

RECOMMANDATION UIT-R BS.562-3<sup>\*,\*\*</sup>**Evaluation subjective de la qualité du son**

(1978-1982-1986-1990)

L'Assemblée des radiocommunications de l'UIT,

*considérant*

- a) que les essais subjectifs d'écoute permettent d'évaluer la gêne causée à l'auditeur par toute dégradation du signal utile au cours de sa transmission depuis la source du programme jusqu'à l'auditeur;
- b) que cette évaluation implique qu'une séquence de programme dégradée doit être comparée à la séquence originale, laquelle doit être d'«excellente qualité» ou dégradée d'une façon «imperceptible»;
- c) que, pour réaliser ces évaluations d'une manière comparable les unes avec les autres, les conditions d'écoute, la composition de l'équipe d'auditeurs et les séquences de programme doivent autant que possible être normalisées;
- d) qu'il serait souhaitable de disposer, tant pour les programmes son que pour les programmes de télévision, d'une seule et même échelle d'évaluation,

*recommande*

**1** que les échelles ci-dessous soient utilisées pour l'évaluation subjective de la qualité ou de la dégradation du son en radiodiffusion (pour les images de télévision, voir la Recommandation UIT-R BT.500). D'après la nature et l'objet des essais, on déterminera quelle est celle de ces deux échelles qui convient le mieux.

---

\* Cette Recommandation intéresse aussi la Commission d'études 9 de la normalisation des télécommunications.

\*\* La Commission d'études 6 des radiocommunications a apporté des modifications rédactionnelles à cette Recommandation en 2002 conformément aux dispositions de la Résolution UIT-R 44.

### 1.1 Echelle de qualité et de dégradation à cinq notes\*

TABLEAU 1

Qualité	Dégradation
5 Excellente	5 Imperceptible
4 Bonne	4 Perceptible, mais non gênante
3 Assez bonne	3 Légèrement gênante
2 Médiocre	2 Gênante
1 Mauvaise	1 Très gênante

### 1.2 Echelle de comparaison à sept notes

Pour certaines catégories d'essais subjectifs, il peut être plus pratique d'utiliser une échelle de comparaison; dans ce cas, on pourrait utiliser l'échelle à sept notes suivante:

TABLEAU 2

3 Bien meilleur
2 Meilleur
1 Légèrement meilleur
0 Identique
-1 Légèrement plus mauvais
-2 Plus mauvais
-3 Beaucoup plus mauvais

## 2 Présentation des résultats

Les résultats obtenus avec des équipes d'auditeurs spécialisés et les résultats obtenus par des équipes d'auditeurs non spécialisés doivent être présentés séparément. Toute précision doit être fournie sur les conditions d'écoute et les niveaux sonores; les méthodes statistiques d'analyse des résultats d'essai doivent être décrites.

NOTE 1 – L'Annexe 1 ci-après fournit quelques considérations générales sur les méthodes d'évaluation, les conditions d'écoute, le choix des auditeurs, etc.

---

\* Compte tenu de l'importante documentation existante sur des expériences faites avec des échelles à six notes, il est souhaitable de pouvoir convertir ces résultats en échelles à cinq notes, de manière que ces informations puissent être utilisables. Pour convertir les résultats d'une échelle dans une autre, on se heurte à des incertitudes. Cependant, on peut, en première approximation, utiliser la relation linéaire pour convertir une note  $A_6$  obtenue dans une expérience faite avec une échelle à six notes en une note  $A_5$  de l'échelle à cinq notes correspondante:

$$A_5 = 5,8 - 0,8 A_6$$

Lors de la présentation de résultats ainsi convertis, il convient d'indiquer que la conversion a été faite.

## ANNEXE 1

## 1 Généralités

Les séquences de programme d'essai doivent comprendre des silences pour permettre, en l'absence du signal utile, l'évaluation subjective de la dégradation causée par le bruit propre de la transmission. Par ailleurs, les essais doivent exclure toute évaluation de défauts dont les effets audibles, dans certains cas, pourraient ne pas être une gêne et même donner l'impression subjective d'une amélioration. Les séquences de programme doivent donc être exemptes de tout défaut audible qui soit semblable à l'un de ceux produits par le système en essai. Si cela est impossible, il convient d'indiquer clairement les limitations qui en découlent quant à la validité des résultats.

Pour les essais utilisant l'échelle à cinq notes du § 1.1, on doit utiliser un signal lumineux pour indiquer à tout moment à l'auditeur la source (dégradée ou non) du programme qu'il est en train d'écouter. Pour contrôler l'attention de l'auditeur et la reproductibilité de son jugement, on doit prévoir certains essais exécutés de manière aléatoire, sans prévenir l'auditeur, dans lesquels la séquence dégradée sera remplacée par une séquence non dégradée ou vice versa. Pour l'échelle de comparaison à sept notes, on ne donnera aucune indication qui puisse influencer le jugement de l'auditeur. Pour les essais comparatifs, il est cependant utile de donner de temps en temps une référence, qui peut être la séquence non dégradée; cette référence peut être indiquée par un signal lumineux.

La quantité de résultats qu'il est nécessaire de recueillir dépend de facteurs non indépendants tels que le degré de confiance dans les résultats statistiques, l'écart type des mesures et l'importance relative du défaut à déceler. Les suggestions suivantes, considérées comme des principes directeurs, ont pour but d'aider à concevoir un dispositif expérimental bien étudié.

## 2 Choix de l'équipe d'écoute

Bien que, dans un auditoire normal, il puisse se trouver des auditeurs spécialistes\*, leur proportion est vraisemblablement si faible qu'il convient de limiter les essais en laboratoire à des personnes non spécialistes, car le recours à des spécialistes pourrait conduire à des résultats beaucoup plus «critiques» que ceux obtenus avec des auditeurs ordinaires. Pour les conditions d'écoute, on doit être plus sévère qu'en temps normal, mais sans excès. Comme les essais avec des auditeurs non spécialistes ont tendance à être longs, un essai rapide peut être exécuté par des spécialistes. Un plus petit nombre d'auditeurs est suffisant dans ce cas. Toutefois, dans certaines circonstances, des essais avec des auditeurs spécialistes risquent de ne pas remplacer de façon satisfaisante des essais avec des non-spécialistes. En cas de doute, il convient de rechercher les causes de divergences entre les opinions exprimées par les spécialistes et par les non-spécialistes.

Le nombre minimal de non-spécialistes doit normalement être de vingt, et celui des spécialistes de dix. Dans tous les cas, il faut préciser le nombre et la catégorie des auditeurs ainsi que la durée des essais. Pour une transmission radiophonique ou une reproduction de haute qualité, il convient de faire appel uniquement à des auditeurs spécialistes.

---

\* L'expression «auditeurs spécialistes» s'applique aux auditeurs ayant eu récemment une grande expérience de l'évaluation de la qualité ou de la dégradation du son du type étudié dans les essais subjectifs.

### 3 Mode opératoire et durée des essais

La mémoire auditive à long ou à moyen terme étant fort imprécise, on opérera toujours par comparaison instantanée.

Pour les essais utilisant l'échelle à cinq notes, chaque opération de comparaison doit comprendre l'émission, quatre fois de suite, de la même séquence de programme, dans l'ordre suivant:

1. séquence originale,
2. même séquence, avec dégradation,
3. séquence originale (répétée),
4. même séquence, avec dégradation (répétée).

La durée de chaque séquence ne doit pas dépasser 15 à 20 s; elle peut être très courte (quelques secondes) pour certains essais. Dans le cas où la séquence est un élément musical, la phrase ne doit pas être interrompue. L'intervalle entre les présentations 1 et 2 et entre les présentations 3 et 4 doit être d'environ 0,5 à 1 s, tandis que l'intervalle entre 2 et 3 doit être légèrement plus long, par exemple 1,5 s. La durée exacte dépend du type de programme. Le dispositif de commutation instantané ne doit pas introduire de perturbation audible.

Les séquences de programme et dégradations doivent se succéder de façon aléatoire, à condition que la même séquence ne soit jamais émise deux fois de suite avec des dégradations identiques ou différentes.

Aucune séance d'écoute ne doit durer, pour un même auditeur, plus de 15 à 20 min environ sans interruption. Si de telles séances doivent se succéder, il faut les séparer par des périodes de repos d'une durée équivalente.

Pour les essais avec l'échelle de comparaison à sept notes utilisant deux types de dégradation, on pourra utiliser une présentation semblable, dans l'ordre suivant:

1. Type 1,
2. Type 2,
3. Type 1 (répété),
4. Type 2 (répété).

Les types 1 et 2 doivent être intervertis de façon aléatoire. On peut en outre présenter une condition de référence au début de chaque groupe de quatre présentations et, dans ce cas, on doit l'indiquer de manière efficace (par exemple, par un signal lumineux).

### 4 Choix des séquences de programme

Selon l'objectif précis qui a été fixé, et selon, en particulier, la catégorie à laquelle appartient le type de transmission radiophonique ou de reproduction en essai, on doit utiliser les séquences de programme suivantes:

- soit un choix représentatif de programme type,
- soit une sélection d'un petit nombre de séquences choisies spécialement pour leur comportement très critique à l'égard des défauts particuliers dus au système en essai. Par exemple, pour l'évaluation de rapports de protection, une séquence d'essai convenablement critique serait constituée par des paroles sur le programme utile, et une altération par de la musique «pop» sur le programme brouilleur.

Chaque fois que le système est destiné à acheminer un son de haute qualité, il convient d'utiliser le deuxième type de séquence de programme. Pour assurer la compatibilité des résultats d'expériences obtenus à des endroits et à des moments différents, il est préférable d'utiliser la même séquence. Le disque compact SQAM adopté et publié par l'UER constitue une source appropriée de programme numérique de haute qualité à partir de laquelle on peut choisir les extraits à utiliser à cette fin.

Dans tous les cas, le contenu artistique ou intellectuel de chaque séquence doit être tel qu'il ne risque pas de distraire l'auditeur soit par un caractère trop attrayant soit, au contraire, par un caractère désagréable ou fatigant.

## **5 Choix du dispositif de reproduction du son**

Selon la nature de la dégradation à évaluer, on peut utiliser des casques d'écoute ou des haut-parleurs.

Il a été démontré que certains défauts de qualité sont mieux perceptibles avec des casques d'écoute qu'avec des haut-parleurs. Par exemple, le rapport signal/bruit nécessaire pour une écoute sans bruit avec casques d'écoute peut dépasser de 10 dB la valeur obtenue avec des haut-parleurs, pour une même intensité sonore. On relève des différences du même ordre pour les pertes de qualité causées par les clics (erreurs binaires dans la transmission numérique), les distorsions de quantification, les distorsions dues aux non-linéarités, les distorsions de phase, etc.

Cependant, d'autres défauts de qualité sont mieux perceptibles dans la reproduction par haut-parleurs. En particulier, les effets qui influencent les caractéristiques de l'image sonore stéréophonique entre les haut-parleurs devraient être évalués avec reproduction par haut-parleurs. Ces effets sont dus, par exemple, aux pertes de qualité imputables aux différences qui peuvent exister entre les voies A et B.

Si l'on veut obtenir des évaluations comparables dans la plus grande mesure possible, il serait souhaitable de faire usage de casques d'écoute. La reproduction par casques est indépendante des propriétés géométriques et acoustiques des salles d'écoute et des salles de régie; en principe, il est donc possible de définir cette condition d'écoute avec plus de précision et elle est facilement reproductible sans erreurs systématiques. Il n'en est pas ainsi dans le cas de la reproduction par haut-parleurs.

De plus, avec des casques d'écoute, les essais d'évaluation peuvent être effectués avec un grand nombre d'auditeurs à la fois et dans des conditions d'écoute identiques.

## **6 Niveau sonore**

### **6.1 Reproduction par haut-parleurs**

Lorsqu'on utilise le signal utile à un niveau de crête élevé, le niveau sonore doit être mesuré avec un sonomètre sans pondération, avec la constante de temps dite «lente» normalisée par la CEI (Publication 123). Pour les autres signaux et pour mesurer le bruit ambiant, on utilisera un sonomètre avec la pondération *A* et avec la constante de temps dite «lente» normalisée par la CEI dans la Publication 123. Pour une séquence de programme dont on veut mesurer le niveau sonore dans les conditions particulières de l'expérience, et en un emplacement déterminé de la salle d'écoute, ce niveau sera pris, par définition, comme égal à la valeur maximale indiquée par le sonomètre au cours de chacune des séquences. Lors de l'évaluation de signaux de haute qualité à fort niveau, le niveau d'écoute sera de 80 à 90 dB.

Le niveau sonore pris en compte pour définir précisément les conditions dans lesquelles ont été faits les essais sera la moyenne des niveaux sonores mesurés aux différents emplacements occupés par les auditeurs. Par rapport à cette valeur moyenne, l'écart correspondant à un emplacement déterminé devra être aussi limité que possible. Une valeur de  $\pm 4$  dB apparaît raisonnable. Toutes les mesures devront être effectuées en présence des auditeurs.

## 6.2 Reproduction par casques d'écoute

Pour éviter de mesurer le niveau sonore dans le canal auditif, dans le cas de la reproduction par casques d'écoute, il y a lieu de régler ce niveau de telle manière que l'intensité sonore soit égale à celle d'un champ sonore de référence. Pour déterminer l'égalité des intensités sonores, l'auditeur doit être placé dans un champ sonore de référence comme indiqué au § 6.1.

Pour comparer l'intensité sonore dans les casques avec celle du champ sonore de référence, les signaux sont présentés à l'auditeur alternativement (et non simultanément). Les casques reçoivent un signal de même nature que celui du champ sonore de référence et ils sont réglés sur la même intensité sonore, selon l'appréciation des auditeurs.

Afin de garantir le niveau sonore adéquat dans les casques d'écoute, il faut utiliser la valeur moyenne de toutes les estimations comparatives de l'intensité sonore pour les essais à effectuer.

## 7 Conditions d'écoute

On s'efforcera en général de minimiser l'effet de masque dû au bruit ambiant, notamment pour établir les tolérances relatives à une transmission sonore de haute qualité.

Le niveau moyen du bruit ambiant devra toujours être indiqué. Et lorsque ce bruit est, de toute évidence, susceptible de produire un effet de masque non négligeable, son spectre moyen sera également indiqué.

Par ailleurs, toutes les précautions seront prises pour éviter au maximum que le ou les auditeurs ne soient gênés ou distraits par certains aspects de leur environnement (température, lumière, objets ou personnes en mouvement, etc.).

### 7.1 Reproduction par haut-parleurs

Chaque fois que l'écoute est faite à l'aide de haut-parleurs, on doit donner toute précision essentielle concernant les dimensions et le temps de réverbération de la salle, la disposition des auditeurs dans cette salle et la distance qui les sépare du ou des haut-parleurs.\*

Un document décrit les caractéristiques techniques des haut-parleurs utilisés en Fédération de Russie.

### 7.2 Reproduction par casques d'écoute

Chaque fois que l'écoute est faite à l'aide de casques, on doit donner toute précision essentielle concernant la désignation du type de casque.

Il convient de définir les spécifications techniques relatives aux caractéristiques des casques. L'UER a publié un texte qui propose un programme visant à établir une norme internationale applicable aux casques d'écoute de haute qualité.

---

\* Pour les propriétés acoustiques des salles d'écoute, voir le Rapport UIT-R BS.797.

## **8 Evaluation de caractéristiques particulières des équipements\*, des programmes, des studios, etc.**

### **8.1 Rapports de protection**

L'évaluation du rapport de protection exige une méthode d'essai légèrement différente. Dans ce cas, la séquence de programme non perturbée de référence doit être telle que la qualité de son reproduit par le récepteur soit appropriée au système de radiodiffusion pour lequel le récepteur est conçu.

### **8.2 Programmes enregistrés, studios, etc.**

Il n'existe aucune méthode uniforme pour évaluer la qualité des enregistrements. L'OIRT propose des méthodes de travail, notamment pour les enregistrements destinés à l'échange international de programmes (Recommandations N° 63/1 et 91 de l'OIRT) et des méthodes d'évaluation des propriétés acoustiques des studios de radiodiffusion et des salles de concert (Recommandation N° 68 de l'OIRT). Un document contient des renseignements sur les conditions appliquées en Fédération de Russie pour les évaluations subjectives de haute qualité (conditions d'écoute, choix de la méthode, nombre d'auditeurs et choix de ces auditeurs).

### **8.3 Applications de l'évaluation subjective de la qualité du son**

Dans des études effectuées en Fédération de Russie, on a cherché à définir des spécifications pour des essais subjectifs en radiodiffusion. Les applications de l'évaluation subjective ont été divisées en trois volets:

- enregistrement du son pour l'échange de programmes,
- studios, salles et autres lieux d'écoute,
- équipements.

Ces trois domaines ont encore été subdivisés en groupes qui sont présentés dans un tableau.

Les spécifications de l'évaluation sont fondées sur la pratique internationale, qui trouve son expression dans les textes de l'ISO, de la CEI et de l'OIRT; par exemple, pour le niveau de bruit, l'éclairage, les instructions, la position des haut-parleurs et des auditeurs, etc., dans la Publication 543 de la CEI; pour les protocoles, dans la Recommandation N° 68/1 de l'OIRT.

## **9 Evaluations subjectives des systèmes audio multidimensionnels**

Dans un document, il est indiqué que les méthodes décrites dans la présente Recommandation permettent de mesurer comme il convient les évaluations subjectives des défauts du signal causés par les non-linéarités, le brouillage ou le bruit.

Toutefois dans certains domaines, tels que ceux du son panoramique ou de la télévision à haute définition, le problème à étudier est plus complexe. Les évaluations subjectives sont nécessaires pour concevoir ou choisir un système audio à voies multiples et, dans ce cas, il faut tenir compte de nouveaux facteurs qui vont au-delà de la distorsion dans une voie unique. C'est ainsi qu'il faut tenir compte de la mesure dans laquelle la localisation est possible et de l'efficacité de la reproduction d'un champ sonore multidimensionnel en un point donné. On doit alors disposer de nouvelles méthodes d'évaluation qui vont au-delà de celles actuellement examinées dans la présente Recommandation.

---

\* Dispositifs et appareils tels que compresseurs, compresseurs-extenseurs, enregistreurs, etc.

Par exemple, on a examiné dans l'étude de Oghusi (NHK) qui utilise une échelle pluridimensionnelle les attributs suivants:

- largeur de l'image apparente,
- effet panoramique,
- taille apparente de la pièce,
- localisation verticale et horizontale,
- caractère naturel,
- réalisme,
- agrément,
- correspondance entre le son et l'image,
- adaptation de l'image sonore à l'image visuelle.

Les observateurs étaient invités à évaluer un certain nombre de systèmes différents pour ces attributs. On a établi un tableau des dissemblances dans lequel deux axes perceptuels perpendiculaires semblent correspondre au réalisme perçu (lien le plus fort avec le nombre de voies) et la coïncidence entre le son et l'image (lien le plus fort avec l'existence d'une source centrale). Dans le cas étudié par la NKH, le système a été choisi en fonction de la qualité globale et c'est ce critère que le Japon recommande d'adopter en pareils cas.

---