



UNIÓN INTERNACIONAL DE TELECOMUNICACIONES

Adiciones al Manual sobre Telefonometría



UIT-T

SECTOR DE
NORMALIZACIÓN DE LAS
TELECOMUNICACIONES
DE LA UIT

1999



UNIÓN INTERNACIONAL DE TELECOMUNICACIONES

Adiciones al Manual sobre Telefonometría

UIT-T
SECTOR DE
NORMALIZACIÓN DE LAS
TELECOMUNICACIONES
DE LA UIT

1999

© UIT 1999

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. écrit de l'UIT.

ADICIONES AL MANUAL SOBRE TELEFONOMETRÍA

ÍNDICE

	<i>Página</i>
Adición al § 2.3.6 del Manual sobre telefonometría.....	1
Adición al § 3.2.5 del Manual sobre telefonometría.....	7
Adición al § 4.8 del Manual sobre telefonometría.....	9

ALGUNAS REPERCUSIONES DEL EFECTO LOCAL

1 Introducción

Durante varios años se estudió el efecto local y se ha llegado a algunas conclusiones importantes desde el punto de vista del abonado como hablante y como oyente. Estas conclusiones se relacionan con la repercusión que tiene sobre el abonado el efecto local al escuchar su propia voz y la forma en que, en consecuencia, cambia su nivel de conversación así como cierta repercusión del efecto local cuando el abonado escucha en condiciones de ruido ambiente de nivel moderado a alto. Estas repercusiones se resumen en las figuras 1 a 3.

La relación entre el efecto local para el hablante y el efecto local para el oyente de un teléfono dado depende básicamente de dos factores:

- a) la geometría de su microteléfono; y
- b) si existen o no algunas características de ganancia o pérdida no lineales en el trayecto de efecto local.

En la cláusula 4 figuran algunas directrices para los diseñadores de aparatos telefónicos.

Se proporciona también información sobre la ocurrencia cada vez más frecuente del eco para el hablante con retardo corto, que puede percibirse como un efecto local desagradable para el hablante.

2 Efecto local para el hablante

La figura 1 muestra que hay una gama preferida para el efecto local cuando el abonado habla en condiciones de silencio y que la diferencia entre el efecto local que pudiera ser motivo de objeción, o demasiado apagado, es del orden de 20 dB. (Estos resultados se obtuvieron mediante pruebas en las que sólo se hablaba y han de confirmarse mediante pruebas de conversación.) La gama preferida se encuentra entre 7 y 12 dB para el índice de enmascaramiento para el efecto local (STMR, *sidetone masking rating*) (Recomendación P.76 [1]).

La gama aceptable es mayor y se encuentra entre 1 y 17 dB para el STMR (aunque debe señalarse que el STMR que aumenta hasta un valor de más de 17 dB es probable que afecte solamente al nivel de conversación y esto, sólo marginalmente). Esta gama corresponde a la diferencia entre las dos curvas en el nivel de apreciación del 50%. No se propone que el valor de 17 dB deba considerarse en modo alguno como un valor máximo. Sin embargo, para un valor del STMR de más de 20 dB, la conexión suena como «muerta».

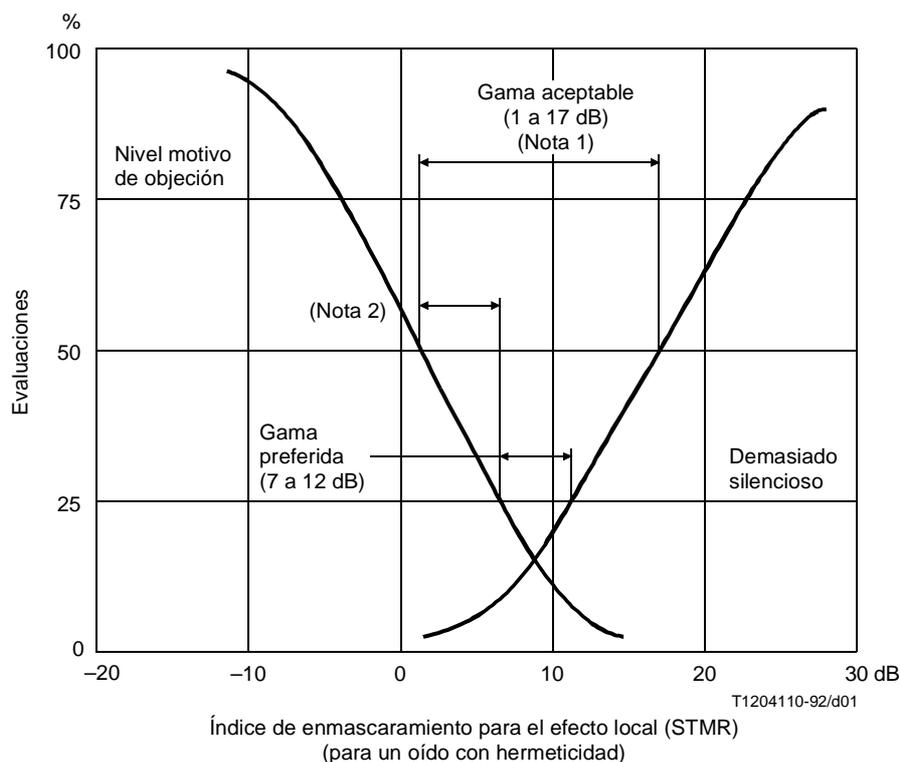
La figura 2 muestra la forma en que cambia el nivel de conversación con el efecto local. Estos resultados se obtuvieron por medio de pruebas de conversación, para una conexión con una atenuación de valor próximo al del equivalente preferido. En las conexiones telefónicas en que el índice de sonoridad global (OLR, *overall loudness rating*) está en la gama preferida, los valores de STMR pueden encontrarse en la gama de STMR preferida indicada anteriormente. Sin embargo, en conexiones con mucha atenuación el valor de STMR debe ser próximo o incluso superior a 12 dB para incitar al abonado a hablar más fuerte. En conexiones con poca atenuación puede permitirse que el valor de STMR sea a veces inferior a 7 dB, pero sólo en raras ocasiones debe ser de sólo 1 dB, por ejemplo, en aparatos telefónicos con control del volumen de recepción. En la Recomendación G.121 [2] se interpretan estos resultados para la planificación de la transmisión.

La tensión de la señal vocal será también función del nivel de ruido ambiente para las mismas condiciones de la conexión. En teléfonos modernos con transmisores lineales puede ser conveniente que el valor de STMR se encuentre en la gama de 10 a 15 dB si los mismos se van a utilizar en ambientes ruidosos.

3 Efecto local para el oyente

Un nivel alto de ruido ambiente en el entorno del abonado perturba la palabra recibida de dos maneras:

- i) por el ruido captado por el micrófono del aparato telefónico y transmitido al receptor del microteléfono a través del trayecto de efecto local eléctrico;
- ii) por el ruido que penetra por la fuga del pabellón del auricular del microteléfono.



NOTA 1 – Las condiciones de conversación determinarán qué parte de esta gama es aceptable para una conexión dada.

NOTA 2 – Esta parte de la gama aceptable (1 a 7 dB) deberá tratarse con precaución, por ejemplo, en conexiones con poca atenuación (véase la Recomendación G.121) o cuando hay control de volumen en recepción.

Figura 1 – Curvas que muestran los niveles del efecto local que son motivo de objeción y demasiado silenciosos, junto con la gama preferida para el abonado como hablante

Los estudios realizados han mostrado que a bajas frecuencias el trayecto de fuga del pabellón del auricular predomina sobre el trayecto de efecto local eléctrico, casi de la misma manera en que la señal del efecto local humano predomina sobre el efecto local para el hablante. Por tanto, son aplicables las ponderaciones realizadas en el cálculo del STMR y se ha elaborado el índice de efecto local para el oyente (LSTR, *listener sidetone rating*) (Recomendación P.76) que utiliza la sensibilidad de efecto local del ruido ambiente (véase la cláusula 9/P.64 [3]) en el método de evaluación del STMR (véase la Recomendación P.79 [4]).

En la figura 3 se muestran los resultados de pruebas subjetivas de la nota media de opinión (MOS, *mean opinion score*) en función del LSTR, (utilizando en este caso una escala de notas medias de opinión de 0 a 10). En cada caso, el LSTR se obtuvo utilizando Δ_{Sm} (véanse las Recomendaciones P.10 [5], P.64, P.79 y 3.3.17, Parte C del Manual sobre telefonometría [6]) para convertir las sensibilidades de efecto local S_{meST} a S_{RNST} antes de calcular el LSTR, o se aplicó como una corrección ponderada al STMR según se describe en A.4.3.3/G.111 [7]. Los niveles de ruido ambiente fueron comparables a 55-59 dBA.

De acuerdo con estos resultados, en la Recomendación G.121 se indica que se debe tratar de lograr un valor de 13 dB para el LSTR.

Este valor de 13 dB se basa en un LSTR de 10 dB (que puede considerarse un valor mínimo), cuando no es posible ninguna otra mejora de la nota media de opinión aumentando el LSTR (véase la figura 3) más un margen de 3 dB que refleja el hecho de que en algunos entornos de oficina el ruido ambiente puede exceder los valores utilizados en estos experimentos. Otras pruebas indican también que pudiera ser más apropiado un valor mayor.

El valor satisfactorio para una conexión telefónica dada dependerá de factores tales como el nivel del ruido ambiente, el OLR de la conexión, los niveles de conversación utilizados, etc. En particular, los teléfonos modernos con transmisores lineales captan mejor el ruido ambiente de fondo. En este caso, puede ser conveniente que $LSTR > 15$ dB.

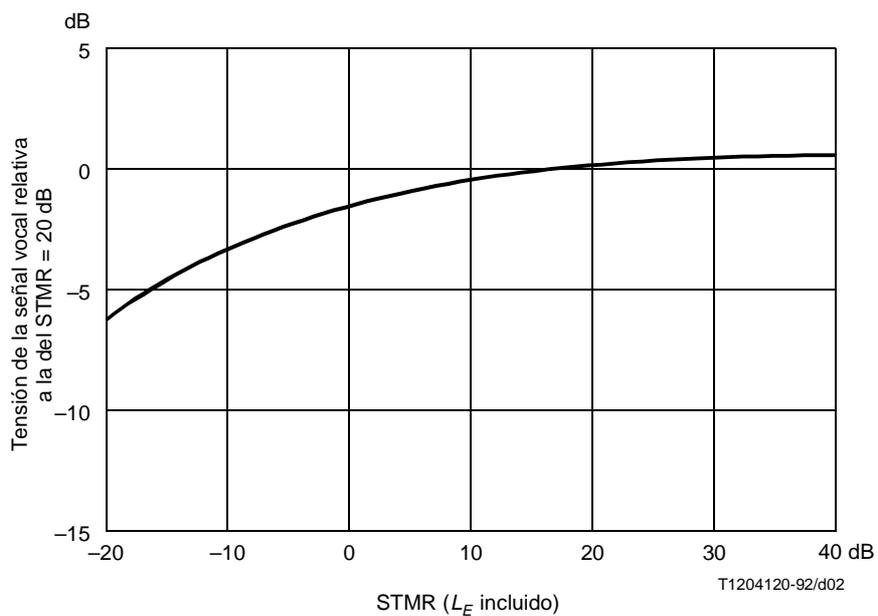


Figura 2 – Tensión de la señal vocal en función del STMR

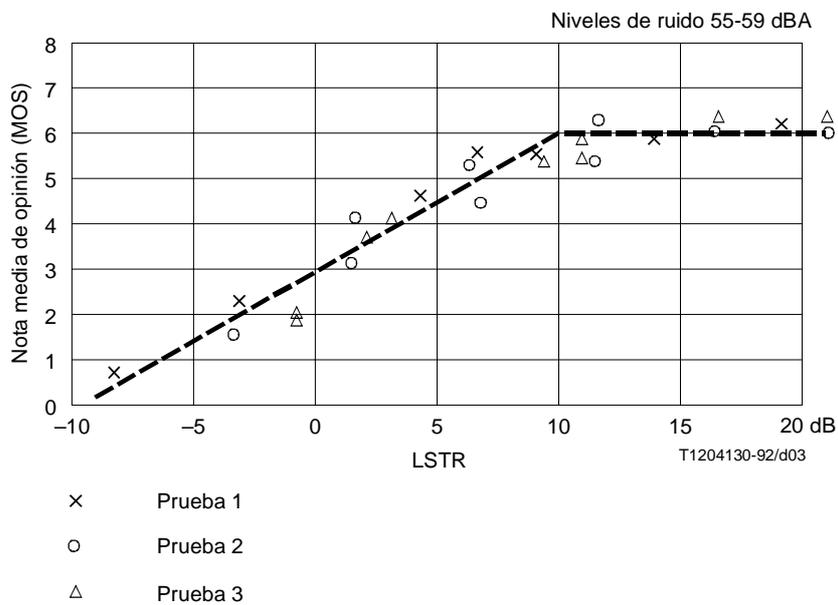


Figura 3 – MOS en función del LSTR calculada a partir de los resultados de diferentes pruebas

4 Relación entre el efecto local para el hablante y el efecto local para el oyente

4.1 Teléfonos con características lineales de efecto local

En los teléfonos con características lineales de ganancia o pérdida en el trayecto de efecto local, la geometría del microteléfono determina la relación entre los niveles de efecto local para el hablante y para el oyente. Aparentemente, los dos aspectos más importantes de la geometría del microteléfono son los siguientes: la distancia de la boca a la entrada del transmisor y el tamaño del obstáculo creado por el extremo del transmisor.

Para la entrada de la voz, un microteléfono con un extremo del transmisor grande situado cerca de la boca sufre una mayor presión acústica en la entrada del transmisor que un microteléfono con un extremo del transmisor ubicado más lejos de la boca (efecto de distancia) o uno que tenga un extremo de transmisor pequeño (efecto de obstáculo). Sin embargo, para las entradas en un campo de ruido ambiente difuso, la presión acústica en la entrada del transmisor es independiente del tamaño y de la forma del microteléfono. Así pues, si ambos microteléfonos poseen el mismo nivel STMR, uno de ellos, el del extremo de transmisor grande situado cerca de la boca, tendrá una ganancia eléctrica menor en su trayecto de efecto local, con lo cual el valor de LSTR será mayor.

Se ha mostrado que la diferencia entre los niveles de LSTR y STMR para una muestra de 26 aparatos telefónicos lineales guarda una elevada correlación con el logaritmo de la distancia entre la entrada del transmisor (el centro de la abertura externa del micrófono en la superficie del microteléfono) y el centro del anillo de labios de la boca artificial cuando el microteléfono está colocado en la posición LRGP de prueba (véase la Recomendación P.64). La relación empírica es la siguiente:

$$LSTR - STMR = 33 - 20 \log(d)$$

donde la distancia d , que va desde la entrada del transmisor hasta el centro del anillo de labios, se mide en mm. En esta relación puede haber pequeñas perturbaciones del orden de ± 1 dB, lo cual depende del tamaño del obstáculo que presenta el extremo del transmisor del microteléfono.

NOTA – Esta relación está basada en mediciones de teléfonos con microteléfonos relativamente convencionales. No se puede aplicar a los microteléfonos de gran tamaño ni a los auriculares del operador cuya entrada del transmisor está detrás de la posición equivalente de los labios.

4.2 Teléfonos con características no lineales de efecto local

En el trayecto de efecto local eléctrico se pueden utilizar características de ganancia o pérdida no lineales para aumentar la diferencia LSTR – STMR. Por ejemplo, es frecuente que los transmisores de carbón sean menos sensibles a los niveles de entrada más bajos de ruido ambiental que a los niveles de entrada más altos de la voz. Una característica de este tipo puede introducirse en teléfonos con micrófono lineal mediante la utilización de diversos circuitos de ganancia no lineal.

Si se utiliza la misma función de ganancia no lineal en los trayectos de emisión y de efecto local del teléfono, se puede obtener una diferencia LSTR – STMR aproximada midiendo la diferencia en las sensibilidades en emisión debida a las entradas de voz y de ruido ambiente, DELSM, como se describe en la Recomendación P.64. La diferencia STMR puede calcularse según el método dado en el anexo A/G.111. Sin embargo, si los trayectos de emisión y de efecto local no tienen las mismas características de ganancia no lineal (por ejemplo, circuito de control automático de ganancia en el trayecto de recepción que influencia el efecto local), el método DELSM dará resultados erróneos. En ese caso, los valores de LSTR y STMR deben medirse directamente.

5 Eco para el hablante con retardo corto percibido como efecto local

El eco para el hablante puede tener un efecto perjudicial sobre la calidad de transmisión con retardos de algunos milisegundos, aunque el retardo no sea suficientemente largo para que sea percibido como una señal de eco separada del efecto local. Estos ecos pueden producirse, por ejemplo, debido a reflexiones del puerto del circuito de enlace analógico de una centralita digital o en llamadas analógicas locales a través de una central digital. A menos que el híbrido que convierte la centralita digital o central a 4 hilos en un circuito analógico a 2 hilos esté bien adaptado, se producirá alguna reflexión. Debido a los tiempos del procesamiento digital, estas señales de eco para el hablante tienen un retardo de algunos milisegundos. El efecto local proporciona un enmascaramiento beneficioso del eco para el hablante con retardo corto de bajo nivel, pero a medida que aumenta el nivel del eco para el hablante, éste interactúa con el efecto local de una manera desagradable (efecto local con resonancia, efecto de lluvia en un barril, etc.).

El efecto objetivamente mensurable del eco para el hablante con retardo corto es que produce ondulaciones en la respuesta en frecuencia de efecto local. La señal de eco para el hablante reflejada se añade a la señal de efecto local directo con una relación de fase que aumenta la señal en algunas frecuencias y las disminuye en otras. El espaciamiento entre las ondulaciones es igual a la recíproca del retardo. Cuando la señal de eco para el hablante reflejada es pequeña con respecto al efecto local directo, las ondulaciones son pequeñas. A medida que aumenta la magnitud de la señal de eco para el hablante, aumentan las ondulaciones hasta que las crestas son de 6 dB por encima de la señal en fase y las depresiones son muy profundas debido a la compensación fuera de fase casi exacta. En niveles aún más altos de eco para el hablante (o niveles más bajos de efecto local) la magnitud de la ondulación disminuye de nuevo, pero la señal predominante es el eco para el hablante retardado.

Los usuarios perciben de modo diferente el eco para el hablante con retardo corto combinado con el efecto local que un nivel equivalente de efecto local puro, aunque es posible que no puedan detectar que está presente una señal de eco separada. Por tanto, una medida simple de efecto local, como el STMR, no es adecuada para describir el efecto de la señal combinada. El eco para el hablante, aun con tiempos de retardo muy cortos, debe tratarse como una degradación distinta de la calidad de transmisión. La Recomendación P.11 [8] y [9] contienen cierta orientación sobre cómo pueden tenerse en cuenta el efecto local y el eco para el hablante al predecir la calidad de una conexión telefónica, pero se sigue examinando este asunto.

Referencias

- [1] Recomendación P.76 del CCITT (1988), *Determinación de índices de sonoridad; principios fundamentales*.
- [2] Recomendación UIT-T G.121 (1993), *Índices de sonoridad de sistemas nacionales*.
- [3] Recomendación UIT-T P.64 (1997), *Determinación de las características de sensibilidad en función de la frecuencia de los sistemas telefónicos locales*.
- [4] Recomendación UIT-T P.79 (1993), *Cálculo de índices de sonoridad de aparatos telefónicos*.
- [5] Recomendación UIT-T P.10 (1998), *Vocabulario de términos sobre calidad de transmisión telefónica y aparatos telefónicos*.
- [6] *Manual sobre telefonometría*, UIT, Ginebra, 1993.
- [7] Recomendación UIT-T G.111 (1993), *Índices de sonoridad en una conexión internacional*.
- [8] Recomendación UIT-T P.11 (1993), *Efectos de las degradaciones de la transmisión*.
- [9] Adición al § 3.2.5 del Manual sobre telefonometría, *Voz artificial de conversación*, UIT, Ginebra, 1999.

VOZ ARTIFICIAL DE CONVERSACIÓN

La voz artificial descrita en la Recomendación P.50 se utiliza como señal de prueba unidireccional para evaluar la calidad de funcionamiento de, por ejemplo, códecs de conversación de baja velocidad. Es necesario ampliar el área de aplicación de estas señales artificiales para evaluar la calidad de dispositivos que deben ser operados por señales de conversación, tales como compensadores de eco, conmutadores de voz en teléfonos de altavoz y dispositivos de interpolación digital de la voz (DSI, *digital speech interpolation*). A tal fin, señales artificiales deben simular las conversaciones de los seres humanos. Por lo tanto, estas señales deben incluir no solamente periodos de conversación activa, realizados mediante la voz artificial de la Recomendación P.50, sino también periodos de pausa. Las señales artificiales deben asimismo ser señales bidireccionales a fin de simular un entorno de «hablante y oyente».

Una señal que alternativamente ubica voz artificial de la Recomendación P.50 de duración arbitraria y silencio (secuencia de ceros) de duración arbitraria en un eje de tiempo, no es adecuada debido a que la conversación humana tiene características propias. Por ejemplo, los seres humanos no pueden hablar durante un largo periodo de tiempo sin hacer una pausa y se dan muy pocos casos de dos personas que hablen simultáneamente durante un tiempo prolongado.

Por lo tanto, en las señales artificiales deben simularse algunas características estadísticas temporales de la señal de conversación real. Una señal artificial que satisfaga estas características se denomina «señal de conversación artificial».

La figura 1 ilustra varios diagramas temporales de referencia de la potencia que se produce cuando dos personas hablan entre sí. En las voces artificiales de conversación, existen elementos que caracterizan correctamente la señal de conversación real, a saber, las características estadísticas tales como índices de ocurrencia, valores medios y distribución acumulada de la duración de los periodos de conversación, de pausa, de doble conversación y de silencio mutuo.

La figura 2 ilustra un modelo de transmisión de estado para la generación de la voz artificial de conversación. Una conversación puede estar en uno de los cuatro estados siguientes: conversación individual (sólo habla un partícipe y el otro permanece en silencio, y viceversa), conversación simultánea y silencio mutuo, empleándose un modelo de transición de estados entre los mismos. Después de permanecer en un estado durante un tiempo T , cambia de un estado a otro con una probabilidad de transición p_i . Las distribuciones acumuladas de la duración T de cada estado se suponen exponenciales y la duración T de cada estado varía según una variable aleatoria x . Durante los periodos de silencio, la salida se compone de secuencias de ceros. Durante los periodos de conversación, la salida está formada por la voz artificial de la Recomendación P.50.

Los parámetros de la figura 2 que deben optimizarse son los siguientes: la duración media estimada de la conversación individual, de la conversación simultánea y del silencio mutuo, así como la probabilidad de transición p_1 . Las probabilidades de transición p_2 y p_3 se fijan al 50%. Dichos parámetros fueron optimizados para minimizar la desviación media de los siguientes parámetros en señales artificiales con respecto a la señal de conversación real: duración media y régimen de ocurrencia de los intervalos de habla, pausa, conversación simultánea y silencio mutuo. A fin de conseguir los valores objetivo para la optimización de los parámetros, se utilizaron valores medios medidos para señales de conversación en japonés, inglés americano e italiano. En la Recomendación P.59 se especifican los valores objetivo para la optimización de parámetros así como los valores óptimos de los mismos.

Las características temporales estadísticas de la voz artificial de conversación tienden a ser estables conforme aumenta la duración de las señales. Las características a largo plazo de la voz artificial de conversación tienen un tiempo de convergencia de diez minutos, lo cual se corresponde con las características a largo plazo de la señal de conversación humana de aproximadamente diez horas. Ello revela la efectividad de la utilización en las pruebas de la voz artificial de conversación en lugar de la señal de conversación humana.

La voz artificial de conversación no se genera normalmente mediante una implementación en tiempo real del proceso de la Recomendación P.59, sino que generalmente está almacenada en memoria de sólo lectura (ROM), o en medios de almacenamiento magnético u óptico. En ambos casos, normalmente se proporcionan sistemas de calibración. Éstos constan de secuencias de tonos con una relación conocida con el nivel de la voz artificial de conversación.

Procedimiento

La voz artificial de conversación no requiere calibración. Sólo precisa un ajuste de nivel mediante la utilización de una secuencia de tonos que está disponible en soporte analógico o digital (ROM, CD, DAT, MO, registros de vídeo MIC o cintas analógicas).

La voz artificial de conversación puede utilizarse como una señal eléctrica o puede alimentar una boca artificial a fin de obtener una voz artificial acústica de conversación. En este último caso, la boca artificial debe incluir un igualador. La exactitud del igualador de boca se verifica midiendo el espectro medio de la voz artificial acústica de conversación en el punto de referencia boca (MRP).

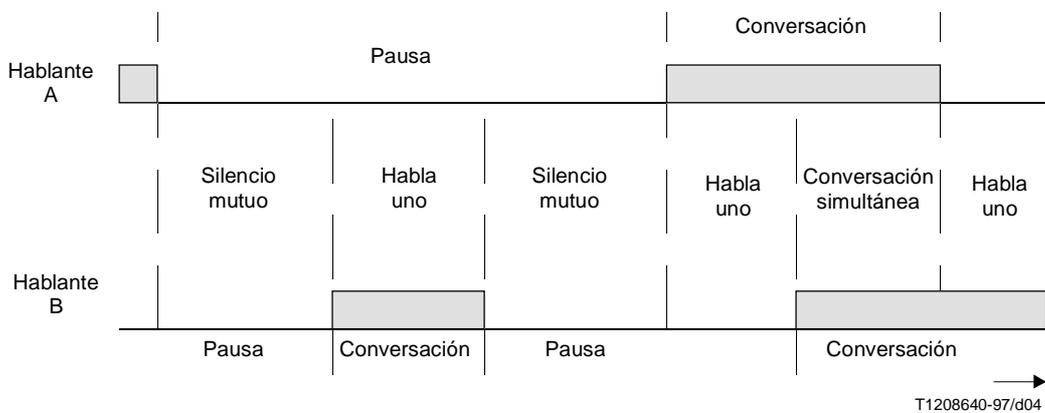
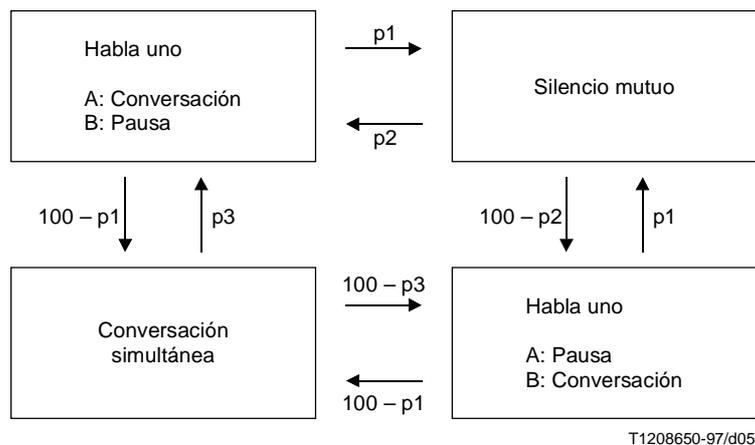


Figura 1 – Diagrama de potencia temporal en conversación



p_i Probabilidad de transición
 T Duración en cada estado
 \hat{T} Duración media estimada
 $0 < x < 1$ Variable aleatoria
 $T = -\hat{T} \ln(1 - x)$

Figura 2 – Modelo de transición de estados para conversación

MODELOS DE CÁLCULO PARA LA ESTIMACIÓN DE LA OPINIÓN DE LOS CLIENTES SOBRE LA CALIDAD DE LAS COMUNICACIONES DE CONVERSACIÓN EN LAS REDES TELEFÓNICAS

1 Introducción

Es de gran importancia proporcionar a los clientes, con una alta probabilidad, una calidad de transmisión adecuada de telefonía vocal. Se ha determinado que los métodos tradicionales de planificación de la transmisión, que asignan una gama de valores posibles a cada uno de los parámetros de transmisión individuales, no son plenamente adecuados para las redes modernas. Algunas de las razones de ello son la introducción de nuevas tecnologías, tales como los códecs de baja velocidad, y la actual desregulación de los mercados.

En estas condiciones cambiantes, los operadores de red necesitan directrices así como un método satisfactorio para evaluar, incluso en la etapa de planificación, la calidad de transmisión de las conexiones y para predecir si el usuario quedará satisfecho. La metodología debe basarse en valores calculados y en valores de planificación, ya que no es posible que mediante medidas pueda verificarse la calidad de funcionamiento de todos los posibles trayectos a través de redes reales o simuladas.

Ello ha creado un renovado interés en las estimaciones de opinión mediante modelos de cálculo que manejen los efectos combinados de diferentes tipos de degradación de la transmisión. A continuación se hace una breve presentación de los modelos que han sido utilizados en el ámbito técnico del UIT-T para ayudar en la planificación de la transmisión. Sin embargo, antes de presentar los distintos modelos de cálculo, se discute el concepto de «opinión del cliente» a fin de evitar interpretaciones erróneas sobre las pruebas subjetivas y los estudios sobre los clientes.

2 Las «opiniones de los clientes» como base para los modelos de cálculo

La parte crucial en el desarrollo de un modelo de cálculo es la obtención de la información básica sobre las opiniones que los clientes tienen de la calidad de una conversación telefónica, y en particular, lo relacionado con los diversos parámetros de transmisión involucrados.

Un problema general es que las opiniones sobre la calidad de la conversación varían de unas personas a otras e, igualmente, varían según el momento y el contexto y circunstancias bajo las que se desarrolla la comunicación.

La respuesta del cliente también se verá influenciada por la forma en que se hagan las preguntas sobre la calidad. Un ejemplo clásico de ello lo constituye los niveles de audición. Los experimentos diseñados para establecer «cual es el nivel de audición más confortable» conducen a un resultado 3 dB inferior a si el objetivo estuviera orientado a obtener respuesta a la pregunta «a qué nivel se consigue la mayor calidad de conversación».

Un método directo de obtener la opinión de la gente es, por supuesto, *realizar estudios sobre los clientes* en redes reales. Sin embargo, durante una llamada telefónica la gente normalmente no evalúa la calidad de la transmisión de la voz de forma consciente, sino que da por sentado que se produce un cierto nivel de calidad, en contraste con lo que ocurre cuando se escucha música en un equipo de alta fidelidad. (Sólo si la calidad de la conversación cae frecuentemente por debajo de un cierto nivel suele reaccionar el usuario, probablemente quejándose bastante enérgicamente al operador de la red.) Por lo tanto, idealmente un cliente debe ser entrevistado justamente al terminar una llamada de la que se conocen más o menos en su totalidad los parámetros de transmisión. Dichas investigaciones se han realizado y se continúan realizando incluso introduciendo en algunos casos degradaciones intencionadas de la transmisión. El proceso es bastante incómodo y costoso pero debiera generar los datos más fiables.

En una primera instancia se suelen utilizar métodos más económicos para la evaluación subjetiva de las degradaciones de la transmisión, normalmente en la forma de *pruebas subjetivas en condiciones controladas* en laboratorio. Dichas pruebas se realizan con un conjunto limitado y específico de variaciones de parámetros de transmisión y un número limitado de personas en la prueba (en las Recomendaciones de la serie P se describe como deben realizarse dichas pruebas). Normalmente, los resultados se presentan como notas medias de opinión (MOS, *mean opinion scores*). Sin embargo, la definición de las condiciones de las pruebas y la interpretación de los resultados deben realizarse con sumo cuidado.

Los valores de MOS para una prueba subjetiva se presentan a menudo con «límites de confianza» estadísticos. No obstante, debe recordarse que estos límites sólo se aplican a dicha instancia de prueba en concreto. El valor del MOS de una prueba no debe considerarse como un valor absoluto, sino relativo. Por lo tanto, una prueba objetiva correctamente dispuesta debe incluir una referencia a las condiciones de referencia, es decir, es preciso que las personas que participen en la prueba juzguen degradaciones de naturaleza conocida. Utilizando este MOS de referencia, los valores MOS de prueba reales se normalizan de forma que los resultados de diferentes laboratorios de prueba se puedan comparar (es decir, se podría decir que los equipos de prueba están «calibrados»). Para producir las degradaciones de referencia se utiliza a menudo un aparato de referencia de ruido modulado (MNRU, *modulated noise reference unit*) MOS. Éste se ha convertido en el procedimiento normalizado cuando se evalúan códecs de baja velocidad mediante pruebas subjetivas (el hecho de que una «degradación de referencia» suene a menudo de forma bastante distinta a una «degradación de prueba» constituye una fuente de incertidumbre).

A título de ejemplo, la figura 1 muestra cuatro curvas de MNRU de referencia obtenidas por cuatro laboratorios de pruebas en un proyecto común para evaluar un determinado códec LD-CELP de baja velocidad (véase la Recomendación G.729). Como puede apreciarse, la distribución de los valores de MOS es de casi 1,0 unidades, lo cual indica la necesidad de normalizar los valores de MOS medidos por los distintos laboratorios para el códec real bajo prueba.

Teniendo en cuenta todos los elementos necesarios, un modelo de cálculo para la predicción de opiniones debe ajustarse a un compromiso entre las distintas pruebas subjetivas y estudios de clientes. El primer paso consiste en utilizar los resultados de las pruebas subjetivas de laboratorio con un cierto «enjuiciamiento ingenieril». Un segundo paso lo constituye la verificación de los resultados del modelo mediante comparaciones con resultados obtenidos de estudios de clientes en condiciones de red reales.

En la literatura técnica se describen varios modelos de cálculo que se presentan a continuación. Debido a la variabilidad de los estudios y de las pruebas subjetivas, no debiera extrañar que las predicciones de los mismos puedan dar lugar a resultados ligeramente distintos bajo las mismas condiciones de red.

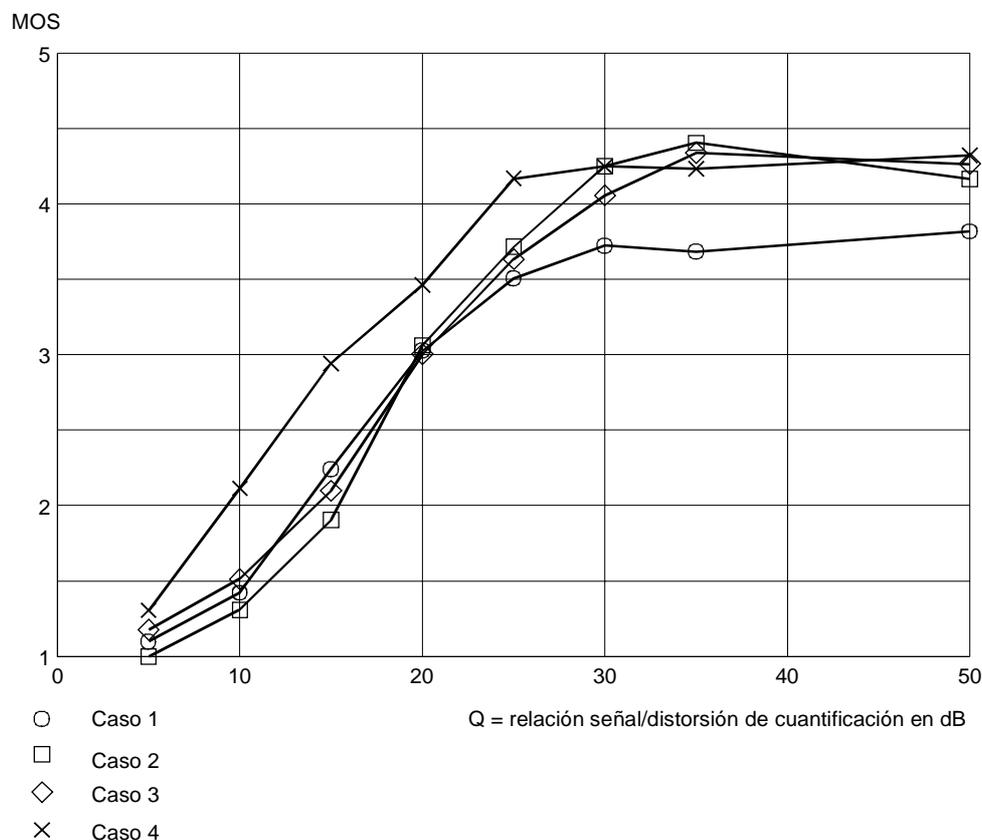


Figura 1 – Curvas MOS de referencia, utilizando las MNRU, de cuatro laboratorios de pruebas subjetivas

T1208630-97/d06

3 Modelos de cálculo descritos en publicaciones del UIT-T

3.1 Generalidades

El Suplemento 3 (1993) a las Recomendaciones de la serie P contienen descripciones de cuatro distintos métodos de cálculo para la predicción de la calidad de transmisión a partir de mediciones objetivas, en concreto:

- El modelo de «índices de transmisión», una contribución de los Estados Unidos y Canadá.
- El modelo «CATNAP83», de British Telecom.
- El modelo de «índice de información», una contribución de Francia.
- El modelo «OPINE» de la NTT.

El anexo A/P.11 (*Libro Azul*) describe el modelo de «índice de calidad de transmisión», que constituye una síntesis de los cuatro modelos del Suplemento 4 de las Recomendaciones de la serie P, aunque con simplificaciones y con un número restringido de parámetros.

Todos estos modelos fueron desarrollados antes de la actual tendencia de desregulación de los mercados de telecomunicación, es decir, en una situación en la que las Administraciones de telecomunicación podían controlar sus redes con un nivel de detalle técnico elevado. La atenuación y el ruido eran las degradaciones principales y no se consideraba el impacto de códecs de baja velocidad avanzados, al menos no con una gran precisión.

El apéndice I/G.101 (1996) describe el «modelo-E» que ha sido desarrollado por un grupo de trabajo del ETSI, Instituto Europeo de Normas de Telecomunicación (*European Telecommunications Standards Institute*). Éste utiliza algunos algoritmos y conceptos de los modelos descritos en el Suplemento 3 de las Recomendaciones de la serie P, complementados con los resultados de pruebas subjetivas recientes. Entre las nuevas características incluidas destacan los efectos subjetivos de los códecs de baja velocidad y la posibilidad de juzgar las opiniones de los clientes sobre la calidad de la conversación en relación con sus expectativas de calidad del medio de comunicaciones.

3.2 El modelo «índice de transmisión»

Los parámetros incluidos son (OR), ruido del circuito, ruido ambiente en el lado de recepción, STMT, qdu (ruido de cuantificación), anchura de banda y distorsión de atenuación, eco del oyente, eco del hablante.

Los resultados de cálculo obtenidos para una conexión se presentan como «índices-R», los cuales pueden ser transformados en porcentajes de clientes que pueden encontrar que la conexión es «buena o mejor» (GOB, *good or better*), o «pobre o peor» (POW, *poor or worse*).

Los comentarios siguientes en la descripción del modelo son de especial interés: «Un motivo importante para la introducción de la escala R ha sido la observación de que las pruebas subjetivas pueden verse afectadas por diversos factores tales como el grupo sujeto a la prueba, el tipo de prueba y la gama de condiciones de prueba que se incluyen en la misma. Se ha detectado que estos factores pueden causar cambios en las notas medias de opinión en una condición determinada, así como en la desviación típica. Por lo tanto, no es fácil establecer una relación biunívoca entre una determinada condición de transmisión y una opinión subjetiva en términos de nota media de opinión o de porcentaje de índices que se califican como buenos o excelentes. La introducción de una escala de índices de transmisión tiende a reducir dichas dificultades, ...»

Los datos de opinión subjetiva han sido recopilados de pruebas de laboratorio así como de estudios de campo bastante elaborados.

Nótese que todas las pruebas se realizaron utilizando microteléfonos del tipo Western Electric-500 o equivalentes.

3.3 El modelo «CATNAP83»

Este modelo centra su atención en los efectos subjetivos de la atenuación de los circuitos, la distorsión atenuación-frecuencia, el ruido del circuito, ruido de cuantificación, ruido ambiente y los trayectos del efecto local, para una gama de valores razonablemente amplia de estas características en cualquier combinación posible.

La estructura del modelo permite que el proceso de evaluación refleje las relaciones causa-efecto que relacionan las condiciones de entrada (propiedades de la conexión, entorno acústico, características de las capacidades auditivas, sonidos de conversación e idioma de los participantes, etc.) con los resultados obtenidos (satisfacción de los participantes o estimación de la calidad de funcionamiento).

En la utilización práctica del modelo (en la forma de un programa de computadora), deben especificarse las conexiones en forma de elementos y cantidades tales como niveles de ruido, tipos específicos de teléfonos, longitud de los cables con indicación de la resistencia y capacitancia por kilómetro, así como atenuadores con pérdidas específicas. El programa calcula índices de sonoridad, niveles de conversación y notas de opinión. Éstas últimas se ofrecen como nota de esfuerzo de audición (Y_e) y nota media de conversación (Y_c).

3.4 El modelo «índice de información»

La teoría de este modelo también considera las características fundamentales de la relación causa-efecto, tal como hace el modelo CATNAP. Tiene en cuenta la atenuación de transmisión, el ruido del circuito, el ruido ambiente, la distorsión atenuación/frecuencia, efecto local y diversas distorsiones que tienen lugar en la transmisión digital.

El resultado del modelo es el «índice de información» (I). Junto a la descripción del modelo se presentan varios ejemplos que comparan los valores calculados de I con los valores subjetivos de MOS.

3.5 El modelo «OPINE»

El modelo OPINE analiza la atenuación de transmisión, el ruido del circuito, la distorsión atenuación/frecuencia, la distorsión de cuantificación, el eco del hablante y el efecto local. Hace un modelado del proceso psicológico-auditivo que realizan los seres humanos sobre la calidad telefónica basado en dichos factores (por lo tanto, es del mismo tipo genérico que los modelos CATNAP e índice de información).

Se han elegido cinco elementos psicológicos que afectan a la calidad de una conversación telefónica:

- 1) Distorsión de la conversación en función de la distorsión atenuación/frecuencia.
- 2) Atenuación o exceso de la sonoridad efectiva en la conversación.
- 3) Ruido existente durante los intervalos de conversación y los de silencio.
- 4) Degradación producida por el eco del hablante.
- 5) Degradación producida por el efecto local.

Cada elemento psicológico se asocia con un índice de calidad (PI, *performance index*). El MOS para una conexión se estima a partir de la suma de todos los PI.

3.6 El modelo de «índice de calidad de la transmisión»

Tal como se indica en el Suplemento 3 de las Recomendaciones de la serie P, se trata de un modelo de opinión de conversación sencillo para predecir el efecto combinado del OLR y del ruido psfométrico. Incluye también la influencia del efecto local (STMR), el ruido ambiente y la distorsión de atenuación.

El modelo predice el MOS, porcentajes de «bueno y excelente» y de «pobre y malo»,

3.7 El «modelo E»

Un grupo de trabajo del ETSI (Instituto Europeo de Normas de Telecomunicación), ha trabajado sobre un informe técnico relativo a la calidad de transmisión de la conversación extremo a extremo en redes telefónicas (denominado ETR 250). Se ha desarrollado un modelo de cálculo para la estimación de la percepción de la calidad de la comunicación de conversación de los usuarios de la red para una conexión, utilizando parámetros de transmisión «convencionales» de la red, así como «factores de degradación del equipo» especiales para los códecs involucrados. Los resultados se han incorporado en el UIT-T conjuntamente con las revisiones de las Recomendaciones G.101 y G.113.

El modelo del ETSI (modelo E) se fundamenta en gran medida en el modelo de «índices de transmisión» (TR, *transmission ratings*) descrito en 3.2, aunque hace uso de características propias de otros modelos incluidos en el Suplemento 3 de las Recomendaciones de la serie P. Sin embargo, la estructura es diferente a la del modelo TR. El principio fundamental del modelo ETSI se basa en un concepto que estableció J. Allnatt hace más de 20 años y que es utilizado en el modelo OPINE de la NTT: «los factores psicológicos de una escala psicológica son aditivos».

El modelo ETSI combina el efecto de los diversos parámetros de transmisión en un factor índice R , a partir del cual puede predecirse la reacción del usuario, por ejemplo, porcentajes de conexiones que se consideran «buena o mejor», «pobre o peor» o incluso tan mala que provocarían una terminación temprana de la llamada, así como cuales serían las notas que se otorgarían en una experimento MOS. El factor R se compone de los términos siguientes:

$$R = R_o - I_s - I_d - I_e + A$$

- R_o representa, en principio, la relación básica señal de voz/ruido;
- I_s, I_d, I_e también se denominan factores de degradación;
- I_s representa degradaciones que tienen lugar simultáneamente con la señal de voz, tal como una conexión de sonido demasiado alto, efecto local elevado y distorsión de cuantificación de sistema MIC;
- I_d representa degradaciones retardadas, tales como el eco del oyente y del hablante, así como un retardo absoluto demasiado elevado;
- I_e representa las degradaciones de transmisión causadas por equipos especiales tales como determinados códecs de baja velocidad, equipos multiplicadores de circuitos digitales (DCME, *digital circuit multiplication equipment*), etc. (Este factor es un concepto nuevo);
- A se denomina «factor de expectativa». Representa la «ventaja de acceso» de que gozan determinados sistemas, en particular los sistemas de telefonía móvil, en relación a los sistemas fijos convencionales por hilos. (La percepción de «buena calidad» está íntimamente ligada a la satisfacción de las expectativas de los clientes; en concreto un área es la cubierta por los conceptos de «usabilidad» «utilidad» y «actitudes». El factor de expectativa es un nuevo concepto que no había sido empleado en otros modelos de cálculo.)

El modelo E no incluye los efectos específicos de la distorsión de atenuación y de la respuesta en frecuencia del teléfono. El motivo de ello es que en un mercado desregulado el operador no conoce con gran precisión dichos parámetros de forma que es suficiente utilizar los valores planificados de los índices de sonoridad. Sin embargo, el manejo del microteléfono por parte del usuario introduce una gran variabilidad en la respuesta en frecuencia de las pérdidas del auricular.

En la práctica, el modelo ETSI proporciona casi los mismos valores R que el modelo TR para las degradaciones «convencionales». Asimismo, el modelo TR puede disponerse de tal forma que tenga la misma estructura de términos aditivos para las degradaciones incluidas en el mismo. Los resultados de la aplicación del modelo ETSI a conexiones típicas están en consonancia con los resultados de otros modelos y de pruebas subjetivas que han sido publicadas. En particular, las degradaciones motivadas por códecs de baja velocidad pueden ser predichas bastante bien por el modelo, incluso mejor que la metodología hasta ahora utilizada de las «unidades de distorsión de cuantificación». (En una planificación real, la forma más conveniente de utilizar el modelo E es poniendo límites a la suma de los factores de degradación tal como se hace en la Recomendación G.113.)

En el anexo E/G.113 se describe como evaluar los códecs de baja velocidad, es decir, como transformar los valores MOS que se han obtenido de forma subjetiva en «factores de degradación del equipo».

La naturaleza aditiva del modelo E hace que sea sencillo realizar una actualización si la evidencia de pruebas subjetivas con parámetros «antiguos» indica que ello es necesario, o si se necesita analizar nuevos tipos de degradaciones.

4 Desarrollos posteriores relativos a modelos «cognoscitivos y de percepción»

Esta es un área en la que hay en marcha una abundante investigación y constituye objeto de estudio en la Comisión de Estudio 12 del UIT-T. El principal motivo es la necesidad de desarrollar métodos fiables y objetivos para la evaluación de los códecs de baja velocidad (de los que continuamente se están introduciendo nuevos modelos). Actualmente, las características de funcionamiento de dichos códecs debe ser investigada mediante costosas pruebas subjetivas a fin de obtener resultados realistas y fiables.

