



UNION INTERNATIONALE DES TÉLÉCOMMUNICATIONS

UIT-T

SECTEUR DE LA NORMALISATION
DES TÉLÉCOMMUNICATIONS
DE L'UIT

P.330

(03/2003)

SÉRIE P: QUALITÉ DE TRANSMISSION
TÉLÉPHONIQUE, INSTALLATIONS TÉLÉPHONIQUES
ET RÉSEAUX LOCAUX

Lignes et postes d'abonnés

**Dispositifs de traitement vocal pour une
amélioration de l'acoustique**

Recommandation UIT-T P.330

RECOMMANDATIONS UIT-T DE LA SÉRIE P
QUALITÉ DE TRANSMISSION TÉLÉPHONIQUE, INSTALLATIONS TÉLÉPHONIQUES ET RÉSEAUX
LOCAUX

Vocabulaire et effets des paramètres de transmission sur l'opinion des usagers	Série	P.10
Lignes et postes d'abonnés	Série	P.30 P.300
Normes de transmission	Série	P.40
Appareils de mesures objectives	Série	P.50 P.500
Mesures électroacoustiques objectives	Série	P.60
Mesures de la sonie vocale	Série	P.70
Méthodes d'évaluation objective et subjective de la qualité	Série	P.80 P.800
Qualité audiovisuelle dans les services multimédias	Série	P.900

Pour plus de détails, voir la Liste des Recommandations de l'UIT-T.

Recommandation UIT-T P.330

Dispositifs de traitement vocal pour une amélioration de l'acoustique

Résumé

La présente Recommandation concerne les caractéristiques de transmission génériques et les principes de performance et d'essai des dispositifs de traitement vocal pour une amélioration de l'acoustique (SPDA, *speech processing device for acoustic enhancement*) destinés aux terminaux, indépendamment de l'application.

Un dispositif de traitement de la parole pour l'amélioration de l'acoustique est défini comme étant une fonction de traitement du signal intégrée dans les terminaux améliorant les signaux vocaux. Les fonctions d'amélioration de la voix sont notamment la gestion de la réduction de l'écho et du bruit acoustique.

La présente Recommandation a pour objet de définir un cadre permettant de spécifier les impératifs de performance des terminaux assurant la fonction SPDA et, le cas échéant, les essais qui peuvent être effectués sur de tels terminaux pour contrôler le respect de ces impératifs. Elle traite des caractéristiques génériques applicables tant aux terminaux analogiques que numériques.

Source

La Recommandation P.330 de l'UIT-T, élaborée par la Commission d'études 12 (2001-2004) de l'UIT-T, a été approuvée le 16 mars 2003 selon la procédure définie dans la Résolution 1 de l'AMNT.

AVANT-PROPOS

L'UIT (Union internationale des télécommunications) est une institution spécialisée des Nations Unies dans le domaine des télécommunications. L'UIT-T (Secteur de la normalisation des télécommunications) est un organe permanent de l'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.

L'Assemblée mondiale de normalisation des télécommunications (AMNT), qui se réunit tous les quatre ans, détermine les thèmes d'étude à traiter par les Commissions d'études de l'UIT-T, lesquelles élaborent en retour des Recommandations sur ces thèmes.

L'approbation des Recommandations par les Membres de l'UIT-T s'effectue selon la procédure définie dans la Résolution 1 de l'AMNT.

Dans certains secteurs des technologies de l'information qui correspondent à la sphère de compétence de l'UIT-T, les normes nécessaires se préparent en collaboration avec l'ISO et la CEI.

NOTE

Dans la présente Recommandation, l'expression "Administration" est utilisée pour désigner de façon abrégée aussi bien une administration de télécommunications qu'une exploitation reconnue.

Le respect de cette Recommandation se fait à titre volontaire. Cependant, il se peut que la Recommandation contienne certaines dispositions obligatoires (pour assurer, par exemple, l'interopérabilité et l'applicabilité) et considère que la Recommandation est respectée lorsque toutes ces dispositions sont observées. Le futur d'obligation et les autres moyens d'expression de l'obligation comme le verbe "devoir" ainsi que leurs formes négatives servent à énoncer des prescriptions. L'utilisation de ces formes ne signifie pas qu'il est obligatoire de respecter la Recommandation.

DROITS DE PROPRIÉTÉ INTELLECTUELLE

L'UIT attire l'attention sur la possibilité que l'application ou la mise en œuvre de la présente Recommandation puisse donner lieu à l'utilisation d'un droit de propriété intellectuelle. L'UIT ne prend pas position en ce qui concerne l'existence, la validité ou l'applicabilité des droits de propriété intellectuelle, qu'ils soient revendiqués par un Membre de l'UIT ou par une tierce partie étrangère à la procédure d'élaboration des Recommandations.

A la date d'approbation de la présente Recommandation, l'UIT n'avait pas été avisée de l'existence d'une propriété intellectuelle protégée par des brevets à acquérir pour mettre en œuvre la présente Recommandation. Toutefois, comme il ne s'agit peut-être pas de renseignements les plus récents, il est vivement recommandé aux responsables de la mise en œuvre de consulter la base de données des brevets du TSB.

© UIT 2003

Tous droits réservés. Aucune partie de cette publication ne peut être reproduite, par quelque procédé que ce soit, sans l'accord écrit préalable de l'UIT.

TABLE DES MATIÈRES

	Page
1	Domaine d'application 1
2	Références normatives..... 1
3	Définitions et abréviations 2
4	Traitement de l'écho acoustique 3
4.1	Composants analogiques 3
4.2	Unités fonctionnelles d'un limiteur d'écho acoustique 3
4.3	Interaction du limiteur d'écho du terminal avec l'équipement traitant le signal dans le réseau 4
4.4	Retard 4
4.5	Spécification de la réduction de l'écho acoustique 5
5	Réduction du bruit 8
5.1	Composants analogiques 8
5.2	Unités fonctionnelles d'un système de réduction du bruit..... 8
5.3	Interaction entre l'équipement de réduction du bruit du terminal et l'équipement de traitement du signal du réseau..... 9
5.4	Temps de traitement de réduction du bruit..... 9
5.5	Spécifications du système de réduction du bruit 10

Recommandation UIT-T P.330

Dispositifs de traitement vocal pour une amélioration de l'acoustique

1 Domaine d'application

La présente Recommandation concerne les caractéristiques de transmission génériques et les principes de performance et d'essai des dispositifs de traitement vocal pour une amélioration de l'acoustique (SPDA) destinés aux terminaux, indépendamment de l'application.

Un dispositif de traitement de la parole pour l'amélioration de l'acoustique est défini comme étant une fonction de traitement du signal intégrée dans les terminaux améliorant les signaux vocaux. Les fonctions d'amélioration de la voix sont notamment la gestion de la réduction de l'écho et du bruit acoustique.

La présente Recommandation a pour objet de définir un cadre permettant de spécifier les impératifs de performance des terminaux assurant la fonction SPDA et, le cas échéant, les essais qui peuvent être effectués sur de tels terminaux pour contrôler le respect de ces impératifs. Elle traite des caractéristiques génériques applicables tant aux terminaux analogiques que numériques. On trouvera dans la Rec. UIT-T P.340 [13] les caractéristiques s'appliquant exclusivement aux terminaux mains-libres.

Les méthodes d'essai applicables aux paramètres définis dans la présente Recommandation sont exposées dans la Rec. UIT-T P.502 [15].

En cas d'utilisation d'un simulateur de tête et de torse HATS pour les essais, la Rec. UIT-T P.581 [16] s'applique.

2 Références normatives

La présente Recommandation se réfère à certaines dispositions des Recommandations UIT-T et textes suivants qui, de ce fait, en sont partie intégrante. Les versions indiquées étaient en vigueur au moment de la publication de la présente Recommandation. Toute Recommandation ou tout texte étant sujet à révision, les usagers de la présente Recommandation sont invités à se reporter, si possible, aux versions les plus récentes des références normatives suivantes. La liste des Recommandations de l'UIT-T en vigueur est régulièrement publiée. La référence à un document figurant dans la présente Recommandation ne donne pas à ce document, en tant que tel, le statut d'une Recommandation.

- [1] Recommandation UIT-T G.114 (2003), *Temps de transmission dans un sens*.
- [2] Recommandation UIT-T G.121 (1993), *Equivalents pour la sonie des systèmes nationaux*.
- [3] Recommandation UIT-T G.122 (1993), *Influence des systèmes nationaux sur la stabilité et l'écho pour la personne qui parle dans les connexions internationales*.
- [4] Recommandation UIT-T G.131 (1996), *Réduction de l'écho pour le locuteur*.
- [5] Recommandation UIT-T G.167 (1993), *Dispositifs pour la réduction de l'écho acoustique*.
- [6] Recommandation UIT-T G.168 (2002), *Annuleurs d'écho pour les réseaux numériques*.
- [7] Recommandation UIT-T G.223 (1988), *Hypothèses pour le calcul du bruit sur les circuits fictifs de référence pour la téléphonie*.
- [8] Recommandation UIT-T P.10 (1998), *Terminologie relative à la qualité de la transmission téléphonique et aux appareils téléphoniques*.
- [9] Recommandation UIT-T P.50 (1999), *Voix artificielle*.

- [10] Recommandation UIT-T P.51 (1996), *Bouche artificielle*.
- [11] Recommandation UIT-T P.78 (1996), *Méthode subjective de détermination des équivalents pour la sonie, conforme à la Recommandation P.76*.
- [12] Recommandation UIT-T P.79 (1999), *Calcul des équivalents pour la sonie des postes téléphoniques*.
- [13] Recommandation UIT-T P.340 (2000), *Caractéristiques de transmission et paramètres de qualité vocale des terminaux mains-libres*.
- [14] Recommandation UIT-T P.501 (2000), *Signaux d'essai à utiliser en téléphonométrie*.
- [15] Recommandation UIT-T P.502 (2000), *Méthodes d'évaluation objective des systèmes de communication vocale utilisant des signaux de test complexes*.
- [16] Recommandation UIT-T P.581 (2000), *Utilisation du simulateur de tête et de torse pour les essais des terminaux mains-libres*.
- [17] Recommandation UIT-T P.800 (1996), *Méthodes d'évaluation subjective de la qualité de transmission*.
- [18] Recommandation UIT-T P.832 (2000), *Evaluation subjective des performances des terminaux mains-libres*.
- [19] Manuel de la téléphonométrie de l'UIT-T, 1992.
- [20] Recommandation UIT-T G.161 (2002), *Interaction des équipements de réseaux de traitement de signaux*.
- [21] Recommandation UIT-T G.108.2 (2003), *Annuleurs d'écho: planification de la transmission*.

3 Définitions et abréviations

Les définitions pertinentes indiquées en [8] sont applicables, ainsi que les définitions suivantes:

3.1 dispositif de traitement vocal pour une amélioration de l'acoustique (SPDA, *speech processing device for acoustic enhancement*): toute fonction de traitement du signal intégrée dans les terminaux et qui est destinée à améliorer les signaux vocaux.

3.2 écho acoustique (AE, *acoustic echo*): signal différé et réfléchi résultant du trajet acoustique entre l'écouteur ou le haut-parleur et le microphone d'un terminal mains-libres ou à combiné.

3.3 annuleur d'écho acoustique (AEC *acoustic echo canceller*): dispositif qui abaisse le niveau de l'écho acoustique moyennant des effets négligeables sur les signaux vocaux des interlocuteurs local et distant. Afin de suivre les variations du trajet d'écho acoustique, la réduction de l'écho acoustique est généralement mise en œuvre par identification adaptative de la réponse impulsionnelle du trajet d'écho acoustique.

3.4 commande d'affaiblissement: dispositif qui abaisse le niveau d'écho acoustique par introduction d'affaiblissements variables dans les signaux audio reçus ou émis.

3.5 processeur non linéaire (NLP, *non-linear processor*): dispositif réduisant l'écho résiduel qui n'est pas supprimé par le dispositif de réduction de l'écho acoustique. Il utilise un traitement non linéaire pour abaisser l'écho à un niveau où il n'est pas perçu par le sujet à l'autre extrémité de la communication. Un écrêteur est un dispositif couramment utilisé à cet effet.

3.6 dispositif de réduction du sifflement: dispositif qui modifie certaines caractéristiques des signaux émis ou reçus afin d'améliorer la marge de stabilité du terminal. Cette fonction est généralement mise en œuvre par un processeur harmonique. Pour éviter de perturber le réseau, il

convient d'éviter l'emploi de tels dispositifs dans des terminaux susceptibles d'être utilisés sur des connexions comportant des annuleurs d'écho électrique conformes à la Rec. UIT-T G.168 qui ne peuvent pas fonctionner avec des trajets d'écho dépendant du temps (par exemple déplacement de fréquence).

3.7 bruit de fond: défini comme étant le signal ajouté au signal vocal utile de l'extrémité locale. Il est dû dans une large mesure au signal acoustique que détecte le microphone en plus du signal vocal de l'extrémité locale.

3.8 bruit de confort: insertion d'un bruit pseudo-aléatoire pendant les périodes de silence (sans signaux vocaux actifs) par un dispositif SPDA.

3.9 réduction de bruit (NR, *noise reduction*): fonction qui diminue les effets gênants et fatigants du bruit de fond. Autrement dit, la fonction NR abaisse le niveau du bruit de fond dans le but d'améliorer la qualité globale perçue du signal.

3.10 évaluateur de bruit: dispositif qui fait, par calcul, une évaluation des caractéristiques du bruit de fond gênant. Pour les systèmes classiques fonctionnant dans le domaine fréquentiel, c'est la densité spectrale du bruit qui est calculée. Pour les systèmes fonctionnant dans le domaine temporel, la valeur estimée est l'autocorrélation du bruit. D'habitude, l'évaluateur de bruit agit pendant les périodes sans parole (déterminées par le détecteur VAD) mais d'autres systèmes sont possibles.

3.11 filtrage du bruit: traitement qui consiste à appliquer à l'entrée (autrement dit le signal bruité) le filtrage calculé par un système de réduction du bruit. Ce traitement est inclus dans le dispositif de réduction du bruit (NR). Le filtrage peut être appliqué dans le domaine temporel à échantillon par échantillon (convolution) ou en bloc dans le domaine fréquentiel (affaiblissement spectral à court terme).

3.12 effet Lombard: en présence d'un fort bruit de fond, les sujets ont tendance à parler plus fort, ce qui est appelé "l'effet Lombard".

3.13 détecteur d'activité vocale (VAD, *voice activity detector*): dispositif qui fait la distinction entre les périodes de silence (pas de signal vocal actif), les périodes pendant lesquelles un des interlocuteurs parle (la personne proche ou la personne à distance) et les périodes où les deux personnes parlent (signaux vocaux proches et distants actifs simultanément).

4 Traitement de l'écho acoustique

Les composants d'un terminal utilisés pour limiter l'écho acoustique sont les éléments analogiques (microphone et haut-parleur), l'annuleur d'écho acoustique et des dispositifs non linéaires.

4.1 Composants analogiques

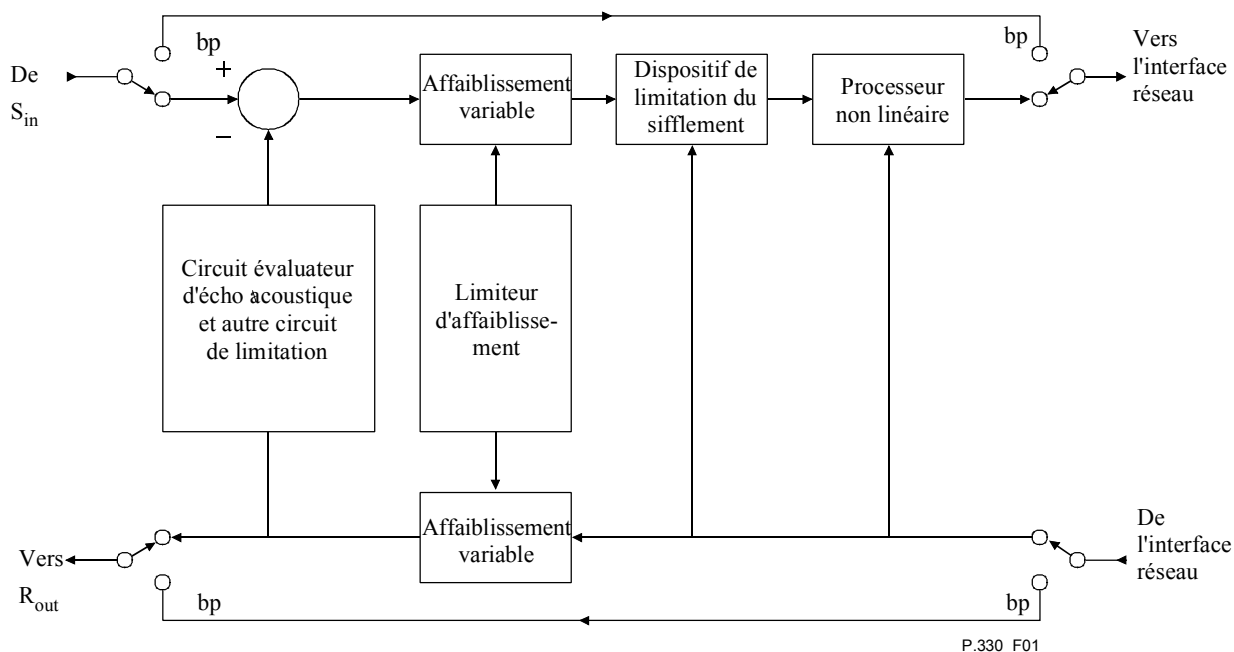
L'emplacement et le ou les types de microphone et de haut-parleur déterminent le niveau de l'écho acoustique (couplage acoustique) avant tout traitement du signal. En l'absence de réduction numérique, l'affaiblissement d'adaptation pour l'écho (ERL, *echo return loss*) d'un système mains-libres, par exemple, se situe entre 20 et -20 dB. Autrement dit, en l'absence de toute réduction de l'écho numérique, le niveau de l'écho acoustique peut être supérieur de 20 dB à celui du signal initialement reçu. On peut modifier l'affaiblissement ERL en changeant l'emplacement ou le type de microphone (emploi d'un microphone directif, par exemple) et de haut-parleur.

4.2 Unités fonctionnelles d'un limiteur d'écho acoustique

Les unités fonctionnelles d'un dispositif AEC sont les dispositifs ou parties de dispositif qui sont implémentés dans l'unité de traitement et contribuent au fonctionnement général des dispositifs pour la réduction de l'écho acoustique. Aucune restriction n'est imposée à leur mode d'implémentation.

Les unités fonctionnelles peuvent être combinées pour une meilleure performance. Elles peuvent utiliser tous les signaux disponibles dans le terminal (par exemple, les divers signaux provenant de plusieurs microphones disposés de manière à former un réseau acoustique). En outre, elles peuvent être associées à d'autres fonctions (par exemple, codage en sous-bandes de signaux vocaux) pour une implémentation rationnelle, sous réserve qu'elles ne modifient pas les caractéristiques propres de ces fonctions lorsqu'elles sont en service.

Le schéma fonctionnel d'une unité de traitement courante est présenté dans la Figure 1.



**Figure 1/P.330 – Schéma fonctionnel d'une unité de traitement courante (partie AEC)
(bp désigne les trajets de dérivation utilisés pour les essais)**

4.3 Interaction du limiteur d'écho du terminal avec l'équipement traitant le signal dans le réseau

La réduction de l'écho acoustique est la plus simple et la plus efficace lorsqu'elle a lieu dans le terminal. Toutefois, les équipements du réseau peuvent aussi appliquer une réduction de l'écho acoustique, et dans ces conditions on obtient une mise en cascade des fonctions de réduction. Le traitement complémentaire assuré par le réseau empêchera toute dégradation de la qualité globale perçue.

La réduction de l'écho électrique est obtenue par l'équipement du réseau, comme indiqué dans la Rec. UIT-T G.168. Si l'action de cet équipement est insuffisante, il convient d'utiliser dans le terminal un dispositif limitant partiellement l'écho électrique. Les interactions entre les deux dispositifs de traitement sont examinées dans la Rec. UIT-T G.161. Les aspects de la planification de la transmission des annuleurs d'écho sont traités dans la Rec. UIT-T G.108.2.

4.4 Retard

Les valeurs spécifiées ci-dessous correspondent au temps de transmission supplémentaire qui peut résulter du traitement de réduction AEC. Quel que soit ce retard, il faut satisfaire aux objectifs de planification de la transmission.

On trouvera dans la Rec. UIT-T G.114 des informations générales sur les temps de transmission et dans la Rec. UIT-T G.131 des règles applicables à la réduction de l'écho dans le réseau.

4.4.1 Temps de traitement

Le traitement nécessaire à la réduction de l'écho nécessite un certain temps, ce qui entraîne un retard dans le terminal, retard appelé "temps de traitement".

4.4.2 Temps de propagation aller-retour de l'écho (EPDn) – Interface de réseau

L'audibilité de l'écho dépend du temps de propagation aller-retour sur le trajet de l'écho. Elle nécessite de calculer la réponse impulsionnelle (trajet de l'écho). Il convient d'utiliser à cette fin le bruit stationnaire de la bande passante. Le terminal peut être placé dans pratiquement n'importe quel environnement réverbérant ou non réverbérant, puisque le premier écho acoustique sera dû au couplage direct. Le bruit acoustique devrait satisfaire aux conditions définies au § 5.4/P.340. L'AEC est d'abord conditionné au moyen d'un signal au niveau de l'interface de réception (R_{in}).

4.5 Spécification de la réduction de l'écho acoustique

On trouvera également les prescriptions en matière de performance dans le paragraphe 8/P.340 et les méthodes d'essai dans la Rec. UIT-T P.502.

4.5.1 Trajet de l'écho acoustique

L'utilisation de salles ou d'espaces réels ayant des caractéristiques appropriées est recommandée. Les trajets d'écho simulés par des dispositifs électroniques tels que les réverbérateurs numériques à répartition de la réflexion invariante avec le temps peuvent être utilisés également si le terminal a des accès extérieurs "côté usager". Dans ce dernier cas, les réglages du simulateur électronique doivent être conformes aux valeurs recommandées pour les salles ou enceintes réelles et la forme de l'enveloppe de la réponse impulsionnelle simulée doit être analogue à la réponse impulsionnelle du trajet d'écho réel:

- pour les systèmes de téléconférence, le temps de réverbération moyen dans la bande passante de transmission doit être généralement de 400 ms; le temps de réverbération dans l'octave la plus basse ne doit pas être supérieur au double de cette valeur moyenne; le temps de réverbération dans l'octave la plus élevée ne doit pas être inférieur à la moitié de cette valeur. Le volume d'une salle d'essai typique doit être de l'ordre de 90 m³;
- pour les terminaux et visiophones mains-libres, le temps de réverbération moyen dans la bande passante de transmission doit être généralement de 500 ms; le temps de réverbération dans l'octave la plus basse ne doit pas être supérieur au double de cette valeur moyenne; le temps de réverbération dans l'octave la plus élevée ne doit pas être inférieur à la moitié de cette valeur. Le volume d'une salle d'essai typique doit être de l'ordre de 50 m³;
- pour les radiotéléphones mobiles, on peut utiliser une enceinte simulant l'intérieur d'une voiture ou bien une voiture réelle. Le "temps de réverbération moyen" typique est de 60 ms. Le volume de l'enceinte doit être de l'ordre de 2,5 m³.

NOTE – Il est recommandé d'éviter des salles extrêmement longues (Longueur \gg Largeur, Hauteur) et des salles avec des plafonds extrêmement bas (Hauteur \ll Longueur, Largeur) et, si possible, les salles dont les trois dimensions sont proches les unes des autres.

Il convient d'éviter les surfaces murales larges, plates et parallèles et les parties de surfaces produisant une réflexion sonore large bande, particulièrement les surfaces murales à une hauteur de salle moyenne (grossièrement 0,8 à 1,8 m au-dessus du sol) car elles peuvent créer des "flutter" échos et des perturbations de type "flutter" écho (écho, rugosité), si le montage d'essai est dans une position défavorable.

Le fait de mesurer la distribution locale dépendant de la fréquence des niveaux de pression acoustique dans une salle donnée se trouvant en conditions stables peut aider à déterminer la position optimale du montage d'essai.

Comme suggestion générale, la distance minimale entre le montage d'essai et les surfaces murales de la salle devrait être de 1 m indépendamment des propriétés acoustiques de ces surfaces. Ceci peut prévenir les perturbations dues aux réflexions initiales et une élévation du niveau de pression acoustique qui peut apparaître localement aux basses fréquences. La même recommandation s'applique aux surfaces de grands éléments de mobilier qui réfléchissent le son.

4.5.2 Paramètres et limites recommandés

4.5.2.1 Equivalent pondéré de couplage du terminal – Monologue simple (TCLwst)

Affaiblissement pondéré entre les interfaces de réseau R_{in} et S_{out} lorsque l'AEC fonctionne normalement et que l'utilisateur local n'émet aucun signal¹.

Avant chaque essai, le terminal est allumé.

Les valeurs recommandées pour chaque type de terminal mains-libres sont indiquées dans les Recommandations UIT-T pertinentes (par exemple Rec. UIT-T P.341 pour les terminaux mains-libres à large bande ou Rec. UIT-T P.342 pour les terminaux mains-libres numériques).

4.5.2.2 Equivalent pondéré de couplage du terminal – Conversation simultanée (TCLwdt)

Affaiblissement pondéré entre les interfaces de réseau R_{in} et S_{out} lorsque l'AEC fonctionne normalement et que l'utilisateur local et l'utilisateur distant sont simultanément actifs¹.

Les valeurs recommandées pour chaque type de terminal mains-libres sont indiquées au paragraphe 8/P.340.

4.5.2.3 Affaiblissement des signaux vocaux reçus dans des conditions de conversation simultanée (Artdt)

Affaiblissement des signaux reçus (au point R_{out}) qui est inséré par l'AEC dans des conditions de conversation simultanée.

La réponse en fréquence du côté réception pendant une conversation simultanée devrait de préférence être la même qu'en conditions de monologue. Dans la pratique, il n'est toutefois pas possible d'appliquer une réduction de l'écho qui assure un affaiblissement suffisant pendant la conversation simultanée sans modifier la réponse en fréquence.

4.5.2.4 Affaiblissement des signaux vocaux émis dans des conditions de conversation simultanée (Asdt)

Affaiblissement des signaux émis (au point S_{out}) qui est inséré par l'AEC dans des conditions de conversation simultanée.

La réponse en fréquence du côté émetteur pendant la conversation simultanée devrait de préférence être la même que celle en conditions de monologue. Dans la pratique, il n'est toutefois pas possible d'appliquer une annulation de l'écho qui assure un affaiblissement suffisant pendant la conversation simultanée sans modifier la réponse en fréquence.

4.5.2.5 Distorsion des signaux vocaux reçus dans des conditions de conversation simultanée (Drtdt)

Distorsion non linéaire totale du signal au point R_{out} qui peut être produite par l'AEC dans des conditions de conversation simultanée.

¹ La pondération est effectuée conformément à la règle spécifiée dans la Rec. UIT-T G.122 (calcul de l'équivalent pour la sonie du trajet d'écho pour le locuteur). Des précautions doivent être prises pour éviter le masquage possible des effets d'amorçage d'oscillations par la pondération (à l'étude).

Pour toutes les applications, la distorsion supplémentaire au point R_{out} comparée aux conditions de monologue doit être faible.

4.5.2.6 Distorsion des signaux vocaux émis dans des conditions de conversation simultanée (Dsdt)

Distorsion non linéaire totale du signal au point S_{out} qui peut être produite par l'AEC dans des conditions de conversation simultanée.

Pour toutes les applications, la distorsion supplémentaire au point S_{out} comparée aux conditions de monologue doit être faible.

4.5.2.7 Temps de montée – Monologue (TRst)

Intervalle de temps entre le début du signal reçu (ou du signal émis) et l'instant où l'affaiblissement sur le trajet de réception (ou sur le trajet d'émission) atteint [3] dB. Dans ce cas, l'autre côté est inactif.

4.5.2.7.1 Côté réception (TRst-r)

Pour toutes les applications, TRst-r ne sera pas supérieur à [20 ms].

4.5.2.7.2 Côté émission (TRst-s)

Pour toutes les applications, TRst-s ne sera pas supérieur à [20 ms].

4.5.2.8 Temps de montée – Conversation simultanée (TRdt)

Intervalle de temps entre le début du signal reçu (ou du signal émis) et l'instant où l'affaiblissement sur le trajet de réception (ou sur le trajet d'émission) atteint la valeur Ardt (ou Asdt). Dans ce cas, le signal dans le sens opposé de transmission est maintenu à un niveau spécifié.

4.5.2.8.1 Côté réception (TRdt-r)

TRdt-r doit être inférieur à [20 ms] si l'affaiblissement est supérieur à 6 dB.

4.5.2.8.2 Côté émission (TRdt-s)

TRdt-s doit être inférieur à [20 ms] si l'affaiblissement est supérieur à 6 dB.

4.5.2.9 Temps de convergence (Tc)

Le temps de convergence est l'intervalle de temps entre l'instant où un signal d'essai spécifié est appliqué au port R_{in} du terminal (toutes les fonctions de l'émetteur AEC ayant été réinitialisées et ensuite activées) et l'instant où le signal de l'écho renvoyé au port S_{out} est affaibli par une valeur prédéfinie. L'utilisateur local n'est pas actif.

4.5.2.10 Temps de maintien après conversation simultanée (THdt)

Temps qui s'écoule entre la fin d'un événement de conversation simultanée et l'instant où l'affaiblissement de l'écho retrouve une valeur spécifiée (lorsqu'un signal est reçu en continu de l'utilisateur distant).

Dans toutes les applications, l'affaiblissement du signal au port S_{out} doit être d'au moins [20 dB] après un temps THdt de [1] seconde.

4.5.2.11 Affaiblissement de couplage de terminal temporairement pondéré – Monologue (TCLtst)

L'affaiblissement d'adaptation pour l'écho de R_{in} vers S_{out} est mesuré selon la procédure définie pour ERLtst dans la Rec. UIT-T P.502.

4.5.2.12 Affaiblissement d'adaptation pour l'écho temporairement pondéré – Parole simultanée (ERLtdt)

L'affaiblissement d'adaptation pour l'écho de R_{in} vers S_{out} est mesuré selon la procédure définie pour ERLtdt dans la Rec. UIT-T P.502.

5 Réduction du bruit

Le but principal d'un système de réduction du bruit (NR, *noise reduction*) utilisé dans un dispositif est d'abaisser les effets gênants et fatigants du bruit de fond transmis. Les techniques utilisées à cet effet sont, selon le cas, analogiques, numériques ou une combinaison analogique-numérique.

5.1 Composants analogiques

Les composants analogiques d'un système NR sont notamment le microphone et tout circuit analogique reliant le microphone au CODEC (convertisseur analogique-numérique). Il existe plusieurs techniques utilisées pour abaisser le bruit de fond qui s'appuient uniquement sur des composants analogiques:

- a) la proximité du microphone par rapport à la bouche de la personne qui parle est un facteur décisif de l'établissement du rapport signal/bruit (SNR, *signal to noise ratio*). Le fait de rapprocher le microphone de la bouche produit des améliorations du rapport SNR évidentes et significatives (SNRE, *SNR enhancement*). On peut utiliser des microphones additionnels pour améliorer le rapport SNR;
- b) le trajet du signal analogique est généralement conçu de manière à présenter une réponse en fréquence qui est celle d'un filtre passe-haut. Lorsque le bruit comporte de fortes composantes à basse fréquence (par exemple le bruit d'une automobile), cette technique de filtrage améliore le rapport SNR (mesuré sur toute la bande). La répercussion est une perte notable de timbre dans la qualité vocale (surtout des voix masculines);
- c) des microphones peuvent être conçus pour rendre un gain directionnel passif. Le type le plus couramment utilisé pour les voitures est le microphone différentiel de premier ordre. Celui-ci peut être à simple transducteur utilisant deux ports. Avec un champ de bruit diffus et l'orientation correcte du microphone, un montage à plusieurs microphones différentiels de premier ordre hypercardioïdes améliore le rapport SNR de 6 dB par rapport à un microphone omnidirectionnel. On peut aussi utiliser des microphones différentiels d'ordre supérieur. Les montages à plusieurs microphones utilisant uniquement des techniques passives sont utilisables mais ont peu de chances de se généraliser étant donné que ces montages doivent être très grands pour avoir un effet sur les composantes basse fréquence de la parole. Ils peuvent contenir jusqu'à 16 éléments pouvant offrir un gain directionnel d'environ 20 dB à certaines fréquences.

5.2 Unités fonctionnelles d'un système de réduction du bruit

Les unités fonctionnelles d'un système de réduction du bruit sont des dispositifs ou des parties de dispositifs, implémentés dans l'unité de traitement, qui contribuent à la fonction générale de réduction du bruit. Il n'y a aucune restriction quant à la manière de les implémenter.

Un schéma fonctionnel d'une unité de traitement courante est présenté dans la Figure 2.

On peut implémenter deux types courants de techniques de réduction du bruit numérique dans un processeur de signal numérique (DSP, *digital signal processor*) ou dans un autre type de microprocesseur de terminal utilisant un seul microphone. On utilise couramment les techniques suivantes:

- a) suppression du bruit dans toute la bande: au cours des pauses de la conversation, le bruit est abaissé de manière significative tant que son énergie est inférieure à un niveau seuil.

Pendant la conversation active, l'affaiblissement est supprimé, ce qui permet le passage de la parole et du bruit. Cela produit un effet de pompage de bruit indésirable si l'affaiblissement est réglé à un niveau trop élevé;

- b) suppression des bruits de sous-bande: le signal transmis est réparti en sous-bandes au moyen d'un algorithme FFT, transformée de Fourier rapide (FTT, *fast Fourier transform*). Seules les bandes de fréquences qui présentent un bruit stationnaire sont atténuées tandis que les bandes transmettant les signaux vocaux ne le sont pas. La méthode bien connue de la soustraction spectrale est une de ces techniques. Dans la pratique, les niveaux de suppression du bruit se situent entre 6 et 15 dB. L'inconvénient de ces techniques est l'existence d'un compromis entre le niveau de réduction du bruit et la distorsion du signal vocal original. Dès lors, il est difficile de trouver un réglage qui fonctionne dans toutes les conditions de bruit (rapport SNR et type de bruit). En conditions de faible rapport SNR, toutefois, le signal vocal est quelque peu dégradé si le niveau de réduction est élevé.

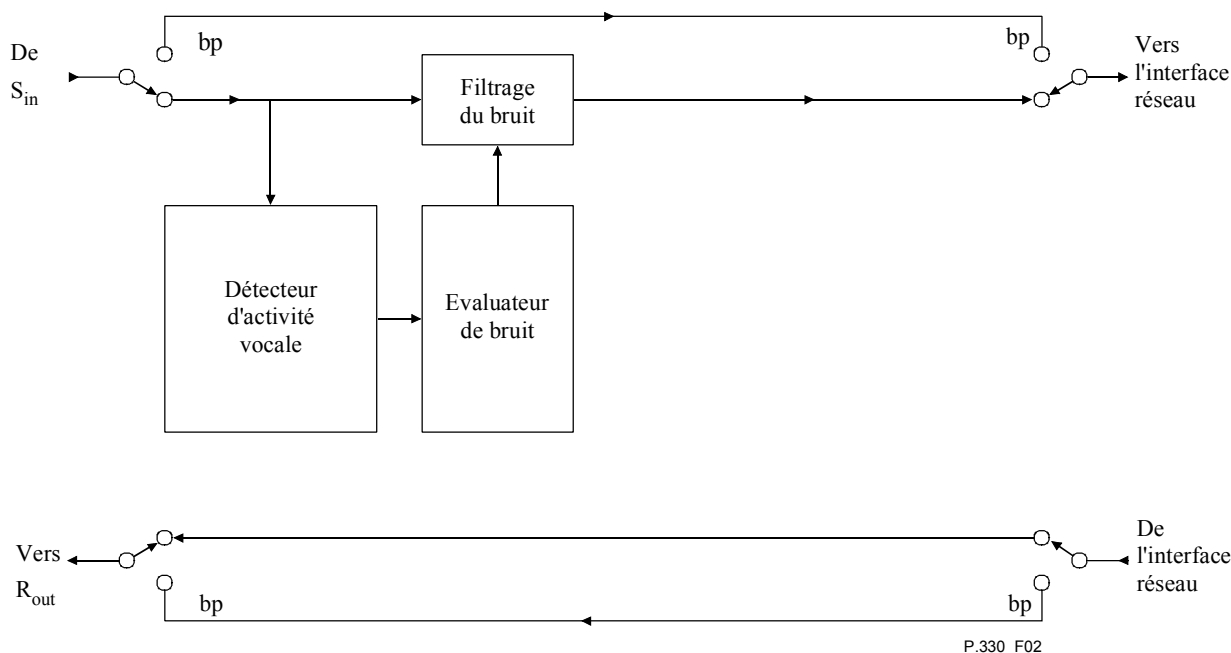


Figure 2/P.330 – Schéma fonctionnel d'une unité de traitement (partie NR) courante (bp désigne le trajet de dérivation du signal pour les besoins de l'essai)

5.3 Interaction entre l'équipement de réduction du bruit du terminal et l'équipement de traitement du signal du réseau

L'équipement du réseau peut comporter un traitement de réduction du bruit. Cela conduit à mettre en cascade les fonctionnalités de réduction du bruit. Cette composante de réseau additionnelle empêchera toute dégradation de la qualité globale perçue.

5.4 Temps de traitement de réduction du bruit

Le temps de traitement de réduction du bruit dépend dans une large mesure de la technique utilisée pour le filtrage du bruit. Il convient en tout cas de respecter les objectifs de planification de la transmission.

On trouvera des renseignements généraux sur les temps de transmission dans la Rec. UIT-T G.114.

5.5 Spécifications du système de réduction du bruit

5.5.1 Bruit ambiant

Les caractéristiques acoustiques du cadre des essais sont présentées au § 4.5.1. Il convient de prendre en compte l'effet des conditions ambiantes dans le cas d'une prise de son à distance; en effet, la réverbération du bruit de fond peut être considérée comme une dégradation additionnelle.

La diffusion de signaux d'essai de bruit de fond est examinée au § 7.10/P.340.

Les signaux d'essai du bruit de fond doivent comporter des signaux réels tels que ceux de murmures confus, de bruits de bureau, de bruits de rue, de bruits de voiture (moteurs, conditions de déplacement à diverses vitesses) et d'autres signaux de bruit de fond simulés (selon la vocation de l'équipement).

Il conviendra de faire varier les niveaux des signaux de bruit de fond de manière à obtenir des rapports SNR de l'ordre de $[-3 \text{ dB à } 30 \text{ dB}]$.

NOTE – On trouvera certains des signaux d'essai correspondants dans la Rec. UIT-T P.501 et les méthodes d'essai dans la Rec. UIT-T P.502. D'autres signaux d'essai sont à l'étude.

5.5.2 Paramètres et limites recommandées

Il conviendra de mesurer tous les paramètres spécifiés pour différentes valeurs du rapport SNR de la gamme $[-3 \text{ dB à } 30 \text{ dB}]$.

En ce qui concerne les paramètres relatifs à la mesure de l'affaiblissement du niveau du signal ou à son temps de transmission, on trouvera les procédures de mesure (méthodes et stimulus) dans les Recommandations UIT-T P.501, P.502 et P.340, qui s'appliquent avec les restrictions suivantes:

- le niveau du signal vocal doit être supérieur de 10 dB au moins au niveau de bruit;
- le bruit doit être stationnaire.

Des méthodes d'essai pour tous les autres cas sont à l'étude (bruit non stationnaire, faibles valeurs de rapport SNR, mesure de la distorsion).

Tous les paramètres définis ci-dessous correspondent à des conditions de conversation simple. En raison des interactions éventuelles entre l'annuleur AEC et le traitement de réduction du bruit intégré dans le terminal, il convient d'examiner également les paramètres en conditions de conversation simultanée (à l'étude).

5.5.2.1 Affaiblissement des signaux vocaux émis en conditions calmes (Asqc)

L'affaiblissement du signal émis (au point S_{out}) qui est introduit par la réduction NR en conditions calmes.

5.5.2.2 Distorsion des signaux vocaux à l'émission en conditions calmes (Dsqc)

La distorsion de signal non linéaire totale au point S_{out} pouvant être produite par la réduction NR en conditions calmes.

Pour toutes les applications, la distorsion supplémentaire à S_{out} par comparaison avec S_{in} doit être aussi faible que possible. En principe, aucune distorsion additionnelle ne doit être introduite par la NR.

5.5.2.3 Affaiblissement des signaux vocaux émis dans des conditions de bruit (Asnc)

L'affaiblissement du signal émis (au point S_{out}) qui est introduit par la NR pendant des événements bruyants.

En principe, aucune distorsion additionnelle ne doit être introduite par la NR.

5.5.2.4 Distorsion des signaux vocaux émis dans des conditions de bruit (D_{snc})

La distorsion non linéaire totale du signal au point S_{out} qui peut être produite par la NR pendant des événements bruyants.

Pour toutes les applications, la distorsion supplémentaire à S_{out} par comparaison avec S_{in} doit être aussi faible que possible. En principe, aucune distorsion additionnelle ne doit être introduite par NR.

5.5.2.5 Temps d'adaptation (TA, *adaptation time*)

Le temps d'adaptation est l'intervalle entre l'instant où un signal d'essai de bruit spécifié est appliqué au port S_{in} du terminal (après réinitialisation et activation de toutes les fonctions de NR) et l'instant où le signal d'essai de bruit renvoyé au port S_{out} est stable à ± 1 dB près par rapport au niveau de bruit réduit à long terme (voir Figure 3). Les utilisateurs locaux et distants ne sont pas actifs.

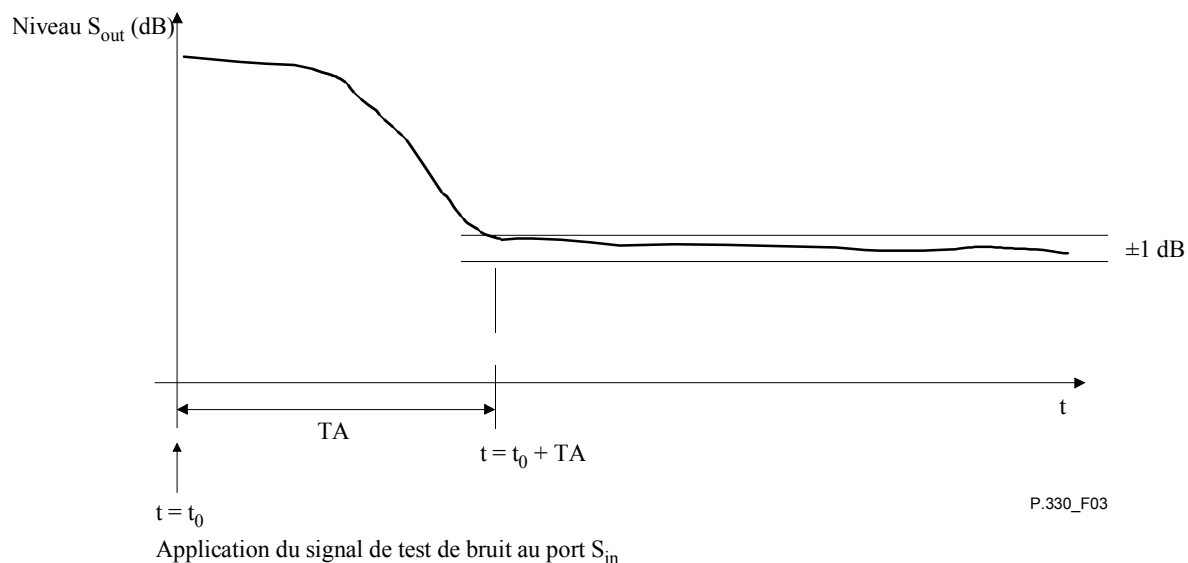


Figure 3/P.330 – Définition du temps d'adaptation (TA)

5.5.2.6 Temps d'adaptation après un événement vocal (TA_{se})

Le temps qui s'écoule entre la fin d'un événement vocal et l'instant où l'affaiblissement du bruit se rétablit à une valeur spécifiée.

Pour toutes les applications ayant des niveaux élevés de bruit de fond [-3 dB < SNR < 15 dB], l'affaiblissement du signal de bruit à S_{out} doit être d'au moins [6 dB] après l'adaptation TA_{se} de [100] millisecondes.

5.5.2.7 Affaiblissement du bruit de terminal – Sans signaux vocaux (TNAt_{ns})

L'affaiblissement du bruit de terminal entre S_{in} et S_{out} qui est introduit par la NR dans le signal de bruit de fond en l'absence de signaux vocaux.

5.5.2.8 Distorsion de bruit – En l'absence de signaux vocaux (D_{nns})

La distorsion non linéaire totale du signal au point S_{out} pouvant être produite par la NR sur le signal de bruit de fond en l'absence de tout signal vocal.

Dans toutes les applications, la distorsion au niveau de S_{out} par rapport à S_{in} doit être négligeable.

5.5.2.9 Affaiblissement du bruit de terminal – En présence des signaux vocaux (TNAt_{ps})

L'atténuation de bruit de terminal entre S_{in} et S_{out} qui est introduit par la NR dans le signal de bruit de fond en présence de signaux vocaux (mesure de l'amélioration du signal rapport/bruit).

5.5.2.10 Niveau de bruit de confort et correspondance du spectre – En l'absence de signaux vocaux (CNLMns et CNSMns)

Le bruit de confort au point de sortie S_{out} qui peut être introduit par la NR en l'absence de signaux vocaux doit correspondre au plan du niveau du spectre au bruit de fond présent au point S_{in} .

5.5.2.11 Niveau de bruit de confort et correspondance de spectre – En présence de signaux vocaux (CNLMps et CNSMps)

Le bruit de confort au point de sortie S_{out} qui peut être introduit par la NR en présence de signaux vocaux doit correspondre au plan du niveau du spectre au bruit de fond présent au point S_{in} .

SÉRIES DES RECOMMANDATIONS UIT-T

Série A	Organisation du travail de l'UIT-T
Série B	Moyens d'expression: définitions, symboles, classification
Série C	Statistiques générales des télécommunications
Série D	Principes généraux de tarification
Série E	Exploitation générale du réseau, service téléphonique, exploitation des services et facteurs humains
Série F	Services de télécommunication non téléphoniques
Série G	Systèmes et supports de transmission, systèmes et réseaux numériques
Série H	Systèmes audiovisuels et multimédias
Série I	Réseau numérique à intégration de services
Série J	Réseaux câblés et transmission des signaux radiophoniques, télévisuels et autres signaux multimédias
Série K	Protection contre les perturbations
Série L	Construction, installation et protection des câbles et autres éléments des installations extérieures
Série M	RGT et maintenance des réseaux: systèmes de transmission, circuits téléphoniques, télégraphie, télécopie et circuits loués internationaux
Série N	Maintenance: circuits internationaux de transmission radiophonique et télévisuelle
Série O	Spécifications des appareils de mesure
Série P	Qualité de transmission téléphonique, installations téléphoniques et réseaux locaux
Série Q	Commutation et signalisation
Série R	Transmission télégraphique
Série S	Equipements terminaux de télégraphie
Série T	Terminaux des services télématiques
Série U	Commutation télégraphique
Série V	Communications de données sur le réseau téléphonique
Série X	Réseaux de données et communication entre systèmes ouverts
Série Y	Infrastructure mondiale de l'information et protocole Internet
Série Z	Langages et aspects généraux logiciels des systèmes de télécommunication