

Union internationale des télécommunications

UIT-T

SECTEUR DE LA NORMALISATION
DES TÉLÉCOMMUNICATIONS
DE L'UIT

P.880

(05/2004)

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

Méthodes d'évaluation objective et subjective de la qualité

**Evaluation continue d'une qualité vocale
variable dans le temps**

Recommandation UIT-T P.880

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	Series	P.10
Lignes et postes d'abonnés	Series	P.30
		P.300
Normes de transmission	Series	P.40
Appareils de mesures objectives	Series	P.50
		P.500
Mesures électroacoustiques objectives	Series	P.60
Mesures de la sonie vocale	Series	P.70
Méthodes d'évaluation objective et subjective de la qualité	Series	P.80
		P.800
Qualité audiovisuelle dans les services multimédias	Series	P.900
Aspects relatifs à la qualité de transmission et à la qualité de service aux points de terminaison des réseaux à protocole Internet	Series	P.1000

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

Recommandation UIT-T P.880

Evaluation continue d'une qualité vocale variable dans le temps

Résumé

La présente Recommandation décrit une méthode d'évaluation continue d'une qualité vocale variable dans le temps, désignée sous le sigle CETVSQ (*continuous evaluation of time varying speech quality*), qui peut être utilisée pour évaluer l'incidence des fluctuations dans le temps de la qualité vocale sur la qualité instantanée perçue (à un instant quelconque de la séquence vocale) et sur la qualité globale perçue (à la fin de la séquence vocale). Cette méthode permet de formuler, dans un premier temps, une appréciation instantanée selon une échelle continue avec curseur pendant la séquence vocale et, dans un deuxième temps, une appréciation globale selon une échelle de catégories à cinq notes classique à la fin de la séquence vocale.

Source

La Recommandation UIT-T P.880 a été approuvée le 14 mai 2004 par la Commission d'études 12 (2001-2004) de l'UIT-T selon la procédure définie dans la Recommandation UIT-T A.8.

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 2005

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	Abréviations 1
4	Description de la méthode 1
4.1	Origine et motifs 1
4.2	Préparation des essais 2
4.3	Séance d'écoute 2
4.4	Analyse statistique 5
	BIBLIOGRAPHIE 6

Recommandation UIT-T P.880

Evaluation continue d'une qualité vocale variable dans le temps

1 Domaine d'application

La présente Recommandation définit une méthode d'évaluation subjective de la qualité vocale transmise pendant de longues séquences vocales de qualité ou de durée variable. Cette méthode repose à la fois sur une évaluation continue en cours d'écoute d'une séquence vocale et sur une évaluation globale à la fin de la séquence vocale. Elle permet donc de mesurer, non seulement la qualité globale perçue (comme avec les méthodes généralement recommandées), mais aussi la qualité instantanée perçue, *c'est-à-dire* la qualité perçue à un instant quelconque d'une séquence vocale transmise pour écoute. Dans sa version actuelle, cette méthode n'est pas applicable à la sélection de codecs vocaux. Néanmoins, elle est utile pour diagnostiquer les effets de certaines dégradations sur la qualité instantanée ou globale perçue, notamment les dégradations discontinues fluctuant dans le temps (dues à la perte de paquets IP ou au passage d'un réseau mobile à un autre, par exemple). Elle peut également faciliter la mise au point et la validation d'instruments de mesure objective destinés à prévoir la qualité vocale par détection et analyse des divers types de dégradations présentes dans un signal vocal; en fait, la méthode d'évaluation continue d'une qualité vocale variable dans le temps (CETVSQ, *continuous evaluation of time varying speech quality*) permet de procéder à des appréciations instantanées, ainsi qu'à une évaluation globale de la qualité subjective.

2 Références normatives

- [1] Recommandation UIT-T P.800 (1996), *Méthodes d'évaluation subjective de la qualité de transmission*.
- [2] Recommandation UIT-R BT.500-11 (2002), *Méthodologie d'évaluation subjective de la qualité des images de télévision*.

3 Abréviations

La présente Recommandation utilise les abréviations suivantes:

ACR	évaluation par catégories absolues (<i>absolute category rating</i>)
ANOVA	analyse de la variance (<i>analysis of variance</i>)
CETVSQ	évaluation continue d'une qualité vocale variable dans le temps (<i>continuous evaluation of time varying speech quality</i>)
MOS	note moyenne d'opinion (<i>mean opinion score</i>)
QS	qualité de service
SSCQE	évaluation continue de la qualité avec stimulus unique (<i>single stimulus continuous quality evaluation</i>)

4 Description de la méthode

4.1 Origine et motifs

La mise au point de la méthode d'évaluation continue d'une qualité vocale variable dans le temps (CETVSQ) est motivée par le fait que la qualité de service (QS) des nouveaux réseaux varie, même au cours d'une simple conversation, en raison de certaines dégradations (perte de paquets en téléphonie IP ou passage d'un réseau mobile à un autre, par exemple). Ces dégradations se

caractérisent par des artefacts de qualité transitoires (par opposition aux dégradations "continues" telles que les rapports signal/bruit) d'une densité plus ou moins élevée. Des échantillons courts ne permettent pas de rendre compte de cette densité. De plus, en raison des caractéristiques techniques des réseaux mobiles ou IP, la qualité vocale peut varier sensiblement au cours d'une même communication. Aux fins de l'évaluation de la qualité vocale, les méthodes ordinaires décrites dans la Rec. UIT-T P.800 [1] utilisent des stimuli de courte durée (8 s) pour les essais d'écoute subjectifs. Ces méthodes se prêtent bien à l'évaluation d'une qualité vocale constante dans le temps. Toutefois, sauf si on procède à une évaluation portant sur un très grand nombre d'échantillons, ces méthodes ne permettent pas de tenir compte de la manière dont ces dégradations se produisent et sont réparties dans la réalité. D'autre part, elles ne permettent pas non plus de tenir compte des fluctuations dans le temps de la qualité de longue durée faisant intervenir des processus mnésiques (phénomènes propres à la mémoire) qui ont une incidence sur la qualité globale perçue. L'absence d'une méthodologie permettant d'évaluer la qualité vocale variable dans le temps justifie la présente Recommandation. Par conséquent, pour évaluer l'incidence sur la qualité perçue de ces sortes de dégradations et de leurs fluctuations dans le temps au cours d'une communication donnée, il convient d'utiliser des séquences vocales plus longues que celles qu'utilisent généralement les méthodes subjectives classiques. En outre, l'utilisation d'une méthode d'appréciation continue permet d'étudier l'incidence des dégradations transitoires sur la qualité perçue à tout instant durant l'écoute de la séquence vocale, et sur la qualité globale perçue (à la fin de cette séquence). Cette méthode est inspirée de la méthode d'évaluation continue de la qualité avec stimulus unique (SSCQE, *single stimulus continuous quality evaluation*) utilisée dans le domaine de la vidéo (Rec. UIT-R BT.500-11 [2]). Elle a également été validée pour l'évaluation de la qualité vocale dans le cadre de plusieurs études antérieures ([B-1], [B-2], [B-3], [B-4]).

4.2 Préparation des essais

4.2.1 Stimuli

Le texte à prononcer doit être simple, clair et facile à comprendre. Les séquences vocales courtes doivent être éliminées et, à titre provisoire, des séquences vocales d'une durée de 45 secondes à 3 minutes pourront être utilisées. Les enregistrements de la source, y compris l'environnement et la procédure d'enregistrement, les systèmes d'émission et d'enregistrement, les locuteurs et les niveaux de parole pourraient être les mêmes que ceux qui sont indiqués dans le § B.1/P.800 [1].

4.2.2 Sources

Le nombre et le choix des conditions dépendent de la finalité de l'essai. La seule limite est celle qui est imposée par la durée de l'essai. Si possible, il est recommandé d'inclure dans l'ensemble des conditions d'essai certaines conditions de commande, c'est-à-dire des conditions dont les paramètres physiques ne sont sujets à aucune variation.

4.3 Séance d'écoute

4.3.1 Auditeurs

Vingt-quatre auditeurs non expérimentés, au moins, doivent participer au test. Tous les auditeurs doivent avoir pour langue maternelle la langue utilisée pour le test et aucun d'entre eux ne devrait avoir participé à une expérience subjective dans les six mois qui précèdent.

4.3.2 Présentation audio et environnement d'essai

La présentation audio doit être conforme aux directives données dans la Rec. UIT-T P.800 [1]. Ces directives portent notamment sur le système d'écoute, les niveaux d'écoute et l'environnement d'écoute.

4.3.3 Dispositif d'enregistrement et configuration pour l'appréciation continue de la qualité

Un mécanisme électronique à glissière (résistance variable à curseur) connecté à un ordinateur sera utilisé pour enregistrer l'évaluation continue de la qualité faite par les sujets. Ce dispositif aura les caractéristiques suivantes:

- mécanisme à glissière sans position de réinitialisation (c'est-à-dire, pas de retour automatique à une position prédéfinie);
- course: environ 10 cm;
- fixe ou pouvant être installé sur un bureau;
- enregistrement des échantillons, au niveau de la position du curseur, deux fois par seconde (ce qui constitue une vitesse suffisante pour restituer fidèlement les réponses des sujets);
- le codage de la position du curseur de 0 (bas de l'échelle) à 100 au minimum (haut de l'échelle) permettrait d'obtenir une résolution satisfaisante. La position initiale du curseur doit correspondre au point milieu de l'échelle.

4.3.4 Evaluation

Pour chaque séquence vocale, les sujets doivent procéder à deux opérations, à savoir: une évaluation continue en cours d'écoute et une évaluation globale à la fin de la séquence. Pour plus de précisions et les résultats des études précédentes, voir [B-1] et [B-2].

a) Evaluation continue

Dans un premier temps, il est demandé aux sujets d'évaluer la qualité vocale de la séquence en continu en déplaçant un curseur le long d'une échelle continue de manière que la position de ce curseur indique leur appréciation de la qualité d'écoute à l'instant considéré; les sujets peuvent placer le curseur en n'importe quel point de l'échelle. Cinq notes d'appréciation de qualité d'écoute sont portées le long de l'échelle, à savoir: excellente, bonne, passable, médiocre et mauvaise, pour aider les sujets à associer la position du curseur aux valeurs appropriées de qualité vocale.

Echelle de qualité continue

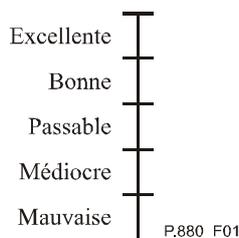


Figure 1/P.880 – Echelle continue utilisée pour l'appréciation de la qualité instantanée

b) Evaluation globale

Dans un deuxième temps, à la fin de chaque séquence, il est demandé aux sujets d'en évaluer la qualité globale selon l'échelle de qualité d'écoute à cinq notes (identique à l'échelle MOS utilisée dans l'évaluation par catégories absolues (ACR)).

Echelle de qualité globale (ACR)

Qualité de la parole	Note associée
Excellente	5
Bonne	4
Passable	3
Médiocre	2
Mauvaise	1

4.3.5 Procédure d'essai

Avant d'évaluer les séquences vocales à tester, les sujets suivent une formation pour s'entraîner à écouter un petit nombre de séquences retenues aux fins de la formation. Ces séquences doivent inclure différents niveaux de qualité et différentes fluctuations de qualité représentatifs de l'éventail des fluctuations dans le temps et des niveaux de qualité auxquels les sujets seront confrontés durant l'essai proprement dit.

En règle générale, l'essai consiste en un certain nombre de séances entrecoupées de pauses. L'ensemble complet de stimuli (séquences) est présenté dans un ordre aléatoire différent aux divers groupes de sujets.

4.3.6 Directives pour les sujets

Le Tableau 1 donne un exemple de directives types (pour des séquences vocales d'une durée de T secondes). Celles-ci doivent être données par écrit (oralement aussi, si nécessaire), avant le début de l'expérience.

Tableau 1/P.880 – Exemple de directives que recevront les auditeurs

Dans le présent test, vous écouterez des séquences vocales d'une durée de T secondes (T-s) par l'intermédiaire du combiné téléphonique. La qualité vocale de chaque séquence peut varier dans le temps de diverses manières. Pour chaque séquence, vous devrez:

- 1) donner votre opinion sur la qualité vocale, pendant toute la durée de la séquence, c'est-à-dire à tout instant, en déplaçant le curseur sur le tableau placé devant vous. L'échelle d'évaluation étant continue, vous pouvez placer le curseur n'importe où, sa position indiquant votre opinion sur la qualité vocale.
Des notes d'appréciation de la qualité (excellente, bonne, passable, médiocre et mauvaise) sont portées en différents points de l'échelle, pour vous aider à placer le curseur. Par exemple, si vous estimez que la qualité correspond exactement à une note intermédiaire entre passable et bonne, vous placerez le curseur au point milieu entre les deux notes correspondantes. Cela dit, vous n'êtes pas tenus de placer le curseur directement en regard d'une note ou exactement à mi-chemin entre deux notes, sauf si l'une de ces positions correspond exactement à votre opinion.
N'oubliez pas de déplacer le curseur à mesure que la qualité varie;
- 2) à la fin de chaque séquence, il vous sera demandé de donner une note de qualité globale qui devrait refléter votre opinion sur la qualité vocale de toute la séquence. A cet effet, appuyez sur le bouton approprié pour indiquer votre opinion selon l'échelle suivante:

Tableau 1/P.880 – Exemple de directives que recevront les auditeurs

<p>Opinion globale sur la qualité vocale de la séquence que vous venez d'entendre:</p> <p>Excellente</p> <p>Bonne</p> <p>Passable</p> <p>Médiocre</p> <p>Mauvaise</p> <p>Vous aurez cinq secondes pour enregistrer votre réponse en appuyant sur le bouton correspondant à votre choix. Il y aura une courte pause avant la présentation de la séquence suivante.</p> <p>Nous commencerons par une courte séance d'exercices pour vous familiariser avec la procédure d'essai. Les tests proprement dits donneront lieu à des séances de 10 à 15 minutes.</p> <p>Nous vous remercions de votre assistance.</p>

4.4 Analyse statistique

Pour chaque sujet, si T -s correspond à la durée de chaque séquence (en secondes), on procède à l'enregistrement d'un fichier de données de valeurs $2 \times T$ -s (correspondant à une note instantanée toutes les 500 ms pendant T secondes), ainsi que d'une valeur scalaire (donnant l'appréciation de la qualité globale). On transforme ensuite linéairement les valeurs instantanées $2 \times T$ -s (de $t = 0$ à $t = 2T - 1$) en valeurs de 1 à 5, selon la relation $S(t) = 1 + 4$ (position du curseur/maximum), où $S(t)$ est la note d'opinion instantanée.

Pour chaque séquence, une note moyenne d'opinion instantanée (et l'écart type correspondant) est obtenue en établissant la moyenne des différentes notes d'opinion instantanée pour l'ensemble des sujets, à chaque instant t (c'est-à-dire, toutes les 500 ms). Pour chaque séquence, une note moyenne d'opinion (MOS) globale (et l'écart type correspondant) est obtenue en établissant la moyenne des différentes notes d'opinion globale pour l'ensemble des sujets selon l'échelle ACR.

Des écarts importants entre la note continue et la note globale sont généralement révélateurs de dégradations transitoires. Tout dépendra du nombre d'occurrences et de la durée des dégradations ainsi que du fait qu'elles soient apparues récemment, par rapport à la note d'opinion globale. Un complément d'étude s'impose avant que l'on soit en mesure d'utiliser ces informations aux fins de la planification de la transmission.

L'analyse statistique (ANOVA, par exemple) permet d'identifier les effets notables présents dans différentes conditions expérimentales.

Remarque: les réponses de certains sujets peuvent être écartées si elles présentent des variations importantes et correspondent à plus de 10% du temps des écarts types entre deux sujets (calculés pour l'ensemble des sujets).

BIBLIOGRAPHIE

- [B-1] Contribution ITU-T [COM 12-94] (1999), *Continuous assessment of time-varying subjective vocal quality and its relationship with overall subjective quality*, (Source FT R&D).
- [B-2] GROS (L.), CHATEAU (N.), Instantaneous and Overall Judgments for Time-Varying Speech Quality: Assessments and Relationships, *Acta Acustica*, 87, 367-377, 2001.
- [B-3] Contribution ITU-T [COM 12-D139] (2000), *Study of the relationship between instantaneous and overall subjective speech quality for time-varying quality speech sequences: influence of recency effect*, (Source FT R&D).
- [B-4] Contribution ITU-T [COM 12-22] (2000), *Assessment of time-varying speech quality: comparison between listening situation and conversational situation*, (Source FT R&D).

SÉRIES DES RECOMMANDATIONS UIT-T

Série A	Organisation du travail de l'UIT-T
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	Gestion des télécommunications y compris le RGT et maintenance des réseaux
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, communication entre systèmes ouverts et sécurité
Série Y	Infrastructure mondiale de l'information, protocole Internet et réseaux de prochaine génération
Série Z	Langages et aspects généraux logiciels des systèmes de télécommunication