



UNIÓN INTERNACIONAL DE TELECOMUNICACIONES

UIT-T

SECTOR DE NORMALIZACIÓN
DE LAS TELECOMUNICACIONES
DE LA UIT

P.83

(03/93)

**CALIDAD DE TRANSMISIÓN
TELFÓNICA**

PRUEBAS SUBJETIVAS DE OPINIÓN

**EVALUACIÓN DE LA CALIDAD DE
FUNCIONAMIENTO SUBJETIVA DE
LOS CODECS DIGITALES DE BANDA
TELFÓNICA Y DE BANDA ANCHA**

Recomendación UIT-T P.83

(Anteriormente «Recomendación del CCITT»)

PREFACIO

El Sector de Normalización de las Telecomunicaciones de la UIT (UIT-T) es un órgano permanente de la Unión Internacional de Telecomunicaciones. El UIT-T tiene a su cargo el estudio de las cuestiones técnicas, de explotación y de tarificación y la formulación de Recomendaciones al respecto con objeto de normalizar las telecomunicaciones sobre una base mundial.

La Conferencia Mundial de Normalización de las Telecomunicaciones (CMNT), que se reúne cada cuatro años, establece los temas que habrán de abordar las Comisiones de Estudio del UIT-T, que preparan luego Recomendaciones sobre esos temas.

La Recomendación UIT-T P.83, revisada por la Comisión de Estudio XII (1988-1993) del UIT-T, fue aprobada por la CMNT (Helsinki, 1-12 de marzo de 1993).

NOTAS

1 Como consecuencia del proceso de reforma de la Unión Internacional de Telecomunicaciones (UIT), el CCITT dejó de existir el 28 de febrero de 1993. En su lugar se creó el 1 de marzo de 1993 el Sector de Normalización de las Telecomunicaciones de la UIT (UIT-T). Igualmente en este proceso de reforma, la IFRB y el CCIR han sido sustituidos por el Sector de Radiocomunicaciones.

Para no retrasar la publicación de la presente Recomendación, no se han modificado en el texto las referencias que contienen los acrónimos «CCITT», «CCIR» o «IFRB» o el nombre de sus órganos correspondientes, como la Asamblea Plenaria, la Secretaría, etc. Las ediciones futuras en la presente Recomendación contendrán la terminología adecuada en relación con la nueva estructura de la UIT.

2 Por razones de concisión, el término «Administración» se utiliza en la presente Recomendación para designar a una administración de telecomunicaciones y a una empresa de explotación reconocida.

© UIT 1994

Reservados todos los derechos. No podrá reproducirse o utilizarse la presente Recomendación ni parte de la misma de cualquier forma ni por cualquier procedimiento, electrónico o mecánico, comprendidas la fotocopia y la grabación en micropelícula, sin autorización escrita de la UIT.

ÍNDICE

	<i>Página</i>
1	Introducción 1
2	Grabaciones de señales en la fuente 2
2.1	Sistema emisor 2
2.2	Muestras vocales 2
3	Selección de los parámetros del experimento 2
3.1	Condiciones del códec 2
3.2	Condiciones de referencia 7
4	Plan del experimento 9
5	Procedimiento de la prueba de escucha 9
5.1	Sistema de recepción 10
5.2	Escalas de opinión 10
5.3	Ruido eléctrico 10
6	Análisis de los resultados 10
	Anexo A – Comparación de las distintas definiciones de la relación señal/ruido 12
	Anexo B – Conjunto de parámetros para determinar la calidad de funcionamiento de códec 13
	B.1 Sistemas de banda estrecha (300-3400 Hz) 13
	B.2 Sistemas de banda ancha (100-7000 Hz) 13
	Referencias 14
	Bibliografía 14

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

(Helsinki, 1993)

1 Introducción

Actualmente se va introduciendo rápidamente la telefonía digital en la red telefónica pública conmutada (RTPC). Desde los años 60, la transmisión digital en la RTPC es sinónimo de MIC de ley A a 64 kbit/s (véase la Recomendación G.711). Actualmente, el 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), suscitan un creciente interés por los métodos de codificación más eficaces que el de MIC. Un ejemplo de este desarrollo es la introducción de MICDA de 32 kbit/s (Recomendación G.726). Recientemente el CCITT ha normalizado un códec de banda ancha de 7 kHz a 64 kbit/s (Recomendación G.722). Actualmente el interés está centrado en la codificación de baja velocidad binaria (LRE, *low rate encoding*), es decir, a velocidades binarias de 16 kbit/s o inferiores.

Cuando el trayecto de transmisión es digital y/o no lineal, las mediciones de objetivos sencillos, como se describe en la Recomendación G.712, son insuficientes para asegurar la adecuada calidad de transmisión. 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 el comportamiento 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 los tramos no lineales y los otros tramos del sistema de transmisión.

Esto supone tener la posibilidad de asignar de forma unívoca una contribución numérica a cada proceso digital así como la posibilidad de utilizar esta contribución asignada junto con otras degradaciones para estimar la calidad de funcionamiento de la conexión telefónica.

La presente Recomendación, basada en el Anexo B/P.80 así como en la experiencia obtenida en diversos experimentos de carácter internacional [2, 3, 4, 5, 6], define un método de prueba específico para evaluar los procesos digitales, de forma que los efectos de distorsión de cuantificación de estos procesos sobre la calidad de funcionamiento de la transmisión puedan tenerse en cuenta en la red telefónica internacional en constante evolución.

Para caracterizar la calidad de funcionamiento de los procesos digitales pueden utilizarse varios métodos. Actualmente dichos métodos consisten en pruebas de sólo escucha 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 CCITT recomienda la utilización del método de 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 lo que también se utilizan en esta Recomendación. Sólo se indicarán las desviaciones con respecto a la utilización del método ACR que utiliza la escala de calidad de escucha.

NOTA – El método de índices por categorías de degradación (DCR, *degradation category rating*) [1] se escribe con detalle en el Anexo D/P.80. Se recomienda emplear este método 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.

La presente Recomendación contiene indicaciones sobre la forma de evaluar la calidad de funcionamiento de los codecs digitales y debe considerarse juntamente con la Recomendación P.80.

Además, es preciso complementar las pruebas de sólo escucha con pruebas de conversación, pero hasta ahora no se ha llegado a un acuerdo sobre un método para incluirlo en esta Recomendación.

2 Grabaciones de señales en la fuente

Esta cláusula es la misma que B.1/P.80, con las siguientes excepciones.

2.1 Sistema emisor

Para sistemas de banda estrecha (300-3400 Hz), si es preciso comparar los resultados obtenidos en los laboratorios y en las administraciones, se recomienda utilizar el sistema intermedio de referencia (IRS, *intermediate reference system*) especificado en la Recomendación P.48 y calibrado de acuerdo con la Recomendación P.64. Sin embargo, se ha determinado que la calidad de funcionamiento de los codecs vocales de baja velocidad binaria puede depender muy significativamente de la característica de frecuencia aplicada a la señal vocal de entrada.

NOTA – El IRS es representativo de los aparatos telefónicos analógicos (del decenio de 1970) y actualmente la Comisión de Estudio 12 está estudiando un «teléfono de referencia» representativo de los aparatos telefónicos modernos (véase Figura 1/P.31).

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

2.2 Muestras vocales

Se recomienda utilizar un mínimo de dos frases por muestra, lo que permite, por ejemplo, evaluar tasas de errores en los bits (BER, *bit-error-rate*) de 0, 1:10000, 1:1000 y 1:100 (véase 3.1.4). Los valores de BER entre 0 y 1:10000 exigirán un número mayor de frases por muestra.

3 Selección de los parámetros del experimento

3.1 Condiciones del códec

3.1.1 Niveles de entrada de la señal vocal

Cuando se evalúa el códec es práctica habitual ajustar los niveles de entrada con relación al punto de saturación del códec. Se recomienda utilizar niveles de entrada de 14, 26 y 38 dB por debajo del punto de saturación del códec (equivalentes a valores de -8, -20 y -32 dBm0). 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 produce una modificación en el nivel de entrada de la señal vocal.

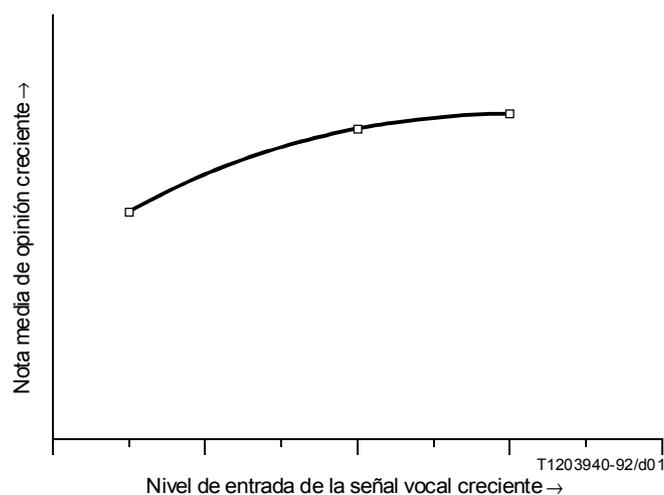


FIGURA 1/P.83

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

3.1.2 Niveles de escucha

Se recomienda utilizar al menos tres niveles de escucha que no deben diferir en más de ± 10 dB del nivel de escucha preferido (véase B.2.1/P.80 para información sobre niveles de escucha preferidos). En la Figura 2 se representa el efecto que produce la modificación del nivel de escucha.

3.1.3 Participantes

3.1.3.1 Diversidad de participantes

Se recomienda una participación mínima de dos hombres y dos mujeres. En la Figura 2 se representa el efecto de participantes distintos. Sin embargo, si se va a probar la dependencia con los participantes se debe emplear el siguiente número de sujetos:

- 8 hombres
- 8 mujeres
- 8 niños

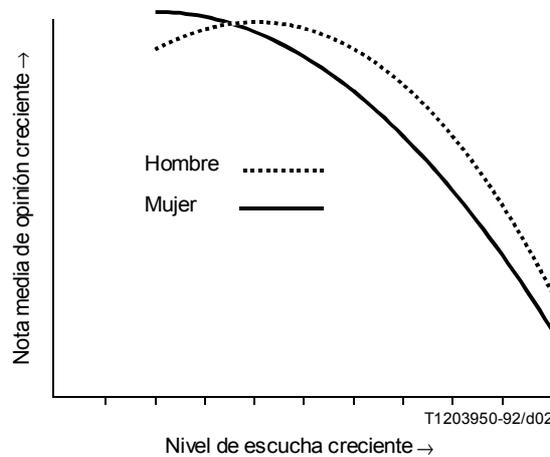


FIGURA 2/P.83

Nota media de opinión en función del nivel de escucha

3.1.3.2 Múltiples participantes

Es importante conocer la robustez del códec frente a diversas señales vocales de entrada y determinar si pueden aparecer efectos negativos tales como cortes de la transmisión, señales parásitas, etc.

Si sólo va a considerarse la utilización del microteléfono, normalmente es suficiente la combinación de dos participantes distintos con niveles de entrada de la señal vocal separados 20 dB.

Sin embargo, si va a utilizarse el modo conferencia; es decir, teléfono de manos libres, donde más de un participante pueda hablar al mismo tiempo hay que asegurar que el algoritmo de codificación puede tratar varios participantes cuya diferencia de nivel de señal vocal puede ser nula.

Para evaluar los efectos de múltiples participantes se recomienda utilizar el método de índices por categorías de degradación (DCR) que hace uso de la escala de 5 notas de 5.2.3 (véase también el Anexo D/P.80) o el método de detectabilidad de respuesta cuantificada que utiliza la escala de 3 notas de 5.2.4 (véase también el Anexo C/P.80).

3.1.4 Errores

Éstos dependerán de la aplicación, es decir, de que los codecs se utilicen en la RTPC convencional (sistemas de líneas) o, por ejemplo, en un entorno radioeléctrico como el de las radiocomunicaciones móviles.

En sistemas de líneas 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 2.2). En algunas circunstancias es necesario efectuar pruebas con valores de BER de hasta 1:100.

Para otras aplicaciones, como las de radiocomunicaciones móviles, los errores pueden ser del tipo ráfaga y, por consiguiente, convendría emplear errores de este tipo.

Normalmente se utiliza el método ACR que hace uso de la escala de calidad de escucha (5.2.1). Sin embargo, en caso de que se suponga una baja calidad, puede ser más adecuado utilizar el método ACR que utilice la escala de esfuerzo de escucha (5.2.2).

3.1.5 Velocidades binarias

El códec debe comprobarse para todas las velocidades binarias con las que es capaz de trabajar; 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 puede aplicarse la carga del sistema (véase la Recomendación P.84). La Figura 3 representa el efecto de la modificación de la velocidad binaria.

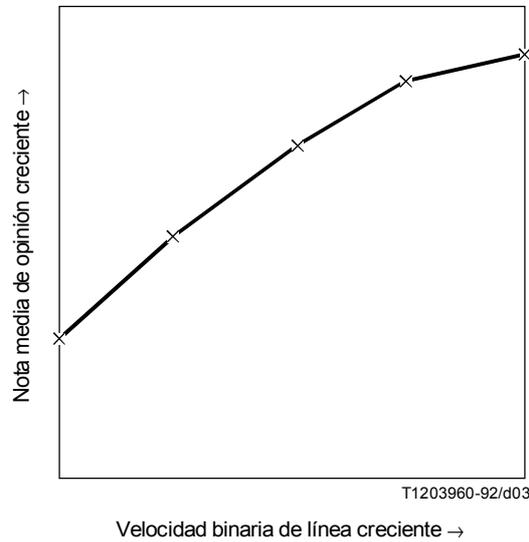


FIGURA 3/P.83
**Nota media de opinión en función de la
 velocidad binaria de línea**

3.1.6 Transcodificaciones

Es habitual evaluar el comportamiento de un códec asociado a un proceso de compansión de ley A o ley μ . Por consiguiente, un códec con entrada y salida analógicas tendrá la configuración mostrada en la Figura 4.

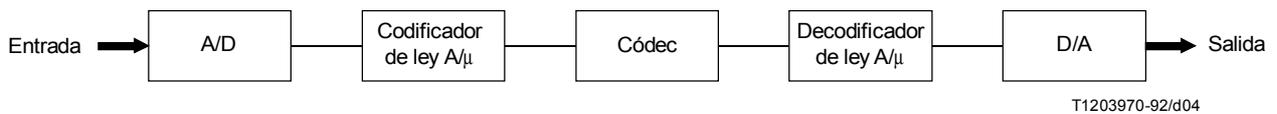


FIGURA 4/P.83
Esquema de transcodificación

3.1.6.1 Conexión asíncrona en cascada

Este método refleja la utilización del códec en la red, basada en la arquitectura de 64 kbit/s. Por consiguiente, cada transcodificación incluye una distorsión de cuantificación adicional, debida a la transición de compansión lineal a compansión de ley A o μ , y a la distorsión de atenuación adicional debida a las tolerancias 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 CCITT recomienda que se comprueben uno, dos y, como mínimo, tres códecs conectados en cascada.

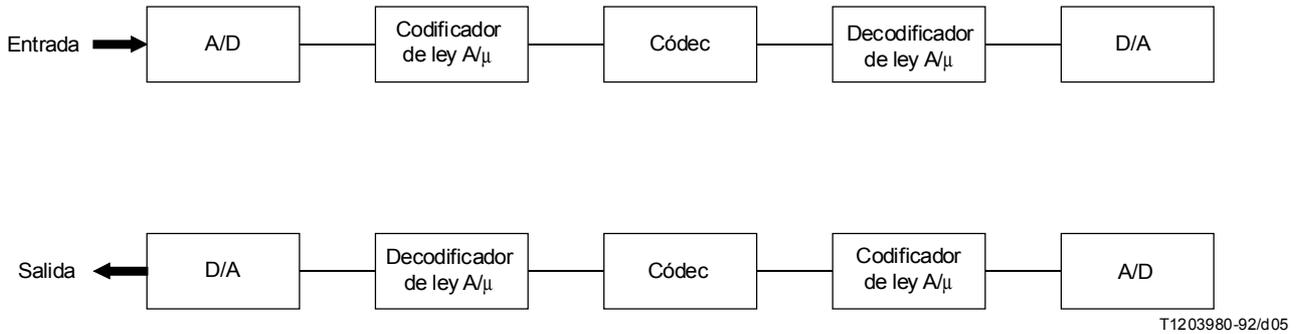


FIGURA 5/P.83
Conexión asíncrona en cascada

3.1.6.2 Conexión síncrona en cascada

La configuración del circuito se representa en la Figura 6 y presenta la ventaja de que existe únicamente un proceso de compansión de ley A o μ y, por consiguiente, aparece una distorsión de cuantificación reducida en comparación con el caso anterior (3.1.6.1).

El CCITT recomienda que se comprueben uno, dos y, como mínimo, tres códecs conectados en cascada.

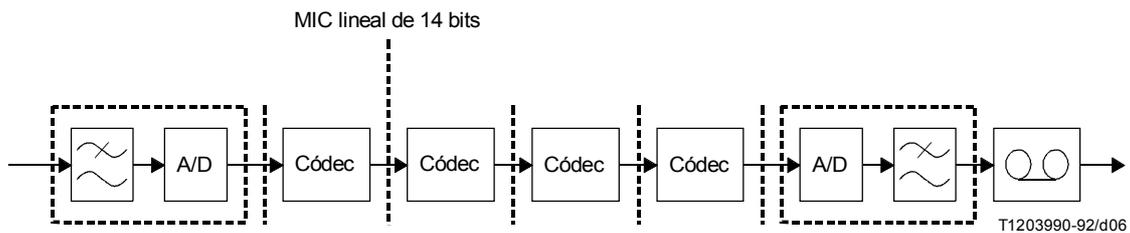
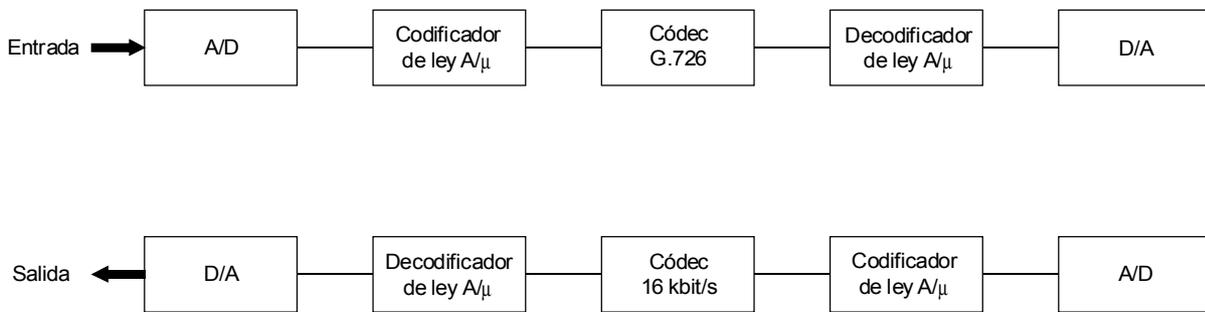


FIGURA 6/P.83
Conexión síncrona en cascada

3.1.6.3 Conexión mixta en cascada

Es importante conocer los efectos de los sistemas en cascada que utilizan codificación a distintas velocidades binarias, y para ello se utiliza la configuración representada en la Figura 7. Es fundamental verificar el comportamiento de las combinaciones más probables.



T1204000-92/d07

FIGURA 7/P.83
Conexión mixta en cascada

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 anuncios grabados	circuito de enlace	DCME de velocidad binaria variable	circuitos de enlace	téfonos sin cordón

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

3.1.7 Desadaptación

En los sistemas de banda ancha (Recomendación G.722) se recomienda investigar los efectos debidos a problemas de desadaptación, es decir, a modos de funcionamiento distintos del transmisor y el receptor, en las siguientes condiciones:

<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

3.1.8 Ruido ambiental (emisión)

Al igual que las consideraciones sobre múltiples participantes efectuadas en 3.1.3.2, el efecto de interacción del ruido ambiental es un factor importante, por las siguientes razones:

- los codecs de baja velocidad binaria pueden hacer uso de procesos de síntesis de la palabra así como de una codificación convencional de la señal; 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.

Deben efectuarse las pruebas suficientes con el ruido apropiado (véase A.1.1.2.2/P.80) para comprobar este efecto y se recomienda utilizar las siguientes relaciones señal/ruido (véase en 2.2.3 las definiciones):

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

Como la calidad percibida será algo inferior a la normal, el CCITT recomienda la utilización del método ACR basado en la escala de esfuerzo de escucha (véase 5.2.2) y en la capacidad de entender el significado de las frases.

NOTAS

1 El ruido puede combinarse eléctricamente con las señales de la cinta magnética 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 la señales vocales y de un nivel elevado de ruido ambiental grabados por separado 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 codecs pueden procesar distintas partes del espectro de audio de formas diferentes, esta modificación de las características vocales puede afectar a la calidad de funcionamiento.

3.1.9 Señales de información de red

En toda red nacional hay diversos tonos o señales de información que se transmiten para beneficio del usuario y algunas que se tratan de instrucciones a 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 cuando estas señales atraviesen algún dispositivo de procesamiento de la palabra, éste no las degrade de tal forma que sean irreconocibles para el usuario o para el equipo que debe actuar de una forma determinada al recibirlas. Es posible que esta última situación, por ejemplo, los tonos de señalización multifrecuencia, sea menos tolerante a la degradación que la primera, pero presenta la ventaja de que es suficiente realizar pruebas objetivas sencillas para determinar el límite admisible de dicha degradación.

Se recomienda probar 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 pertinencia (reconocimiento) de las señales de información se recomienda utilizar el método de índices por categorías de degradación (DCR) con escala de 5 notas del 5.2.3 (véase también el Anexo D/P.80).

3.1.10 Música

Algunos equipos utilizados en la RTPC, especialmente las centralitas privadas, tienen la posibilidad de emitir música al abonado mientras éste espera.

El CCITT recomienda utilizar únicamente pruebas sencillas, para asegurarse de que la música presente una calidad razonable.

3.2 Condiciones de referencia

3.2.1 Condición directa

Esta representa la mejor condición obtenible en el experimento y equivale a un valor infinito de Q_N o Q_W (véase 3.2.2).

3.2.2 Aparato de referencia para ruido modulado (MNRU)

El ruido aleatorio con una amplitud proporcional a la amplitud de la señal instantánea en términos de Q_N o Q_W , de acuerdo con el MNRU especificado en la Recomendación P.81, debe utilizarse como sistema de referencia para expresar la calidad de funcionamiento subjetiva de los procesos digitales, por las siguientes razones:

- a) para codecs de forma de onda y posiblemente para algunos codecs distintos de los de forma de onda la señal procesada a través del MNRU es de percepción muy similar a la señal procesada, lo que da lugar, en principio, a una evaluación más fácil por los participantes en la prueba;
- b) la experiencia ha demostrado que el MNRU es una norma de transferencia útil y permite realizar comparaciones razonables entre distintos laboratorios; y
- c) se dispone ya de una información y experiencia considerables en lo que respecta al MNRU.

Esta gama de sistemas de banda estrecha debe abarcar valores de Q_N entre 5 dB y 35 dB (preferiblemente de 5 a 7 valores distintos) y en sistemas de banda ancha debe ser para valores de Q_W entre 10 dB y 45 dB.

En la Figura 8 se representa el efecto de la modificación de Q_N o Q_W .

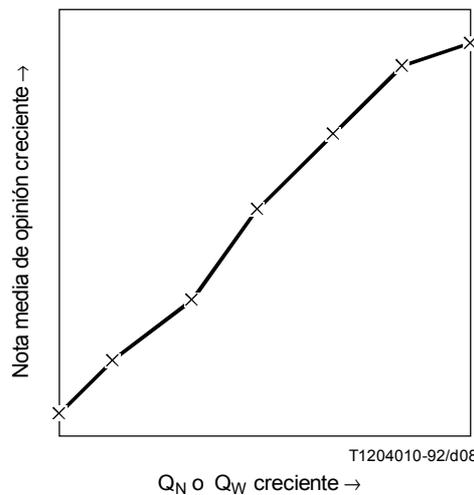


FIGURA 8/P.83
Nota media de opinión en función de Q_N o Q_V

3.2.3 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 CCITT recomienda las siguientes definiciones de la relación señal/ruido, medida en las conexiones, para el ruido en régimen estacionario:

- *Mediciones telefónicas (de banda estrecha) – Ponderación sofométrica* (véase la Recomendación O.41)
SNR(p) = Nivel vocal activo (véase la Recomendación P.56)/medición del ruido con ponderación sofométrica
- *Mediciones de banda ancha – Ponderación A* (véase la Recomendación P.54)
SNR(A) = Nivel vocal activo (véase la Recomendación P.56)/medición 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)/medición del ruido sin ponderación
- *Banda ancha 100-7000 Hz – Sin ponderación*
SNR(W) = Nivel vocal activo (véase la Recomendación P.56)/medición del ruido sin ponderación
- *Ponderación de mensaje C* (véase la Recomendación O.41)
SNR(C) = Nivel vocal activo (véase la Recomendación P.56)/medición del ruido con ponderación de mensaje C
- *Anchura de banda 100-5000 Hz*
SNR(0,1-5) = Nivel vocal activo (véase la Recomendación P.56)/medición del ruido sin ponderación

NOTA – Si se utilizan otras anchuras de banda, las cifras entre paréntesis (0,1-5) deben sustituirse por la anchura de banda de medición.

En el Anexo A figura una comparación de las distintas relaciones señal/ruido según su definición 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 entre 15 y 45 dB.

En la Figura 9 se representa el efecto de la variación de la relación SNR.

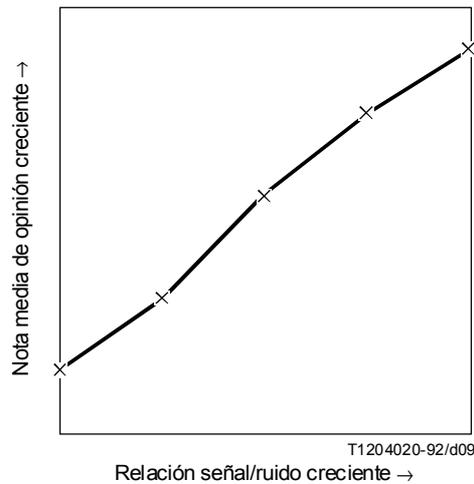


FIGURA 9/P.83

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

3.2.4 Codecs de referencia

Los codecs de referencia tienen dos objetos:

- 1) pueden utilizarse para determinar las reglas de planificación (véase la Recomendación G.113);
- 2) pueden utilizarse como norma para juzgar la calidad de funcionamiento global en función de los parámetros (véase 3.1).

Si se emplean para determinar las reglas de planificación, el CCITT recomienda utilizar una conexión en cascada asíncrona (véase 3.1.6.1) de 1, 2, 4, 8 y 16 codecs de ley A o μ , conforme a las Recomendaciones G.711/712.

Si se emplean para determinar el comportamiento relativo, por ejemplo con respecto al indicado en las Recomendaciones G.726 o G.722, deben aplicarse las consideraciones efectuadas en 3.1.6.1 y 3.1.6.2.

NOTA – Se está estudiando la especificación de un códec G.711 de referencia.

4 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 2 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 razonables. En el Anexo B figura un conjunto de experimentos que el CCITT ha encontrado adecuados para los estudios de los codecs de banda estrecha y banda ancha; esos estudios sirven de base para la formulación de las Recomendaciones de la serie G.700.

Deben extremarse las precauciones cuando se comparen en una misma prueba sistemas con degradaciones muy distintas; por ejemplo, codecs digitales, sistemas múltiplex por división de frecuencia, vocoders, etc.).

5 Procedimiento de la prueba de escucha

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

5.1 Sistema de recepción

Para sistemas de banda estrecha (300-3400 Hz) se recomienda utilizar el extremo receptor del sistema intermedio de referencia 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 de recepción debe ser conforme a la Publicación 581.7 de la CEI.

5.2 Escalas de opinión

El método que debe utilizarse es el de índice por categorías absolutas (ACR) con escala de 5 notas, conforme se describe en B.4.5/P.80.

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

5.2.1 Escala de calidad de escucha

Calidad de la señal vocal:

- 5 Excelente
- 4 Buena
- 3 Regular
- 2 Mediocre
- 1 Mala

5.2.2 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.

5.2.3 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.

5.2.4 Escala de opiniones sobre la detectabilidad

- 3 Objetable.
- 2 Detectable.
- 1 No detectable.

5.3 Ruido eléctrico

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

6 Análisis de los resultados

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

Un objetivo del análisis es determinar una función $Q_2 = F(L)$, donde Q_2 es el valor de Q para el códec y L es la velocidad binaria de la línea. Un método sencillo para calcular esta función consiste en utilizar los valores de 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 su correspondiente valor de la nota media de opinión. Este valor de la 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 – Actualmente se está estudiando la expresión matemática de la curva que mejor se adapta a los resultados de las notas medias de opinión en función de Q .

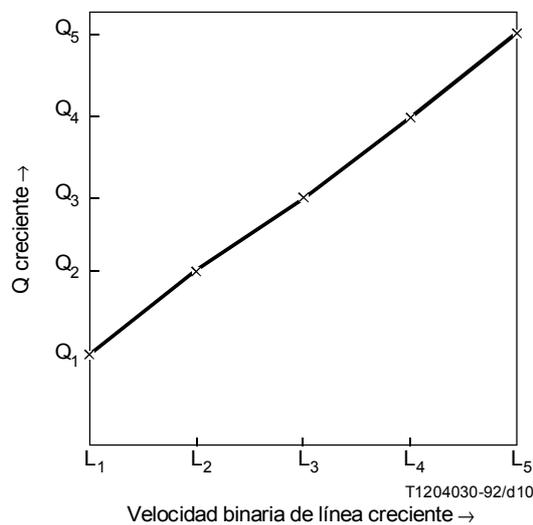


FIGURA 10/P.83
Q en función de la velocidad binaria de línea

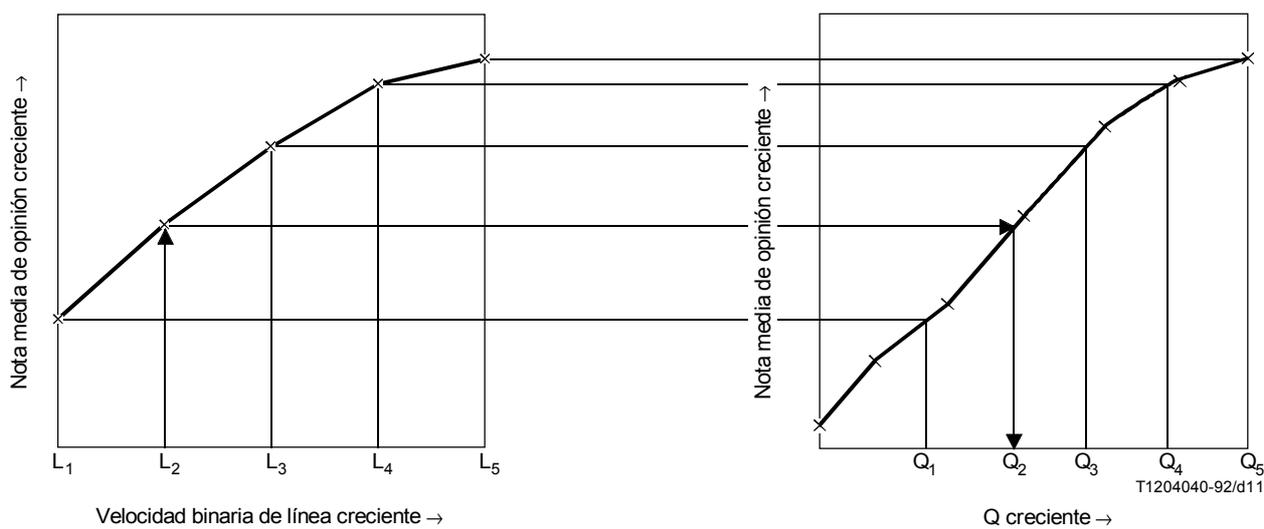


FIGURA 11/P.83
Método gráfico para obtener la Figura 10 a partir de las Figuras 3 y 8

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 la medición de los efectos de este tipo de ruido.
- Ruido (de sala) de Hoth 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.83

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

- Todos los valores indican la diferencia en dB con respecto al ruido sin ponderación.
- El signo negativo significa más silencioso.
- 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.
- 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.
- 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.
- 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)

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

- a) *Experimento 1 – Efectos de los errores, nivel de entrada y nivel de escucha*
- | | |
|-------------------------------------|---|
| Nivel de entrada de la señal vocal: | 3 (14, 26 y 38 dB por debajo del nivel de saturación) |
| Nivel de escucha: | 3 (P + 10, P, P – 10) |
| Errores: | 3 (0, 1:10000, 1:1000) |
| Transcodificación: | 1 transcodificación |
| Ruido ambiental (emisión): | 1 (< 30 dBA) |
- b) *Experimento 2 – Efectos de la transcodificación, nivel de entrada y nivel de escucha*
- | | |
|-------------------------------------|---|
| Nivel de entrada de la señal vocal: | 3 (14, 26 y 38 dB por debajo del nivel de saturación) |
| Nivel de escucha: | 3 (P + 10, P, P – 10) |
| Transcodificación: | «x» |
| Error: | 1 (1:1000) |
| Ruido ambiental (emisión): | 1 (< 30 dBA) |
- c) *Experimento 3 – Efecto del ruido ambiental, ruido de sala, nivel de entrada y nivel de escucha*
- | | |
|-------------------------------------|---|
| Nivel de entrada de la señal vocal: | 3 (14, 26 y 38 dB por debajo del nivel de saturación) |
| Nivel de escucha: | 3 (P + 10, P, P – 10) |
| Transcodificación: | 1 transcodificación |
| Error: | 1 (1:1000) |
| Ruido ambiental (emisión): | 2 (< 30 dBA e «y») |
| 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 del 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, BER, nivel de entrada y nivel de escucha*
- | | |
|-------------------------------------|---|
| Nivel de entrada de la señal vocal: | 2 (20 y 38 dB por debajo del nivel de saturación) |
| Nivel de escucha: | 3 (P + 10, P, P – 10) |
| Transcodificación: | 1 transcodificación |
| Ruido ambiental (emisión): | 1 (< 30 dBA) |
| Velocidades binarias: | 3 (48, 56 y 64 kbit/s) |
| BER: | 3 (0, 1:10000 y 1:1000) |
| Ruido de sala: | 1 (< 30 dBA) |

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

Nivel de entrada de la señal vocal:	2 (20 y 38 dB por debajo del nivel de saturación)
Nivel de escucha:	3 (P + 10, P, P – 10)
Transcodificaciones:	«x», incluidas síncronas y asíncronas
Ruido ambiental (emisión):	1 (< 30 dBA)
Velocidades binarias:	3 (48, 56 y 64 kbit/s)
BER:	3 (0, 1:10000 y 1:1000)
Ruido de sala:	1 (< 30 dBA)

c) *Experimento 3 – Efectos de la desadaptación, nivel de entrada y nivel de escucha*

Nivel de entrada de la señal vocal:	2 (20 y 38 dB por debajo del nivel de saturación)
Nivel de escucha:	3 (P + 10, P, P – 10)
Transcodificación	1 transcodificación
Ruido ambiental (emisión):	1 (< 30 dBA)
Velocidades binarias:	3 (56→ 64, 48→ 56, 48→ 64 kbit/s)
BER:	2 (0 y 1:1000)
Ruido de sala:	1 (< 0 dBA)

donde «P» es el nivel de escucha preferido

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

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

Referencias

- [1] COMBESCURE (P.) y otros: Quality evaluation of speech coded at 32 kbit/s by means of degradation category ratings, *Proc. ICASP 82 (International Conference on Acoustics, Speech and Signal Processing)*, Vol. 2, París, mayo de 1982.
- [2] CCITT Informe de la reunión del Grupo de Trabajo XVIII/2 (Procesamiento de la palabra) COM XVIII – R 28, Anexo 1, pp. 13-39, diciembre de 1983.
- [3] GOODMAN (D.J.), NASH (R.D.): Subjective quality of the same speech transmission conditions in seven different countries, *Proc. ICASP 82 (International Conference on Acoustics, Speech and Signal Processing)*, Vol. 2, París, mayo de 1982.
- [4] MODENA (G.), GOLEMAN (A.), USAI (P.) y COVERDALE (P.): Subjective performance evaluation of the 7 kHz audio coder, *IEEE Global Telecommunications Conference 1986 (Globecom 86)*, Houston, Texas, 1-4 de diciembre de 1986.
- [5] GRUPO MIXTO EUROPEO DE EXPERTOS: Metodología de las pruebas subjetivas destinadas a evaluar codecs de bajas velocidades binarias para las radiocomunicaciones móviles CCITT, COM XII-68, mayo de 1986.
- [6] SWEDISH TELECOMMUNICATION ADMINISTRATION: Report on subjective test on candidate codecs for mobile radio, CCITT, COM XII-147, febrero de 1987.

Bibliografía

Recomendación G.726 *Modulación por impulsos codificados diferencial adaptativa (MICDA) a 40, 32, 24 y 16 kbit/s*, Tomo III, Rec. G.726.

DAUMER (W.R.) y CAVANAUGH (J.R.): A subjective comparison of selected digital codecs for speech, *Bell Systems Technical Journal*, Vol. 57, Nº 9, noviembre de 1978.

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

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

COLEMAN (A.), GLEISS, (N.) USAI (P.): A Subjective Testing Methodology for Evaluating Medium Rate Codecs for Digital Mobile Applications, *Speech Communications*, Vol. 7, pp 151-156, junio de 1988.

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 de 1989.

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

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

