



UNIÓN INTERNACIONAL DE TELECOMUNICACIONES

UIT-T

P.830

SECTOR DE NORMALIZACIÓN
DE LAS TELECOMUNICACIONES
DE LA UIT

(02/96)

CALIDAD DE TRANSMISIÓN TELEFÓNICA MÉTODOS DE EVALUACIÓN OBJETIVA Y SUBJETIVA DE LA CALIDAD

EVALUACIÓN SUBJETIVA DE LA CALIDAD DE FUNCIONAMIENTO DE LOS CÓDECS DIGITALES DE BANDA TELEFÓNICA Y DE BANDA ANCHA

Recomendación UIT-T P.830

(Anteriormente «Recomendación del CCITT»)

PREFACIO

El UIT-T (Sector de Normalización de las Telecomunicaciones) es un órgano permanente de la Unión Internacional de Telecomunicaciones (UIT). Este órgano estudia los aspectos técnicos, de explotación y tarifarios y publica Recomendaciones sobre los mismos, con miras a la normalización de las telecomunicaciones en el plano mundial.

La Conferencia Mundial de Normalización de las Telecomunicaciones (CMNT), que se celebra cada cuatro años, establece los temas que han de estudiar las Comisiones de Estudio del UIT-T, que a su vez producen Recomendaciones sobre dichos temas.

La aprobación de Recomendaciones por los Miembros del UIT-T es el objeto del procedimiento establecido en la Resolución N.º 1 de la CMNT (Helsinki, 1 al 12 de marzo de 1993).

La Recomendación UIT-T P.830 ha sido revisada por la Comisión de Estudio 12 (1993-1996) del UIT-T y fue aprobada por el procedimiento de la Resolución N.º 1 de la CMNT el 6 de febrero de 1996.

NOTA

En esta Recomendación, la expresión «Administración» se utiliza para designar, en forma abreviada, tanto una administración de telecomunicaciones como una empresa de explotación reconocida de telecomunicaciones.

© UIT 1996

Es propiedad. Ninguna parte de esta publicación puede reproducirse o utilizarse, de ninguna forma o por ningún medio, sea éste electrónico o mecánico, de fotocopia o de microfilm, sin previa autorización escrita por parte de la UIT.

ÍNDICE

	<i>Página</i>
1 Alcance.....	1
2 Referencias.....	1
3 Definiciones.....	2
4 Abreviaturas.....	2
5 Convenios.....	3
6 Resumen de los procedimientos de prueba subjetivos.....	3
7 Grabaciones de señales en la fuente.....	3
7.1 Sistemas de grabación.....	3
7.2 Muestras vocales.....	4
8 Selección de los parámetros del experimento.....	5
8.1 Condiciones del códec.....	5
8.2 Condiciones de referencia.....	11
9 Plan del experimento.....	14
10 Procedimiento de la prueba de escucha.....	14
10.1 Sistema de recepción.....	14
10.2 Escalas de opinión.....	14
10.3 Ruido eléctrico.....	15
11 Análisis de los resultados.....	15
Anexo A – Comparación de las distintas definiciones de la relación señal/ruido.....	17
Anexo B – Conjunto de parámetros para determinar la calidad de funcionamiento de códecs.....	17
B.1 Sistemas de banda estrecha (300-3400 Hz).....	18
B.2 Sistemas de banda ancha (100-7000 Hz).....	18
Anexo C – Códec MIC correspondiente a una qdu.....	19
C.1 Interfaz analógica común.....	19
C.2 Interfaz digital.....	19
C.3 Interfaz analógica.....	19
Anexo D – Características de emisión y recepción del sistema IRS modificado.....	20
Bibliografía.....	21

SUMARIO

Esta Recomendación revisada describe los métodos y procedimientos para llevar a cabo evaluaciones subjetivas de la calidad de funcionamiento de códecs de señales vocales digitales.

Las revisiones efectuadas en esta versión de la Recomendación han consistido en la introducción de la nueva información que refleja las prácticas actuales en materia de evaluación subjetiva de la calidad de códecs digitales, incluida la ampliación de la sección relativa a la creación de registros en la fuente así como la adición de dos anexos. Uno de ellos describe la realización de un códec MIC (ley A/ μ) que genera una unidad de distorsión de cuantificación (qdu) en las señales de entrada. El otro anexo nuevo describe las características de transmisión y de recepción del sistema intermedio de referencia (IRS, *intermediate reference system*) modificado. Dichas características se recomiendan como las respuestas en transmisión y recepción que se deben utilizar en situaciones en las que está previsto emplear el códec sometido a prueba en circuitos totalmente digitales.

INTRODUCCIÓN

La telefonía digital en la red telefónica pública conmutada (RTPC) es actualmente de uso común. Desde la década de 1960, la transmisión digital en la RTPC es sinónimo de MIC en ley A o en ley μ a 64 kbit/s (véase la Recomendación G.711). La combinación de un mejor conocimiento de las técnicas de tratamiento de la señal junto con los avances tecnológicos, fundamentalmente las técnicas de integración a gran escala y a muy gran escala (LSI/VLSI, *large scale integration/very large scale integration*) suscitan un creciente interés por los métodos de codificación más eficaces que el de la MIC. Ejemplos de este desarrollo son la introducción del MICDA a 32 kbit/s (véase la Recomendación G.726), la LDCELP a 16 kbit/s (Recomendación G.728) y un códec de banda ancha de 7 kHz a 64 kbit/s (Recomendación G.722). Actualmente el interés está centrado en la codificación de las señales vocales a velocidades binarias inferiores a 16 kbit/s.

La prueba subjetiva es el método de evaluación de la calidad de funcionamiento de códecs digitales más ampliamente usado. Cuando el trayecto de transmisión es digital y/o no lineal, las mediciones objetivas simples, tales como las especificadas en la Recomendación G.712, son insuficientes para garantizar una calidad de funcionamiento de transmisión adecuada. El objeto de una metodología de prueba subjetiva consiste en medir la degradación causada por el tramo no lineal del trayecto de transmisión, de tal forma que la calidad de todo el sistema sea satisfactorio. Para ello, las mediciones:

- a) deben ser fiables; y
- b) deben realizarse de forma que se tengan en cuenta las interacciones principales entre la parte no lineal y la otra parte del sistema de transmisión.

Esto significa tener la posibilidad de asignar una contribución numérica única a cada proceso digital y la posibilidad de utilizar esta contribución asignada junto con otras degradaciones, para estimar la calidad de funcionamiento de las conexiones telefónicas.

PALABRAS CLAVE

Evaluación subjetiva, prueba subjetiva.

EVALUACIÓN SUBJETIVA DE LA CALIDAD DE FUNCIONAMIENTO DE LOS CÓDECS DIGITALES DE BANDA TELEFÓNICA Y DE BANDA ANCHA

(Helsinki, 1993; revisada en 1996)

1 Alcance

La presente Recomendación, está basada en el Anexo B/P.80 así como en la experiencia obtenida en diversos experimentos de carácter internacional [5], [6], [9], [10], [12], [13], [15] y [17]. Además, esta Recomendación define un método de prueba específico para evaluar los procesos digitales. Este método permite que los efectos subjetivos de la distorsión introducidos por el proceso de codificación de las señales vocales se tomen en cuenta cuando se evalúa la calidad de funcionamiento de la red telefónica internacional en constante evolución.

Esta Recomendación contiene indicaciones sobre la forma de evaluar la calidad de funcionamiento de los códecs digitales, y debe considerarse junto con la Recomendación P.80. En las referencias indicadas en «Bibliografía» se dispone de información adicional sobre este tema.

Para caracterizar la calidad de funcionamiento de los procesos digitales pueden utilizarse varios métodos. Actualmente dichos métodos consisten en pruebas de escucha solamente referentes a:

- 1) índices de opinión (categorías);
- 2) comparaciones por pares o múltiples; y
- 3) pruebas de nitidez.

Para la mayoría de las aplicaciones el UIT-T recomienda la utilización del método de evaluación de los índices por categorías absolutas (ACR, *absolute category rating*) que utiliza la escala de calidad de escucha. Sin embargo, a veces es más adecuado y conviene utilizar otras escalas y métodos de índices (por ejemplo, se ha constatado que la escala de esfuerzo en la escucha resulta útil cuando el interés se centra en obtener una buena correlación con las medidas de calidad de conversación), por lo que también se utilizan en esta Recomendación. Sólo se indicarán las desviaciones con respecto a la utilización del método de evaluación ACR que utiliza la escala de calidad de escucha.

NOTA – El método de evaluación de los índices por categorías de degradación (DCR, *degradation category rating*) [8] se escribe con detalle en el Anexo D/P.80. Se recomienda emplear este método cuando la calidad absoluta del material en la fuente es baja (por ejemplo, señales de voz afectadas por ruido de fondo) o cuando las degradaciones digitales son pequeñas. Por consiguiente, puede ser muy útil a la hora de optimizar el sistema. Cabe señalar igualmente que para optimizar los sistemas también puede aplicarse el método del umbral a fin de efectuar comparaciones directas; dicho método se escribe detalladamente en el Anexo E/P.80.

Además, puede existir la necesidad de complementar las pruebas de escucha solamente con pruebas de conversación planificadas y realizadas conforme a la Recomendación P.80. Sin embargo, hasta que no se consideren más detalladamente las circunstancias en las cuales son necesarias pruebas de conversación y se realice una nueva investigación sobre la elección de las condiciones que se deben abarcar en tales experimentos, no se pueden incluir planes detallados en esta Recomendación.

2 Referencias

Las Recomendaciones y demás referencias siguientes contienen disposiciones que, mediante su referencia en este texto, constituyen disposiciones de la presente Recomendación. Al efectuar esta publicación, estaban en vigor las ediciones indicadas. Todas las Recomendaciones y demás referencias son objeto de revisiones, por lo que se preconiza que todos los usuarios de la presente Recomendación investiguen la posibilidad de aplicar las ediciones más recientes de las Recomendaciones y demás referencias citadas a continuación. Se publica regularmente una lista de las Recomendaciones UIT-T actualmente vigentes.

- Norma CEI Publication 581.7 (1986), *High fidelity audio equipment and systems; Minimum performance requirements Part 7 – Loudspeakers*.
- Recomendación UIT-T G.113 (1996), *Degradaciones de transmisión*.
- Recomendación UIT-T G.191 (1993), *Herramientas de soporte lógico para la normalización de la codificación de señales vocales y de audio*.

¹⁾ Anteriormente P.83.

- Recomendación UIT-T G.192 (1996), *Interfaz paralela digital común para las actividades de normalización de señales vocales.*
- Recomendación G.711 del CCITT (1988), *Modulación por impulsos codificados (MIC) de frecuencias vocales.*
- Recomendación G.712 del CCITT (1992), *Características de transmisión de los canales de modulación por impulsos codificados.*
- Recomendación G.722 del CCITT (1988), *Codificación de audio de 7 kHz dentro de 64 kbit/s.*
- Recomendación G.726 del CCITT (1990), *Modulación por impulsos codificados diferencial adaptativa (MICDA) a 40, 32, 24 y 16 kbit/s.*
- Recomendación G.728 del CCITT (1992), *Codificación de señales vocales a 16 kbit/s utilizando predicción lineal con excitación por código de bajo retardo.*
- Recomendación UIT-T O.41 (1994), *Sofómetro para uso en circuitos de tipo telefónico.*
- Recomendación O.131 del CCITT (1988), *Aparato de medida de la distorsión de cuantificación mediante una señal de prueba de ruido pseudoaleatoria.*
- Recomendación O.132 del CCITT (1988), *Aparato de medida de la distorsión de cuantificación que utiliza una señal de prueba sinusoidal.*
- Recomendación UIT-T O.133 (1993), *Aparato de medida de la calidad de funcionamiento de los codificadores y decodificadores de modulación por impulsos codificados.*
- Recomendación UIT-T P.10 (1993), *Vocabulario de términos sobre calidad de transmisión telefónica y aparatos telefónicos.*
- Recomendación P.48 del CCITT (1988), *Especificación de un sistema intermedio de referencia.*
- Recomendación UIT-T P.50 (1993), *Voz artificial.*
- Recomendación UIT-T P.56 (1993), *Medición objetiva del nivel vocal activo.*
- Recomendación UIT-T P.80 (1993), *Métodos de determinación subjetiva de la calidad de transmisión.*
- Recomendación UIT-T P.810 (1996), *Aparato de referencia para ruido modulado.*
- Recomendación UIT-T P.84 (1993), *Método de prueba de escucha subjetiva para la evaluación de equipos de multiplicación de circuito digitales y de sistemas de voz por paquetes.*
- Recomendación Q.35 del CCITT (1988), *Características técnicas de los tonos para el servicio telefónico.*

3 Definiciones

A los fines de la presente Recomendación, se aplican las siguientes definiciones:

- 3.1 **dBov:** dB referidos al punto de sobrecarga de un sistema digital.
- 3.2 **Q:** relación, en dB, entre potencia de la señal vocal y potencia del ruido modulado en el aparato de referencia para ruido modulado descrito en la Recomendación P.810.
- 3.3 **Q_N:** relación Q para un aparato de referencia para ruido modulado de banda estrecha.
- 3.4 **Q_W:** relación Q para un aparato de referencia para ruido modulado de banda ancha.

4 Abreviaturas

A los fines de la presente Recomendación, se utilizan las siguientes abreviaturas:

ACR	Índices por categorías absolutas (<i>absolute category rating</i>)
BER	Tasa de errores en los bits (<i>bit error rate</i>)
DCR	Índices por categorías de degradación (<i>degradation category rating</i>)
eif	Factor de degradación de equipo (<i>equipment impairment factor</i>)

IRS	Sistema intermedio de referencia (<i>intermediate reference system</i>) (Recomendación P.48)
LDCELP	Predicción lineal con excitación por código de bajo retardo (<i>low-delay code excited linear prediction</i>)
MIC	Modulación por impulsos codificados
MICDA	Modulación por impulsos codificados diferencial adaptativa
MNRU	Aparato de referencia para ruido modulado (<i>modulated noise reference unit</i>)
MOS	Nota media de opinión (<i>mean opinion score</i>)
qdu	Unidad de distorsión de cuantificación (<i>quantization distortion unit</i>)
RMS	Valor cuadrático medio (<i>root mean square</i>)
RTPC	Red telefónica pública conmutada
SNR	Relación señal/ruido (<i>signal-to-noise ratio</i>)
UIT-T STL	Biblioteca de herramientas de soporte lógico (<i>software tools library</i>) del UIT-T

5 Convenios

La evaluación subjetiva de códecs de señales vocales digitales se puede efectuar, en principio, utilizando métodos de prueba subjetivos de conversación o de escucha solamente. En la práctica, las pruebas de escucha solamente son el único método viable de evaluación subjetiva durante la etapa experimental de códecs de señales vocales cuando no se dispone de una realización del códec en tiempo real. Esta Recomendación se ocupa de los métodos de evaluación subjetiva de escucha solamente de códecs de señales vocales.

6 Resumen de los procedimientos de prueba subjetivos

Cuando se efectúa una evaluación subjetiva de un códec de señales vocales, se debe tener en cuenta la aplicación a la que está destinado el códec sometido a prueba. El sistema utilizado para grabar material de conversación, las condiciones de prueba experimentales y el sistema utilizado para la entrega a los oyentes del material de conversación procesado, se han de elegir de tal modo que reflejen las principales características de la conexión telefónica de extremo a extremo. En particular, las características de los sistemas de grabación y recepción deben reflejar las características del transmisor y del receptor que serán utilizados en la aplicación prevista del códec.

La evaluación subjetiva de los códecs digitales se ha de llevar a cabo conforme a los siguientes pasos:

- 1) preparación del material de conversación en la fuente, incluida la grabación de los locutores;
- 2) selección de los parámetros del experimento para emplear las características del códec que sean de interés;
- 3) plan del experimento;
- 4) selección de un procedimiento de prueba y realización del experimento;
- 5) análisis de los resultados.

A continuación se describe cada uno de estos pasos.

7 Grabaciones de señales en la fuente

7.1 Sistemas de grabación

Como se indicó anteriormente, es importante utilizar un sistema representativo para el lado de emisión del sistema sometido a prueba. Cuando se prueban códecs digitales de banda estrecha (300-3400 Hz) para aplicaciones de red, el UIT-T recomienda la utilización de un sistema intermedio de referencia (IRS) modificado. Este dispositivo se construye eliminando del emisor del sistema IRS el filtro del sistema de referencia para la determinación de las atenuaciones equivalentes para la nitidez (SRAEN, *système de référence pour la détermination des affaiblissements équivalents pour la netteté*). La característica de emisión del sistema IRS modificado figura en el Anexo D. Se dispone de una realización de soporte lógico de la característica de transmisión del IRS modificado en la biblioteca de herramientas de soporte lógico del UIT-T (UIT-T_u STL, *software tools library*).

NOTA – La biblioteca de herramientas de soporte lógico del UIT-T contiene un número de elementos que son útiles para la elaboración de normas de codificación de señales vocales y de audio. Esas herramientas también son útiles en la preparación de materiales empleados en evaluaciones subjetivas de códecs vocales. Las herramientas que están disponibles en la STL del UIT-T se describen en la Recomendación G.191.

En algunas circunstancias (por ejemplo, cuando se espera que las frecuencias bajas estén fuertemente limitadas), será necesario o conveniente evaluar la calidad de funcionamiento del códec utilizando el lado de emisión del sistema intermedio de referencia (IRS) especificado en la Recomendación P.48 y calibrado de acuerdo con la Recomendación P.64. Sin embargo, se ha encontrado que la calidad de funcionamiento de los códecs vocales de baja velocidad binaria puede depender significativamente de la característica de frecuencia aplicada a la señal vocal de entrada. El sistema IRS es representativo, en particular, de las conexiones telefónicas analógicas (de la década de 1970). Por ello, conviene utilizar en la mayoría de las circunstancias la característica de emisión del sistema IRS modificado.

Si se va a utilizar para audioconferencia un sistema de banda ancha (100-7000 Hz), el lado de emisión debe cumplir las especificaciones de la Publicación 581.7 de la CEI.

7.2 Muestras vocales

El material de conversación debe consistir en frases sencillas, breves y con significado, elegidas al azar y fáciles de entender (por ejemplo, extraídas de publicaciones no técnicas o de periódicos). Además, se deben confeccionar grupos de dos o tres frases de tal manera que no haya ninguna conexión evidente entre el significado de las frases en un mismo grupo. Deben evitarse las frases muy cortas y muy largas ya que el objetivo es que la duración de pronunciación de cada frase sea de dos a tres segundos.

NOTA – Cuando se consideran los errores en los bits, esto permite, por ejemplo, evaluar tasas de errores en los bits (BER, *bit error rate*) de 0, 1:10000, 1:1000 y 1:100 (véase 8.1.4). Los valores de BER por debajo de 1:10000 necesitarán un número mayor de frases por muestra.

7.2.1 Producción del material en la fuente

Los locutores se deben sentar (uno cada vez) en una sala con un tiempo de reverberación menor que 500 ms, y un nivel de ruido de sala por debajo de 30 dBA sin crestas dominantes en el espectro. Las grabaciones se deben efectuar con un sistema de registro de alta calidad.

Los locutores participantes deben pronunciar las frases de manera fluida, pero sin efectos particulares, y adoptar un volumen de voz que les resulte confortable y que puedan mantener en forma constante, evitando producir ruidos como, por ejemplo, el crujido del papel. Las frases se pueden presentar al hablante una por una en un indicador visual controlado por ordenador, o en una lista ubicada a una distancia adecuada para fácil lectura.

Durante la grabación se debe observar un nivel vocal activo. Durante el proceso de grabación se debe mantener el nivel vocal activo de la grabación entre 20 y 30 dB por debajo del punto de sobrecarga del sistema de grabación para cada frase medida por separado. Cualquier frase en la que no se cumpla esta condición debe volverse a grabar.

NOTA – Durante la grabación se debe evitar la introducción de ruidos parásitos (por ejemplo, zumbido de la red de suministro de energía o chasquidos de conmutación).

7.2.2 Ecuación del nivel vocal

La medida de los niveles de la señal en la parte digital de la red internacional viene expresada normalmente en y dBm0 por los ingenieros de telecomunicaciones, es decir, el nivel relativo a 1 mW en 600 ohmios. Sin embargo, los materiales de conversación registrados digitalmente por ordenador se almacenan normalmente en forma de valores enteros. Es más práctico, por tanto, especificar los niveles de dicho material con respecto a la máxima potencia de la señal que pueda ser almacenada en un ordenador, por ejemplo, z dBov (donde los caracteres «ov» significan, por convención, el nivel digital de la señal de sobrecarga). Es necesario entender la relación de estos términos de tal modo que los niveles en dBm0 se puedan expresar en dBov. Esta relación tiene la forma:

$$z \text{ dBov} = y \text{ dBm0} + C \quad (7-1)$$

Para un sistema codificado de ley A, una señal sinusoidal que excita hasta el nivel máximo tiene una potencia dada por $T_{\text{máx}} = 3,14 \text{ dBm0}$ (véase la Recomendación G.711). De forma similar, para un sistema codificado de ley μ , $T_{\text{máx}} = 3,17 \text{ dBm0}$.

La potencia de referencia para dBov es el valor entero máximo para el sistema en uso. Una señal sinusoidal de amplitud máxima (por ejemplo 32 767 para un sistema de 16 bits) tendrá una potencia de $-3,01 \text{ dBov}$. De esta manera, $C = -6,15 \text{ dB}$ para sistemas de ley A, y $C = -6,18 \text{ dB}$ para sistemas de ley μ .

Un valor nominal típico de nivel vocal activo medio (medido conforme a la Recomendación P.56) es -20 dBm0, que corresponde aproximadamente a -26 dBov. Así pues, se debe medir el nivel vocal activo de cada frase, así como el nivel grabado después del ajuste será de -27 ± 1 dBov para el sistema digital, es decir, en formato digital. Para idiomas o voces en los que el nivel de la relación valor de cresta/valor medio es mayor que 23 dB, el nivel medio del material vocal se debe reducir consecuentemente. Se debe utilizar la realización de soporte lógico conforme a la Recomendación P.56, como se encuentra en la STL del UIT-T, para determinar el nivel vocal medio mientras está activo.

Por último, se combinan los grupos de frases (con un periodo de silencio adecuado entre éstas) para producir el material en la fuente final que será utilizado en la prueba subjetiva.

8 Selección de los parámetros del experimento

Las condiciones experimentales se dividen convenientemente en aquellas en las que interviene el códec sometido a prueba y las que sirven de referencia. Las condiciones del códec proporcionan la información esencial relativa a las características del mismo, mientras que las condiciones de referencias permiten efectuar comparaciones razonables cuando el mismo códec se prueba en distintos laboratorios o en el mismo laboratorio en momentos diferentes. Las Figuras 1 a 3 y 8 a 9 son ejemplos que muestran los efectos típicos sobre la nota MOS de los parámetros del experimento.

8.1 Condiciones del códec

Las condiciones de prueba descritas en esta subcláusula servirán para indicar cómo resulta afectada la calidad de funcionamiento del códec por variaciones en los niveles de entrada y de escucha, por diferentes locutores, por presencia de ruido de fondo, etc.

8.1.1 Niveles de entrada de la señal vocal

Es sabido que los códecs de señales vocales digitales son sensibles al nivel de las señales vocales de entrada y que su calidad de funcionamiento puede depender de este nivel. Cuando se evalúa la calidad de funcionamiento de los códecs es práctica habitual ajustar los niveles de entrada con relación al punto de sobrecarga del códec. Se recomienda utilizar niveles de entrada de 14, 26 y 38 dB por debajo del punto de sobrecarga del códec (equivalentes aproximadamente a valores de -8 , -20 y -32 dBm0) para evaluar la repercusión de las variaciones del nivel de entrada en la calidad de funcionamiento del códec. Esto es aproximadamente igual al valor medio medido en el punto de conmutación internacional \pm dos desviaciones típicas. Otras aplicaciones pueden exigir distintos niveles de entrada. En la Figura 1 se representa el efecto que producen las variaciones en el nivel de entrada de la señal vocal en la nota media de opinión (MOS, *mean opinion score*).

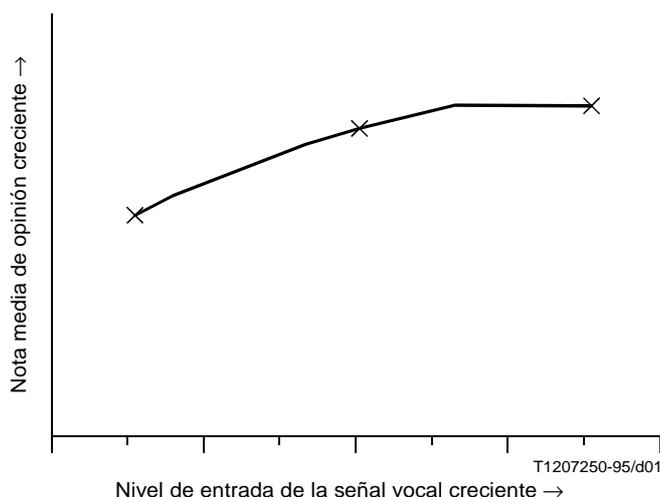


FIGURA 1/P.830

Notas medias de opinión en función de los niveles de entrada de la señal vocal

8.1.2 Niveles de escucha

Cuando se trata de evaluar el efecto del nivel en recepción se recomienda utilizar al menos tres niveles de escucha. Estos niveles no deben diferir en más de ± 10 dB del nivel de escucha preferido (tomado para este caso como de 79 dB de nivel de presión acústica (SPL, *sound pressure level*) en el plano de referencia oído). Conviene señalar que el nivel de escucha preferido está generalmente por debajo del «nivel óptimo de escucha», es decir, el nivel de escucha correspondiente al máximo valor de nota MOS, tal como se define en la Recomendación P.10.

En la Figura 2 se representa el efecto en la nota MOS que producen las variaciones en el nivel de escucha (así como la interacción entre el nivel de escucha con locutores masculinos y femeninos).

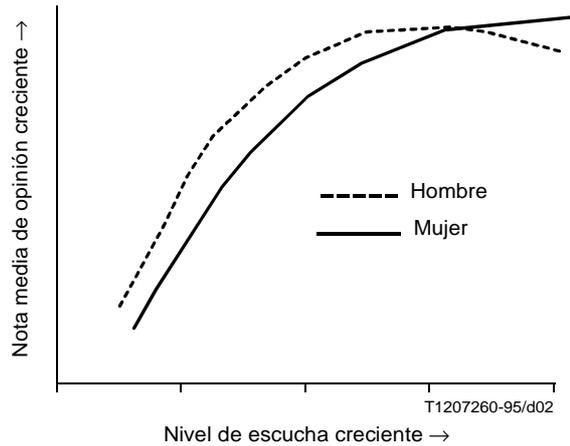


FIGURA 2/P.830

Nota media de opinión en función del nivel de escucha

8.1.3 Locutores

8.1.3.1 Diversidad de locutores

Se recomienda una participación mínima de dos hombres y dos mujeres. En la Figura 2 se representa el efecto de distintos participantes en la nota MOS. Sin embargo, si se va a probar la dependencia del locutor como un factor por sí mismo, se recomienda emplear más participantes, como sigue:

- 8 hombres;
- 8 mujeres;
- 8 niños.

8.1.3.2 Múltiples locutores

Es importante conocer la robustez del códec ante la presencia de diversas señales vocales de entrada simultáneas y determinar si aparecen efectos negativos tales como cortes de la transmisión, señales parásitas, etc.

Si sólo se quiere considerar la utilización del microteléfono, normalmente es suficiente la combinación de dos locutores distintos con niveles de entrada de la señal vocal separados 20 dB. Sin embargo, si el códec va a utilizarse en el modo conferencia; es decir, teléfono de manos libres, donde más de un participante pueda hablar al mismo tiempo, hay que asegurar que la presencia de múltiples locutores no afecte adversamente la calidad de funcionamiento del códec cuando la diferencia de nivel de las señales vocales pueda anularse.

Para evaluar los efectos de múltiples locutores se recomienda utilizar el método de índices por categorías de degradación (DCR) que hace uso de la escala de 5 notas del 10.2.3 (véase también el Anexo D/P.80) o el método de detectabilidad de respuesta cuantificada que utiliza la escala de 3 notas del apartado 10.2.4 (véase también el Anexo C/P.80).

8.1.4 Errores

La selección de las condiciones de error debe hacerse en el contexto de las aplicaciones previstas del códec. Si el códec se ha de utilizar en la RTPC (es decir, sistemas de líneas), los errores que aparecen en los bits deben estar distribuidos aleatoriamente. Si, en cambio, la aplicación prevista es un entorno radioeléctrico como el de las radiocomunicaciones móviles, los errores aparecerán aislados o en ráfagas y el efecto de estos errores será corromper tramas completas de las señales vocales codificadas en lugar de bits aislados.

Para los sistemas de línea se recomienda utilizar errores distribuidos aleatoriamente con una tasa de errores en los bits (BER) entre 0 y 1:1000. Evidentemente, dependerá del número de frases utilizadas en una muestra (véase 7.2). En algunas circunstancias es necesario efectuar pruebas con valores de BER de hasta 1:100. A tal fin se puede utilizar el dispositivo de inserción de errores (EID, *error insertion device*) de la biblioteca STL del UIT-T.

Para otras aplicaciones, como las de radiocomunicaciones móviles, los errores pueden ser del tipo ráfaga y, por consiguiente, convendría emplear errores de este tipo. A tal fin se puede utilizar también el EID de la STL del UIT-T.

Normalmente se utiliza el método de evaluación de índices ACR que hace uso de la escala de calidad de escucha del apartado 10.2.1. Sin embargo, en caso de que se suponga una baja calidad, puede ser más adecuado utilizar el método de evaluación de índices ACR que utilice la escala de esfuerzo de escucha del apartado 10.2.2.

8.1.5 Velocidades binarias

El códec debe ser probado para todas las velocidades binarias con las que es capaz de funcionar (para las especificaciones indicadas en la Recomendación G.722 estas velocidades son de 48 kbit/s, 56 kbit/s y 64 kbit/s y para las de la Recomendación G.726, de 16 kbit/s, 24 kbit/s, 32 kbit/s y 40 kbit/s). Sin embargo, el funcionamiento a ciertas velocidades binarias puede depender de las condiciones de explotación y la carga del sistema puede intervenir (véase la Recomendación P.84). La Figura 3 representa el efecto en la nota MOS de las variaciones de la velocidad binaria.

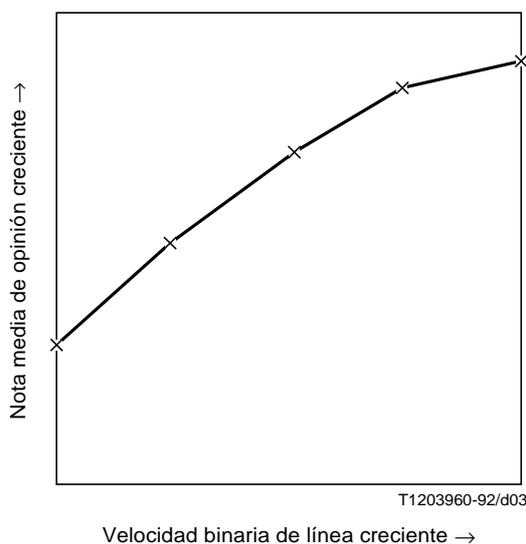


FIGURA 3/P.830
Nota media de opinión en función de la
velocidad binaria de línea

8.1.6 Transcodificaciones

Puesto que es muy probable que se utilicen códecs normalizados por el UIT-T en la RTPC (y en la arquitectura a 64 kbit/s asociada), los códecs vocales se deben evaluar en combinación con un proceso de compresión-expansión de ley A o de ley μ . Por consiguiente, un solo códec insertado en una conexión de un emisor analógico a un receptor analógico tendrá la configuración mostrada en la Figura 4.

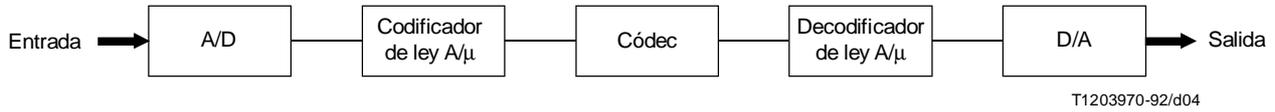


FIGURA 4/P.830
Esquema de transcodificación

8.1.6.1 Conexión asíncrona en cascada

Cuando el códec sometido a prueba se ha de utilizar en la RTPC mixta analógica/digital mixta, se pueden producir situaciones en las que el códec codifica y decodifica las señales vocales dos o más veces, es decir, diferentes segmentos digitales son intercalados con segmentos analógicos. En tales situaciones, la distorsión introducida por el códec crece con cada conversión sucesiva a analógico y de nuevo a digital con recodificación por el códec. Además, cuando el códec se instala en la red cada transcodificación incluye una distorsión de cuantificación adicional debida al proceso lineal de compresión-expansión en ley A o en ley μ , y a la distorsión de atenuación adicional debida a la tolerancia de los filtros de reconstrucción y de antisolapamiento más el ruido acumulado de canal en reposo. Esta configuración se representa en la Figura 5.

El UIT-T recomienda que se comprueben como mínimo uno, dos y tres códecs conectados en cascada. El método recomendado para efectuar la conexión asíncrona en cascada (para simular el proceso de conversión D/A (digital/analógico) y A/D (analógico/digital)) es utilizar el filtro de muestreo ascendente y descendente que figura en la biblioteca de herramientas de soporte lógico (STL) del UIT-T.

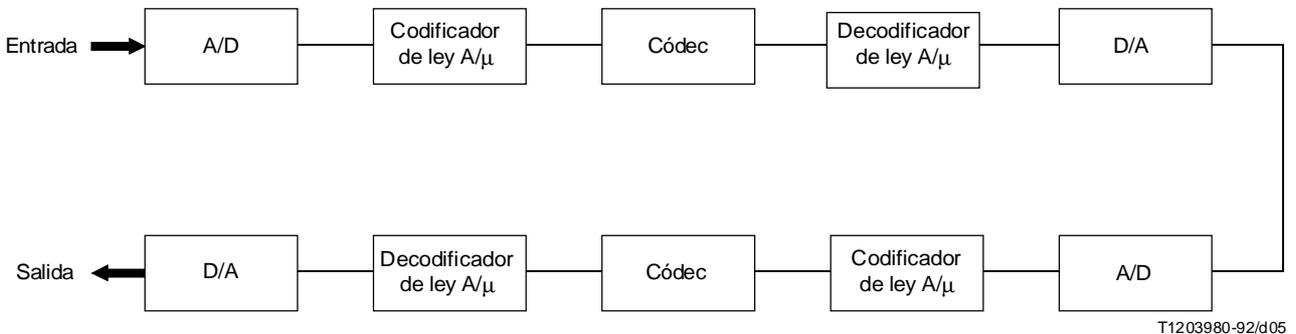


FIGURA 5/P.830
Conexión asíncrona en cascada

8.1.6.2 Conexión síncrona en cascada

La configuración del circuito mostrado en la Figura 6 representa situaciones similares a las del 8.1.6.1. En estas situaciones, las señales vocales se codifican y decodifican múltiples veces pero sin conversiones D/A o ley A/ley μ (es decir, en una red totalmente digital). Por ello, esta situación presenta la ventaja de que existe únicamente un proceso de compresión-expansión de ley A o ley μ y, por consiguiente, aparece una distorsión de cuantificación reducida en comparación con el caso anterior del 8.1.6.1. Sin embargo, continúa produciéndose la acumulación de la distorsión provocada por el códec sometido a prueba y debe ser evaluada.

El UIT-T recomienda que se comprueben en cascada como mínimo una, dos y tres codificaciones y decodificaciones sucesivas efectuadas por el códec.

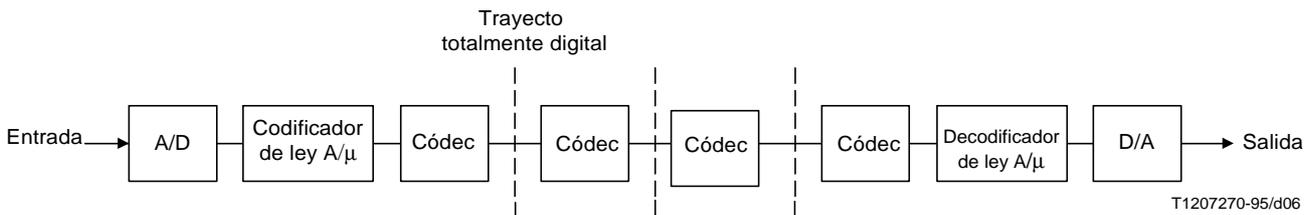


FIGURA 6/P.830

Ejemplo de conexión síncrona en cascada

8.1.6.3 Interfuncionamiento con otras normas de codificación de señales vocales

Es importante conocer los efectos de los sistemas en cascada que utilizan codificaciones a distintas velocidades binarias o diferentes métodos de codificación. En la Figura 7 se muestra un ejemplo de configuración. En [14] se pueden encontrar otros ejemplos de configuraciones de interfuncionamiento. Es fundamental verificar el comportamiento de las combinaciones de códecs más probables.

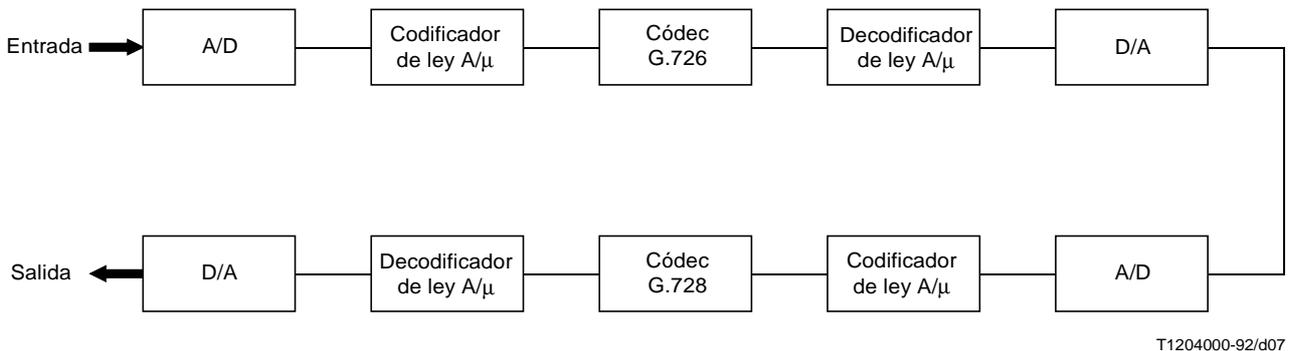


FIGURA 7/P.830

Interfuncionamiento de normas de codificación de señales vocales

Los siguientes ejemplos ilustran posibles combinaciones:

8-16 kbit/s →	64 kbit/s →	32 kbit/s →	64 kbit/s →	16-32 kbit/s →
Radiocomunicaciones móviles, aplicaciones aeronáuticas y locuciones grabadas	Circuitos de enlace	Equipos DCME de velocidad binaria variable	Circuitos de enlace	Teléfonos sin cordón

Equipo de multiplicación de circuitos digitales (DCME, *digital circuit multiplication equipment*).

Para los sistemas de banda ancha (por ejemplo, el especificado en la Recomendación G.722) es preciso comprobar estos sistemas conectados en cascada con sistemas de banda estrecha; por ejemplo, un sistema de la Recomendación G.722 conectado en cascada con uno de la Recomendación G.726, y viceversa.

8.1.7 Desadaptación

Algunos códecs de señales vocales (por ejemplo, los especificados en la Recomendación G.722) pueden funcionar a diversas velocidades binarias. En tales sistemas, es posible que el transmisor y el receptor funcionen en modos diferentes (es decir, diferentes velocidades binarias). Cuando esta situación es posible, es necesario investigar la calidad de funcionamiento de las combinaciones adecuadas de velocidad binaria en los dos extremos de la conexión. Por ejemplo, un códec que satisface las especificaciones de Recomendación G.722 debe ser probado para las siguientes condiciones en las que el codificador y decodificador funcionan a velocidades binarias diferentes:

<i>Velocidad binaria del transmisor</i>		<i>Velocidad binaria del receptor</i>
56 kbit/s	————→	64 kbit/s
48 kbit/s	————→	56 kbit/s
48 kbit/s	————→	64 kbit/s

8.1.8 Ruido ambiental (emisión)

Al igual que las consideraciones sobre múltiples locutores efectuadas en 8.1.3.2, la interacción del ruido ambiental con la voz del locutor es un factor importante, por las siguientes razones:

- los códecs de baja velocidad binaria pueden utilizar técnicas de codificación distintas de la codificación de forma de onda; y
- un códec puede funcionar en entornos ruidosos, como por ejemplo un vehículo en movimiento (si se emplea en radiocomunicaciones móviles) o en una oficina con mucho ruido ambiente.

Conviene efectuar el número de pruebas suficientes con el ruido apropiado (véase A.1.1.2.2/P.80) para examinar este efecto. Se recomienda utilizar las siguientes relaciones señal/ruido (véase en 8.2.3 las definiciones):

- 30 dB para ruido de sala;
- 10 dB y 20 dB para ruido en vehículos.

Teniendo en cuenta que la calidad percibida de la señal vocal afectada por el ruido ambiental será algo inferior a la de la señal vocal en un entorno de fondo silencioso, algunos laboratorios han descubierto que el método de evaluación de índices ACR que utiliza la escala de esfuerzo de escucha (véase 10.2.2) es especialmente útil, ya que está basado en la capacidad de comprender el significado de las frases. Esos laboratorios han constatado que la escala de esfuerzo de escucha dan resultados que están mejor correlacionados con las notas medias de opinión que con las notas de calidad de escucha.

Las técnicas modernas de codificación de señales vocales procuran suprimir o eliminar el ruido de fondo. De esta manera, es posible que el material de conversación que se ha visto afectado por el ruido de fondo, y posteriormente procesado por un códec, tenga mejor calidad que las señales de entrada originales (es decir, las señales vocales afectadas por el ruido). En estas situaciones, se ha visto que la escala de comparación (véase 10.2.5) resulta de una gran utilidad.

Los ruidos ambientales registrados con micrófonos de estudio deben ser filtrados para compensar la diferencia entre fuentes difusas (campo lejano) y fuentes puntuales (campo cercano). En la biblioteca STL del UIT-T se dispone de una realización de soporte lógico de un filtro de este tipo. Las relaciones SNR definidas anteriormente han de ser medidas después de este proceso de filtrado.

NOTAS

1 El ruido puede combinarse eléctricamente con los registros magnéticos de señales fuente, de tal forma que se conozcan con precisión los niveles de ruido y de la señal vocal. No se recomienda efectuar grabaciones de señales fuente en un entorno ruidoso.

2 No debe suponerse de forma automática que la mezcla de señales vocales grabadas por separado y de un nivel elevado de ruido ambiental producirá los mismos efectos que un hablante en un entorno ruidoso, porque las personas suelen adaptar sus características vocales (nivel y espectro) así como su comportamiento en el habla cuando se encuentran en un entorno de esas características. Como algunos códecs pueden procesar las distintas partes del espectro de audiofrecuencias de forma diferente, esta modificación de las características vocales puede afectar a la calidad de funcionamiento.

3 Algunos laboratorios han utilizado una técnica para grabar muestras de señales vocales que aprovecha el hecho de que los locutores ajustan sus niveles de conversación en ambientes ruidosos, pero esta técnica evita los problemas asociados con la grabación en ambientes ruidosos. Los participantes emplean auriculares telefónicos en los que el ruido de fondo se regula a un nivel conocido. La señal vocal del hablante se graba de manera normal, como si se encontrara en un entorno ruidoso. Seguidamente, el ruido se adiciona electrónicamente a la señal vocal grabada a un nivel deseado con relación a la señal vocal.

8.1.9 Señales de información de red

En toda red nacional hay un gran número de tonos o señales de información destinados al usuario y algunos que son instrucciones para los equipos de la red. Estas señales pueden ser originadas en la RTPC o en las redes privadas conectadas a la misma. Es importante que la degradación de estas señales, al atravesar algún dispositivo de procesamiento de la palabra, no las haga irreconocibles para el usuario o para el equipo que debe actuar de una determinada manera al recibirlas. Es posible que esta última situación sea menos tolerante a la degradación que la primera cuando se trata de determinadas señales, tales como los tonos de señalización multifrecuencia (DTMF, *dual tone multi frequency*), pero presenta la ventaja de que es suficiente realizar pruebas objetivas sencillas para determinar el límite admisible de dicha degradación.

Se recomienda probar subjetivamente los tonos de señalización originados en la red conformes a la Recomendación Q.35; como mínimo deben comprobarse los siguientes tonos:

- tono de invitación a marcar;
- tono de llamada del abonado llamado;
- tono de ocupado del abonado llamado;
- tono de ocupación del equipo;
- tono de número inalcanzable.

Para evaluar la idoneidad (reconocimiento) de las señales de información se recomienda utilizar el método de índices por categorías de degradación (DCR) con escala de 5 notas del 10.2.3 (véase también el Anexo D/P.80).

8.1.10 Música

Algunos equipos utilizados en la RTPC, especialmente las centralitas privadas (PABX, *private automatic branch exchange*), tienen la posibilidad de emitir música al abonado mientras éste espera. El UIT-T recomienda utilizar únicamente pruebas sencillas, para asegurarse de que la música presente una calidad razonable.

8.2 Condiciones de referencia

Las condiciones de referencia proporcionan un medio práctico para efectuar comparaciones significativas de los resultados de las pruebas subjetivas de diferentes laboratorios o del mismo laboratorio en momentos diferentes. Estas condiciones incluyen la condición *mejor posible* así como condiciones a las que se han agregado degradaciones conocidas y controladas a los materiales de conversación.

8.2.1 Condición directa

La condición «directa» representa la mejor condición obtenible en el experimento y teóricamente equivale a un valor infinito de relación Q_N o Q_W (véase 8.2.2). Esta condición no tiene, en particular, codificación de la señal vocal de entrada; sólo tiene el mismo filtrado de entrada, el mismo nivel de entrada y el mismo filtrado de salida que la señal vocal que ha sido procesada por el códec sometido a prueba.

8.2.2 Aparato de referencia para ruido modulado (MNRU)

El ruido aleatorio con una amplitud proporcional a la amplitud de la señal instantánea en términos de relación Q_N o Q_W , de acuerdo con el aparato MNRU especificado en la Recomendación P.810, debe utilizarse como sistema de referencia para expresar la calidad de funcionamiento subjetiva de los procesos digitales, por las siguientes razones:

- a) para códecs de forma de onda y posiblemente para algunos códecs distintos de los de forma de onda, la señal procesada a través del aparato MNRU es de percepción muy similar a la señal procesada por estos códecs, lo que facilita, en principio, la evaluación por los participantes en la prueba;
- b) la experiencia ha demostrado que el aparato MNRU es una norma de transferencia útil y permite realizar comparaciones razonables entre distintos laboratorios y entre experimentos efectuados en el mismo laboratorio pero en momentos diferentes; y
- c) se dispone ya de una información y experiencia considerables en lo que respecta al aparato MNRU.

En la prueba subjetiva se debe incluir una gama de valores para la relación Q_N o Q_W . Esta gama para los sistemas de banda estrecha debe abarcar valores de Q_N entre 5 dB y 35 dB (preferiblemente de 5 a 7 valores distintos) y para los sistemas de banda ancha debe ser para valores de Q_W entre 10 dB y 45 dB (también de 5 a 7 valores distintos).

NOTA – La distorsión introducida por el aparato MNRU es similar subjetivamente a la distorsión de cuantificación. Estas distorsiones no son necesariamente similares, cuando se procede a una evaluación subjetiva, a la distorsión producida por las técnicas modernas de codificación de la voz. Las degradaciones de referencia que producen distorsiones muy similares a las degradaciones debidas a las técnicas modernas de codificación se encuentran actualmente en estudio.

En la Figura 8 se representa el efecto de la variación de la relación Q_N o Q_W en la nota MOS.

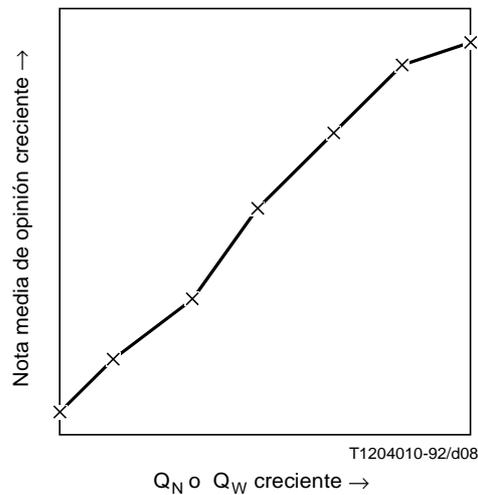


FIGURA 8/P.830

Nota media de opinión en función de Q_N o Q_W

8.2.3 Relación señal/ruido (SNR)

Las administraciones y las compañías de explotación han considerado útil relacionar los efectos de las degradaciones en términos de SNR.

El UIT-T recomienda utilizar las siguientes definiciones de la relación señal/ruido, medida en las conexiones, en lo que respecta al ruido en régimen estacionario:

- *Medidas telefónicas (de banda estrecha) – Ponderación sofométrica* (véase la Recomendación O.41)
 $SNR(p)$ = Nivel vocal activo (véase la Recomendación P.56)/medida del ruido ponderado sofométricamente
- *Medida de banda ancha – Ponderación A* (véase la Recomendación P.54)
 $SNR(A)$ = Nivel vocal activo (véase la Recomendación P.56)/medida del ruido con ponderación A

Si se utilizan definiciones distintas de las indicadas, debe adoptarse el siguiente sistema de notación:

- *Banda estrecha 300-3400 Hz – Sin ponderación*
 $SNR(N)$ = Nivel vocal activo (véase la Recomendación P.56)/medida del ruido sin ponderación
- *Banda ancha 100-7000 Hz – Sin ponderación*
 $SNR(W)$ = Nivel vocal activo (véase la Recomendación P.56)/medida del ruido sin ponderación
- *Ponderación de mensaje C* (véase la Recomendación O.41)
 $SNR(C)$ = Nivel vocal activo (véase la Recomendación P.56)/medida del ruido con ponderación de mensaje C
- *Anchura de banda 100-5000 Hz*
 $SNR(0,1-5 \text{ kHz})$ = Nivel vocal activo (véase la Recomendación P.56)/medida del ruido sin ponderación

NOTA – Si se utilizan otras anchuras de banda, conviene sustituir «0,1-5 kHz» por la anchura de banda de medida.

En el Anexo A figura una comparación de las distintas relaciones señal/ruido establecidas según las definiciones anteriores y para diferentes tipos de espectros de ruido y de otras características sensibilidad/frecuencia de emisión.

Si en la prueba se emplea ruido gaussiano se sugiere para los sistemas de banda estrecha y de banda ancha una gama de SNR entre 15 y 45 dB.

En la Figura 9 se representa el efecto típico de la variación de la relación señal/ruido en la nota MOS.

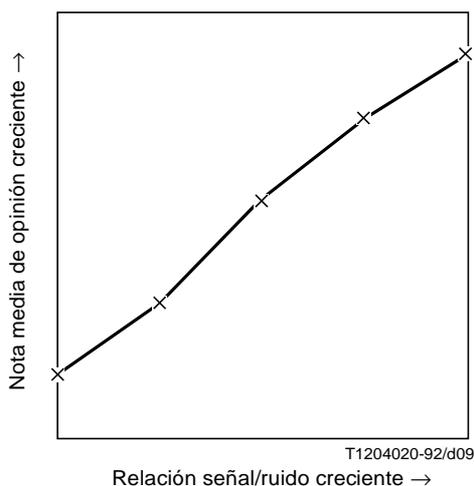


FIGURA 9/P.830

Nota media de opinión en función de la relación señal/ruido

8.2.4 Códex de referencia

Los códecs de referencia tienen dos objetos:

- 1) pueden utilizarse para determinar las reglas de planificación, ya sea en forma de unidades de distorsión de cuantificación (qdu) o de factores de degradación de equipo (eif) (véase la Recomendación G.113 para información sobre qdu y eif); y
- 2) pueden utilizarse como norma para juzgar la calidad de funcionamiento global en función de los parámetros (véase 8.1).

Si se emplean códecs de referencia para determinar las reglas de planificación, el UIT-T recomienda utilizar una conexión en cascada asíncrona (véase 8.1.6.1) de 1, 2, 4, 8 y 16 códecs de ley A o μ , conforme a las Recomendaciones G.711 y G.712. En el Anexo C figura la especificación para un códec G.711 de referencia.

Si se emplean para determinar la calidad relativa, por ejemplo con respecto a la indicada en las Recomendaciones G.726 o G.722, deben aplicarse las consideraciones efectuadas en los apartados 8.1.6.1 y 8.1.6.2.

9 Plan del experimento

Se aplican las consideraciones indicadas en B.3/P.80.

La utilización de todas las combinaciones de los parámetros descritos en la cláusula 7 daría lugar a un experimento demasiado largo desde el punto de vista logístico. Se recomienda realizar un conjunto mínimo de experimentos que, aunque no comprendan todas las combinaciones posibles, ofrezcan datos suficientes como para tomar decisiones con conocimiento de causa. En el Anexo B figura un conjunto de experimentos que el UIT-T ha encontrado adecuados para los estudios de los códecs de banda estrecha y de banda ancha; esos estudios sirven de base para la formulación de las Recomendaciones de la serie G.700.

Deben extremarse las precauciones cuando se comparan, incluso en una misma prueba, sistemas que presentan degradaciones muy distintas; por ejemplo, códecs digitales, sistemas múltiplex por división de frecuencia, vocodificadores, etc.

10 Procedimiento de la prueba de escucha

Esta cláusula es idéntica a la B.4/P.80 con las siguientes excepciones.

10.1 Sistema de recepción

Cuando se prueban códecs digitales de banda estrecha (300-3400 Hz) para aplicaciones de red, el UIT-T recomienda la utilización de una característica de recepción del sistema intermedio de referencia (IRS) modificado. Al igual que en el caso de la característica de emisión del sistema IRS modificado, este dispositivo se construye retirando el filtro SRAEN del receptor del IRS. La característica de recepción del IRS modificado figura en el Anexo D. En la STL del UIT-T se dispone de una realización de soporte lógico de la característica de recepción del IRS modificado. En algunas circunstancias (por ejemplo, cuando se espera que las bajas frecuencias sean limitadas fuertemente), será necesario o conveniente evaluar la calidad de funcionamiento del códec utilizando el extremo receptor del IRS, conforme a la Recomendación P.48 y calibrado según la Recomendación P.64.

Si va a utilizarse un sistema de banda ancha (100-7000 Hz) para audioconferencia, el extremo receptor debe ser conforme a la publicación 581.7 de la CEI.

NOTA – Algunos laboratorios utilizan cascos telefónicos para la presentación del material de conversación a los oyentes durante las evaluaciones subjetivas. Cuando se utilizan cascos telefónicos para la presentación de material de conversación, se debe tomar la precaución de asegurarse de que la presentación proporciona el filtrado adecuado al objeto de simular los efectos de los receptores reales, incluyendo, cuando proceda, los efectos de la fuga acústica. Para mayor información sobre la utilización de cascos telefónicos en lugar de con los micrófonos de los aparatos telefónicos, véase [11].

10.2 Escalas de opinión

Para evaluar los procesos digitales, se recomienda utilizar las siguientes escalas de opinión.

10.2.1 Escala de calidad de escucha

Calidad de la señal vocal:

- 5 Excelente
- 4 Buena

- 3 Regular
- 2 Mediocre
- 1 Mala

10.2.2 Escala de esfuerzo de escucha

Esfuerzo necesario para comprender el significado de las frases:

- 5 Audición perfecta; ningún esfuerzo.
- 4 Es necesaria una cierta atención; ningún esfuerzo apreciable.
- 3 Es necesario un esfuerzo moderado.
- 2 Es necesario un esfuerzo considerable.
- 1 Significado incomprensible, aun con el mayor esfuerzo.

10.2.3 Escala de categorías de degradación

- 5 Degradación inaudible.
- 4 Degradación audible pero no molesta.
- 3 Degradación ligeramente molesta.
- 2 Degradación molesta.
- 1 Degradación muy molesta.

10.2.4 Escala de opiniones sobre la detectabilidad

- 3 Objetable
- 2 Detectable
- 1 No detectable

10.2.5 Escala de comparación

El segundo comparado con el primero es:

- 3 Mucho mejor
- 2 Mejor
- 1 Ligeramente mejor
- 0 Casi igual
- 1 Ligeramente peor
- 2 Peor
- 3 Mucho peor

10.3 Ruido eléctrico

Se debe añadir un ruido gaussiano equivalente de -68 dBmp a la entrada del sistema de recepción para disminuir los efectos del contraste de ruido al comienzo de la pronunciación de las frases.

11 Análisis de los resultados

Se aplican las consideraciones indicadas en B.4.7/P.80.

Uno de los objetivos del análisis es determinar una función $Q_2 = F(L)$, donde Q_2 es el valor equivalente Q para el códec y L es la velocidad binaria de línea. Un método sencillo para calcular esta función consiste en utilizar las notas medias de opinión indicados en las Figuras 3 y 8 y trazar un gráfico de esta función, como el representado en la Figura 10. Este método se ilustra en la Figura 11, donde se ha elegido un valor de velocidad binaria de línea, por ejemplo L_2 , y se ha determinado la nota media de opinión correspondiente. Esta nota media de opinión se utiliza para entrar en el gráfico de la derecha a fin de encontrar el valor de Q , en este caso Q_2 , correspondiente a este valor de nota media de opinión. Los valores de Q para los demás valores de L se obtienen de forma similar, y el conjunto resultante de pares (L_i, Q_i) da lugar a un gráfico como el indicado en la Figura 10.

NOTA – Las operaciones anteriores son más fiables cuando se efectúan sobre curvas ajustadas en lugar de segmentos rectos que unen los puntos observados. Los métodos para calcular la curva que mejor se adapta a las notas MOS en función de Q se encuentran en estudio. Como procedimiento provisional, mientras se esperan los resultados de este estudio, se señala un método de ajuste sigmoide que tiene en cuenta el «efecto de saturación» a valores elevados de Q, y que ha sido utilizado con cierto éxito en el trabajo efectuado por el ETSI sobre los posibles métodos de evaluación para el sistema «Half-Rate GSM» [2].

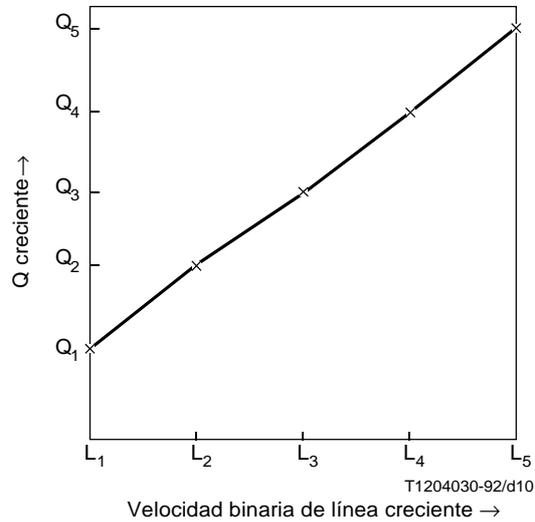


FIGURA 10/P.830
Q en función de la velocidad binaria de línea

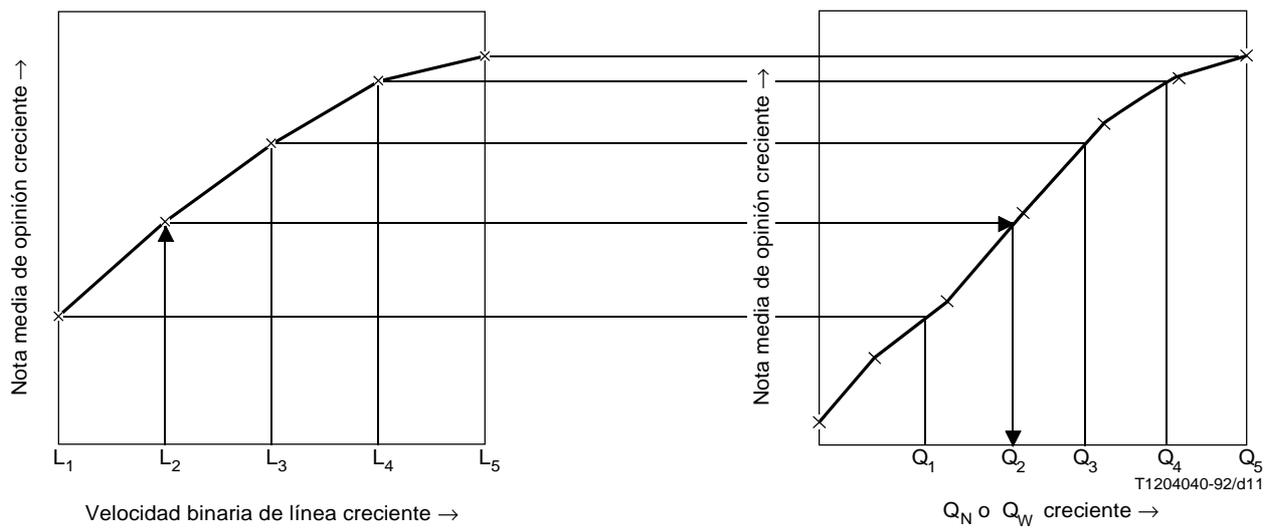


FIGURA 11/P.830
Método gráfico para obtener la Figura 10 a partir de las Figuras 3 y 8

Anexo A

Comparación de las distintas definiciones de la relación señal/ruido

(Este anexo es parte integrante de la presente Recomendación)

En el Cuadro A.1 aparecen algunos resultados de los cálculos efectuados con las ponderaciones descritas en la Recomendación O.41, sobre espectros conocidos que emplean dos tipos de respuesta en frecuencia distintos. Estos espectros se han elegido para representar los que se utilizan normalmente en las pruebas subjetivas.

- Ruido blanco (gaussiano) coherente con el ruido de circuito (sofométrico y de mensaje C) destinado a medir los efectos de este tipo de ruido.
- Ruido de Hoth (de sala) y ruido en el interior de vehículos; se utilizan como ruido ambiental en las pruebas subjetivas (véase A.1.1.2.2/P.80) y representan el ruido o los ruidos captados por el micrófono y medidos en los terminales de la línea telefónica.

CUADRO A.1/P.830

Respuesta	Ponderación	Ruido blanco	Ruido de Hoth	Ruido en el interior de vehículos	Vehículo estacionado
Respuesta plana (banda estrecha)	Sin ponderación	0	0	0	0
	Mensaje C	-1,9	-3,6	-6,4	-4,2
	Sofométrica	-2,6	-2,6	-4,1	-2,8
Respuesta del IRS (banda estrecha)	Sin ponderación	0	0	0	0
	Mensaje C	-1,9	-2,0	-3,7	-2,3
	Sofométrica	-3,6	-2,3	-2,4	-2,3
Respuesta plana (banda ancha)	Sin ponderación	0	0	0	0
	Mensaje C	-5,2	-7,3	-16,5	-12,7
	Sofométrica	-5,9	-6,3	-14,0	-11,3
Respuesta del IRS (banda ancha)	Sin ponderación	0	0	0	0
	Mensaje C	-1,9	-2,1	-4,4	-2,5
	Sofométrica	-3,6	-2,4	-3,1	-2,5

NOTAS

- 1 Todos los valores indican la diferencia en dB con respecto al ruido sin ponderación.
- 2 El signo negativo significa más silencioso (menos ruidoso).
- 3 La respuesta del IRS que aquí se utiliza se refiere a la Recomendación P.48.
- 4 El cálculo del ruido de banda estrecha sin ponderación se ha efectuado en bandas de 1/3 de octava centradas en las frecuencias preferidas definidas en R.266 de la ISO, desde 315 Hz a 3150 Hz inclusive.
- 5 El cálculo del ruido de banda ancha sin ponderación se ha efectuado en bandas de 1/3 de octava centradas en las frecuencias preferidas definidas en R.266 de la ISO, desde 100 Hz hasta 6300 Hz inclusive.
- 6 El cálculo de las ponderaciones de mensaje C y sofométricas se ha efectuado en bandas de 1/3 de octava centradas en las frecuencias preferidas definidas en R.266 de la ISO, desde 100 Hz hasta 5000 Hz inclusive.
- 7 Debe tenerse cuidado al suponer que la ponderación de mensaje C es equivalente a la ponderación sofométrica, puesto que ello sólo es cierto para ruido de tipo gaussiano (véase la Recomendación O.41).

Anexo B

Conjunto de parámetros para determinar la calidad de funcionamiento de códecs

(Este anexo es parte integrante de la presente Recomendación)

En este anexo figuran ejemplos de los experimentos que se deben efectuar para evaluar la calidad de funcionamiento subjetiva de los códecs digitales. Como se indicó en la cláusula 6, es importante ampliar estos experimentos con condiciones y/o experimentos que prueben el códec en las condiciones especiales de su aplicación prevista.

B.1 Sistemas de banda estrecha (300-3400 Hz)

a) Experimento 1 – Efectos de los errores, del nivel de entrada y del nivel de escucha

Niveles de entrada de la señal vocal:	3 (14, 26 y 38 dB por debajo del nivel de sobrecarga)
Niveles de escucha:	3 (P + 10, P, P – 10)
Tasas de error:	3 (0, 1:10000, 1:1000)
Transcodificaciones:	1 transcodificación
Niveles de ruido ambiental (emisión):	1 (< 30 dBA)

b) Experimento 2 – Efectos de la transcodificación, del nivel de entrada y del nivel de escucha

Niveles de entrada de la señal vocal:	3 (14, 26 y 38 dB por debajo del nivel de sobrecarga)
Niveles de escucha:	3 (P + 10, P, P – 10)
Transcodificaciones:	«x»
Tasas de error:	1 (1:1000)
Niveles de ruido ambiental (emisión):	1 (< 30 dBA)

c) Experimento 3 – Efectos del ruido ambiental, del ruido de sala, del nivel de entrada y del nivel de escucha

Niveles de entrada de la señal vocal:	3 (14, 26 y 38 dB por debajo del nivel de sobrecarga)
Niveles de escucha:	3 (P + 10, P, P – 10)
Transcodificaciones:	1 transcodificación
Tasas de error:	1 (1:1000)
Niveles de ruido ambiental (emisión):	2 (< 30 dBA e «y»)
Niveles de ruido de sala:	«z»

donde: «P» es el nivel de escucha preferido;

«x» es el número de combinaciones de transcodificación que van a probarse;

«y» es el nivel de ruido de emisión que va a probarse;

«z» es el número de condiciones de ruido de sala (normalmente 2).

Todos los experimentos deben incluir igualmente las condiciones 5 a 7 del aparato MNRU de banda estrecha.

Estos tres experimentos no pretenden ser exhaustivos y deben completarse con otros para caracterizar más adecuadamente el códec.

B.2 Sistemas de banda ancha (100-7000 Hz)

a) Experimento 1 – Efectos de la velocidad binaria, del BER, del nivel de entrada y del nivel de escucha

Niveles de entrada de la señal vocal:	2 (20 y 38 dB por debajo del nivel de sobrecarga)
Niveles de escucha:	3 (P + 10, P, P – 10)
Transcodificaciones:	1 transcodificación
Niveles de ruido ambiental (emisión):	1 (< 30 dBA)
Velocidades binarias:	«r» (por ejemplo 48, 56 y 64 kbit/s para G.722)
BER:	3 (0, 1:10000 y 1:1000)
Niveles de ruido de sala:	1 (< 30 dBA)

b) Experimento 2 – Efectos de la transcodificación, del nivel de entrada y del nivel de escucha

Niveles de entrada de la señal vocal:	2 (20 y 38 dB por debajo del nivel de sobrecarga)
Niveles de escucha:	3 (P + 10, P, P – 10)
Transcodificaciones:	«x», incluidas síncronas y asíncronas
Niveles de ruido ambiental (emisión):	1 (< 30 dBA)
Velocidades binarias:	«r» (por ejemplo 48, 56 y 64 kbit/s para G.722)

BER:	3 (0, 1:10000 y 1:1000)
Niveles de ruido de sala:	1 (< 30 dBA)
c) <i>Experimento 3 – Efectos de la desadaptación, del nivel de entrada y del nivel de escucha</i>	
Niveles de entrada de la señal vocal:	2 (20 y 38 dB por debajo del nivel de sobrecarga)
Niveles de escucha:	3 (P + 10, P, P – 10)
Transcodificaciones:	1 transcodificación
Niveles de ruido ambiental (emisión):	1 (< 30 dBA)
Velocidades binarias:	«m» (56 → 64, 48 → 56, 48 → 64 kbit/s)
BER:	2 (0 y 1:1000)
Ruido de sala:	1 (< 30 dBA)
donde:	«P» es el nivel de escucha preferido;
	«x» es el número de combinaciones de transcodificación que van a probarse;
	«r» es el número de bits con que funciona el códec;
	«m» es el número de combinaciones entre velocidades binarias diferentes a las que el códec debe funcionar.

Estos experimentos deben incluir igualmente las condiciones del aparato MNRU de banda ancha.

Estos tres experimentos no pretenden ser exhaustivos y deben completarse con otros experimentos para caracterizar más adecuadamente el códec.

Anexo C

Código MIC correspondiente a una qdu

(Este anexo es parte integrante de la presente Recomendación)

C.1 Interfaz analógica común

La interfaz analógica común (véase la Recomendación G.192) es esencial para asegurar condiciones de entrada y de salida idénticas con el objeto de probar los algoritmos destinados a la codificación y procesamiento de la voz. Se espera también proporcionar una qdu exacta para estimar la calidad de red del sistema objetivo.

C.2 Interfaz digital

La interfaz digital conecta el equipo de procesamiento de la voz o de codificación de la voz candidato a la interfaz analógica común. La interfaz digital es conforme a la configuración indicada en la nueva Recomendación sobre herramientas de soporte lógico. Esta interfaz tiene una frecuencia de muestreo de 16 kHz y precisión de cuantificación de hasta 16 bits.

C.3 Interfaz analógica

La interfaz analógica está constituida por convertidores A/D y D/A, un filtro que produce distorsión de atenuación del equipo terminal MIC, codificación y decodificación log-MIC conforme a la Recomendación G.711, y el aparato MNRU según la Recomendación P.810 a fin de obtener la distorsión de una unidad qdu exactamente.

C.3.1 Convertidores A/D y D/A

El convertidor A/D convierte una señal analógica en datos con codificación MIC uniformes con una precisión mínima de 13 bits y una precisión máxima de 16 bits a una frecuencia de muestreo de 16 kHz. La entrada al convertidor A/D debe incluir filtrado antisolapamiento, y la salida del convertidor D/A debe pasar a través de un filtro de reconstrucción. Este filtrado tiene una respuesta de frecuencia plana en la banda de paso y una pérdida suficiente en la banda de corte. La distorsión prevista para esta parte no será superior al valor teórico de la distorsión de cuantificación de la codificación MIC uniforme. El convertidor D/A se debe equilibrar con el convertidor A/D.

C.3.2 Distorsión de atenuación en función de la frecuencia

La distorsión de atenuación en función de la frecuencia de los lados emisión o recepción de la interfaz debe ser conforme al gálibo de la Figura 3/P.66. Sin embargo, para limitar cualquier distorsión de atenuación en la banda de paso debida a una conexión en cascada, la amplitud de la ondulación de la banda de paso debe ser inferior a 0,1 dB.

C.3.3 Aparato MNRU

La distorsión total, correspondiente a una unidad qdu la proporciona un códec MIC cuya relación señal/distorsión es de 35 dB cuando la señal de entrada está comprendida entre 0 y -30 dBm0, y se mide por el método indicado en la Recomendación O.132. La relación S/D (señal/distorsión) para la interfaz analógica común se fija en 2 dB por encima del gálibo de la Recomendación G.712 y se considera equivalente a la de un códec comercial típico (véase la Recomendación G.712).

Puesto que un códec MIC típico de esas características es un dispositivo hipotético, resulta difícil diseñar un códec real que proporcione la relación S/D exacta. Suponiendo que los convertidores A/D y D/A son ideales y que la codificación log-MIC se efectúa mediante consulta de un cuadro, este códec se comportará considerablemente mejor que el equipo MIC comercial característico. Para proporcionar la emulación práctica de un códec MIC típico, se debe incorporar un aparato MNRU (conforme a la Recomendación P.810) con $Q = 36$ dB entre el convertidor A/D y el codificador log MIC, y entre el decodificador log-MIC y el convertidor D/A, como se indica en la Figura C.1.

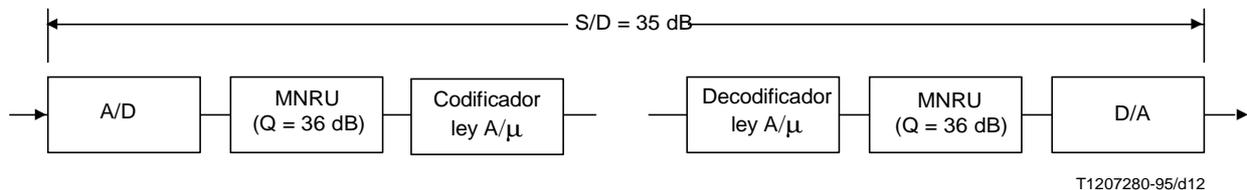


FIGURA C.1/P.830
Configuración para 1-qdu

El sistema se debe calibrar conforme al siguiente procedimiento, en el que se supone que una tolerancia de $\pm 0,1$ dB es aceptable para la relación S/D global de 35 dB:

- 1) se fijan ambos MNRU en $Q = 36$ dB;
- 2) se mide la relación de S/D del sistema completo por el método de la Recomendación O.132, e indicando el resultado por Q_0 ;
- 3) si $34,9 < Q_0 < 35,1$, la calibración está completa, de no ser así;
- 4) se añade ($Q_0 - 35$ dB al ajuste de cada aparato MNRU y se vuelve al paso 2).

Anexo D

Características de emisión y recepción del sistema IRS modificado

(Este anexo es parte integrante de la presente Recomendación)

Las características de frecuencia en emisión y en recepción del sistema IRS tienen su origen en una amplia serie de medidas efectuadas en diversos aparatos telefónicos analógicos contemporáneos a comienzos de la década de 1970 [3] y [4]. A partir de esas medidas, se calcularon las características de respuesta media en emisión y en recepción. Sin embargo, a efectos del equilibrio de sonoridad para el que está diseñado el sistema IRS, fue también necesario incluir un filtro de paso banda de 300 a 3400 Hz, conocido como filtro SRAEN. Los valores de sensibilidad en emisión y en recepción que figuran actualmente en la Recomendación P.48 se componen por tanto de las respuestas medias en emisión y en recepción para un cierto número de aparatos telefónicos, así como en la respuesta del filtro SRAEN.

Si bien estas respuestas pueden ser representativas de conexiones sobre medios de transmisión analógicos de larga distancia, no son representativas de conexiones sobre medios digitales modernos. En particular, el régimen de caída a baja frecuencia da lugar a una degradación de la calidad innecesaria. A efectos de evaluación del códec de baja velocidad binaria, especialmente cuando el códec está ubicado en el microteléfono del aparato telefónico, se puede obtener una mejor característica modificando la respuesta del sistema IRS mediante la eliminación del filtro SRAEN, como se muestra en el Cuadro D.1. A efectos informativos se señala que la respuesta del filtro SRAEN está también incluida en el Cuadro D.1. La tolerancia en los puntos nominales mostrados en el Cuadro D.1 para las características de emisión y de recepción del sistema IRS es de $\pm 2,5$ dB para frecuencias comprendidas entre 200 Hz y 3400 Hz. Por debajo de 200 Hz el régimen de caída del sistema IRS modificado debería ser de al menos 15 dB/octava. Por encima de 3400 Hz debe utilizarse un filtro paso bajo adecuado al objeto de satisfacer los requisitos «anti-aliasing» o de reconstrucción. Las realizaciones de estos filtros están disponibles en la STL del UIT-T.

CUADRO D.1/P.830

Frecuencia (Hz)	Características en emisión del IRS modificado (dBV/Pa)	Características en recepción del IRS modificado (dBPa/V)	Pérdidas de inserción del filtro SRAEN (dB)
100	-31,7	-13,4	14,1
125	-24,7	-7,4	11,4
160	-17,2	-2,4	8,4
200	-13,3	3,2	5,9
250	-10,3	6,7	4,0
300	-8,5	9,2	2,8
315	-8,3	9,7	2,5
400	-7,0	11,3	1,4
500	-6,3	11,9	10,6
600	-6,0	12,1	0,3
630	-5,9	12,1	0,2
800	-4,9	12,3	0,0
1000	-3,7	12,6	0,0
1250	-2,3	12,5	0,0
1600	-0,5	13,1	0,1
2000	0,1	12,9	-0,2
2500	1,3	12,6	-0,5
3000	2,0	13,0	0,5
3150	2,1	12,9	0,3
3500	-0,3	10,9	7,0
4000	-3,5	2,1	33,7
5000	-9,0	-11,7	43,2

Bibliografía

- [1] BOYD (I.), SOUTHCOTT (C.B.): A speech codec for the Skyphone service, *British Telecom Technology Journal*, Vol. 6, No. 2, abril 1988.
- [2] BT Laboratories: Enhanced Equivalent-Q Rating Algorithm, *ETSI/TM/TM5/TCH-HS Document TD 93/126*, diciembre 1993.

- [3] Contribución COM XII-79 del CCITT, *Especificación para un sistema intermedio de referencia*, Periodo de Estudios 1973-1976.
- [4] Contribución COM XII-104 del CCITT, *Resumen y análisis de los resultados de las mediciones del índice de sonoridad subjetivos y objetivos efectuados con once sistemas telefónicos por el laboratorio del CCITT*, Periodo de Estudios 1973-1976.
- [5] CCITT Report of the meeting of Working Party XVIII/2 (Speech processing), *COM XVIII-R 28, Anexo 1*, pp. 13-39, diciembre 1983.
- [6] COLEMAN (A.), GLEISS (N.), SOTSHECK (J.), USAI (P.), SCHEUERMANN (H.): Subjective performance evaluation of the RPE-LTP codec for the Pan-European cellular digital mobile radio system, *IEEE Globecom 1989*, Dallas, Texas, 27-30 de noviembre 1989.
- [7] COLEMAN (A.), GLEISS (N.), USAI (P.): A Subjective Testing Methodology for Evaluating Medium Rate Codecs for Digital Mobile Applications, *Speech Communications*, Vol. 7, pp. 151-166, junio 1988.
- [8] COMBESCURE (P.) *et al.*: Quality evaluation of speech coded at 32 kbit/s by means of degradation category ratings, *Proc. ICASSP 82 (International Conference on Acoustics, Speech and Signal Processing)*, Vol. 2, Paris, mayo 1982.
- [9] CROWE (D.P.): Selection of Voice Codec for the Aeronautical Satellite Service, *European Conference on Speech Communication and Technology*, Vol. 2, S37, pp. 320-323, septiembre 1989.
- [10] DAUMER (W.R.), CAVANAUGH (J.R.): A subjective comparison of selected digital codecs for speech, *Bell Systems Technical Journal*, Vol. 57, No. 9, noviembre 1978.
- [11] DIMOLITSAS (S.), CORCORAN (F.), RAVISHANKAR (C.): Correlation between headphone and telephone handset listener opinion scores for single-stimulus voice coder performance assessments, *IEEE Signal Processing Letters*, Vol. 2, No. 3, marzo 1995.
- [12] EUROPEAN JEG: Subjective testing methodology for the evaluation of low-bit rate codecs for mobile radio, *CCITT, COM XII-68*, mayo 1986.
- [13] GOODMAN (D.J.), NASH (R.D.): Subjective quality of the same speech transmission conditions in seven different countries, *Proc. ICASSP 82 (International Conference on Acoustics, Speech and Signal Processing)*, Vol. 2, Paris, mayo 1982.
- [14] ITU-T Contribution COM 15-20: Transmission quality of interconnected PSTN-digital cellular networks, *COMSAT, Study Period 1993-1996*.
- [15] MODENA (G.), COLEMAN (A.), USAI (P.), COVERDALE (P.): Subjective performance evaluation of the 7 kHz audio coder, *IEEE Global Telecommunications Conference 1986 (Globecom '86)*, Houston, Texas, 1-4 de diciembre 1986.
- [16] RICHARDS (D.L.), BARNES (G.J.): Pay-off between quantizing distortion and injected circuit noise, *Proc. ICASSP 82 (International Conference on Acoustics, Speech and Signal Processing)*, Vol. 2, Paris, mayo 1982.
- [17] Swedish Telecommunication Administration Report: Subjective test on candidate codecs for mobile radio, *CCITT, COM XII-147*, febrero 1987.
- [18] WHEDDON (C.), LINGGARD (R.): *Speech and Language Processing*, Chapman and Hall, 1990.