 (P + 10, P, P – 10) |
| Transcodage: | 1 transcodage |
| Bruit d'origine externe (émission): | 1 (< 30 dBA) |
| Débits binaires: | 3 (56→64, 48→56, 48→64 kbit/s) |
| BER: | 2 (0 et 1:1000) |
| Bruit de salle: | 1 (< 30 dBA) |
- où «P» est le niveau d'écoute préféré;
«x» est le nombre de combinaisons de transcodage à tester.

Les trois expériences doivent toutes inclure également les conditions du MNRU à large bande.

Références

- [1] COMBESCURE (P.) *et autres*: Quality evaluation of speech coded at 32 kbit/s by means of degradation category ratings, *Proc. ICASP 82 (International Conference on Acoustics, Speech and Signal Processing)*, Vol. 2, Paris, mai 1982.
- [2] CCITT Rapport de la réunion du Groupe de travail XVIII/2 (Traitement des signaux vocaux), COM XVIII-R 28, Annexe 1, p. 13 à 39, décembre 1983.
- [3] GOODMAN (D.J.), NASH (R.D.): Subjective quality of the same speech transmission conditions in seven different countries, *Proc. ICASP 82 (International Conference on Acoustics, Speech and Signal Processing)*, Vol. 2, Paris, mai 1982.
- [4] MODENA (G.), COLEMAN (A.), USAI (P.) et COVERDALE (P.): Subjective performance evaluation of the 7 kHz audio coder, *IEEE Global Telecommunications Conference 1986 (Globecom 86)*, Houston, Texas, 1-4 décembre 1986.
- [5] GROUPE MIXTE D'EXPERTS EUROPÉENS Méthode d'essai subjectif pour l'évaluation des codecs à faible débit binaire pour les radiocommunications mobiles, CCITT, COM XII-68, mai 1986.
- [6] ADMINISTRATION SUÉDOISE DES TÉLÉCOMMUNICATIONS Rapport sur les essais subjectifs de codecs susceptibles d'être utilisés pour les radiocommunications mobiles, CCITT, COM XII-147, février 1987.

Bibliographie

Recommandation du CCITT *Modulation par impulsions et codage différentiel adaptatif (MICDA) à 40, 32, 24 et 16 kbit/s*, tome III, Rec. G.726.

DAUMER (W.R.) et CAVANAUGH (J.R.): A subjective comparison of selected digital codecs for speech, *Bell Systems Technical Journal*, Vol. 57, n° 9, novembre 1978.

RICHARDS (D.L.) et BARNES (G.J.): Pay-off between quantizing distortion and injected circuit noise, *Proc. ICASP 82 (International Conference on Acoustics, Speech and Signal Processing)*, Vol. 2, Paris, mai 1982.

BOYD (I.), SOUTHCOTT (C.B.): A speech codec for the Skyphone service, *British Telecom Technology Journal*, Vol. 6, n° 2, avril 1988.

COLEMAN (A.), GLEISS (N.), USAI (P.): A Subjective Testing Methodology for Evaluating Medium Rate Codecs for Digital Mobile Applications, *Speech Communications*, Vol. 7, p. 151-166, juin 1988.

CROWE (D.P.): Selection of Voice Codec for the Aeronautical Satellite Service, *European Conference on Speech Communication and Technology*, Vol. 2, S37, p. 320-323, septembre 1989.

COLEMAN (A.), GLEISS (N.), SOTSCHECK (J.), USAI (P.), SCHEUERMANN (H.): Subjective performance evaluation of the RPE-LTP codec for the Pan-European cellular digital mobile radio system, *IEEE Globecom 89*, Dallas, Texas, 27-30 novembre 1989.

WHEDDON (C.) et LINGGARD (R.): *Speech and Language Processing*, Chapman and Hall, 1990.

Imprimé en Suisse

Genève, 1994