



UNIÓN INTERNACIONAL DE TELECOMUNICACIONES

**UIT-T**

SECTOR DE NORMALIZACIÓN  
DE LAS TELECOMUNICACIONES  
DE LA UIT

**P.330**

(03/2003)

SERIE P: CALIDAD DE TRANSMISIÓN TELEFÓNICA,  
INSTALACIONES TELEFÓNICAS Y REDES LOCALES

Líneas y aparatos de abonado

---

**Dispositivos de procesamiento de señales  
vocales para mejorar la acústica**

Recomendación UIT-T P.330

---

RECOMENDACIONES UIT-T DE LA SERIE P

**CALIDAD DE TRANSMISIÓN TELEFÓNICA, INSTALACIONES TELEFÓNICAS Y REDES LOCALES**

Vocabulario y efectos de los parámetros de transmisión sobre la opinión de los clientes	Serie	P.10
<b>Líneas y aparatos de abonado</b>	<b>Serie</b>	<b>P.30</b>
		<b>P.300</b>
Patrones de transmisión	Serie	P.40
Aparatos para mediciones objetivas	Serie	P.50
		P.500
Medidas electroacústicas objetivas	Serie	P.60
Medidas relativas a la sonoridad vocal	Serie	P.70
Métodos de evaluación objetiva y subjetiva de la calidad	Serie	P.80
		P.800
Calidad audiovisual en servicios multimedios	Serie	P.900

*Para más información, véase la Lista de Recomendaciones del UIT-T.*

## **Recomendación UIT-T P.330**

### **Dispositivos de procesamiento de señales vocales para mejorar la acústica**

#### **Resumen**

En esta Recomendación se describen las características de transmisión genéricas, la calidad de funcionamiento y los principios de prueba de los dispositivos de procesamiento de señales vocales para mejorar la acústica (SPDA), que se utilizan en los terminales, sin importar el tipo de aplicación.

Un dispositivo de procesamiento de señales vocales para mejorar la acústica se define como cualquier función de procesamiento de señales integrada en los terminales que permite mejorar las señales vocales. Estas funciones incluyen el control del eco acústico y la reducción de ruido.

La finalidad de esta Recomendación es definir un marco de trabajo para especificar las restricciones de calidad de funcionamiento de los terminales que incluyen SPDA y, cuando proceda, definir las pruebas que se puedan llevar a cabo en esos terminales para verificar que se cumplen las restricciones. Esta Recomendación comprende las características genéricas aplicables a terminales tanto analógicos como digitales.

#### **Orígenes**

La Recomendación UIT-T P.330 preparada por la Comisión de Estudio 12 (2001-2004) del UIT-T, fue aprobada por el procedimiento de la Resolución 1 de la AMNT el 16 de marzo de 2003.

## PREFACIO

La UIT (Unión Internacional de Telecomunicaciones) es el organismo especializado de las Naciones Unidas en el campo de las telecomunicaciones. El UIT-T (Sector de Normalización de las Telecomunicaciones de la UIT) es un órgano permanente de la 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 Asamblea Mundial de Normalización de las Telecomunicaciones (AMNT), 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 1 de la AMNT.

En ciertos sectores de la tecnología de la información que corresponden a la esfera de competencia del UIT-T, se preparan las normas necesarias en colaboración con la ISO y la CEI.

## 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.

La observancia de esta Recomendación es voluntaria. Ahora bien, la Recomendación puede contener ciertas disposiciones obligatorias (para asegurar, por ejemplo, la aplicabilidad o la interoperabilidad), por lo que la observancia se consigue con el cumplimiento exacto y puntual de todas las disposiciones obligatorias. La obligatoriedad de un elemento preceptivo o requisito se expresa mediante las frases "tener que, haber de, hay que + infinitivo" o el verbo principal en tiempo futuro simple de mandato, en modo afirmativo o negativo. El hecho de que se utilice esta formulación no entraña que la observancia se imponga a ninguna de las partes.

## PROPIEDAD INTELECTUAL

La UIT señala a la atención la posibilidad de que la utilización o aplicación de la presente Recomendación suponga el empleo de un derecho de propiedad intelectual reivindicado. La UIT no adopta ninguna posición en cuanto a la demostración, validez o aplicabilidad de los derechos de propiedad intelectual reivindicados, ya sea por los miembros de la UIT o por terceros ajenos al proceso de elaboración de Recomendaciones.

En la fecha de aprobación de la presente Recomendación, la UIT no ha recibido notificación de propiedad intelectual, protegida por patente, que puede ser necesaria para aplicar esta Recomendación. Sin embargo, debe señalarse a los usuarios que puede que esta información no se encuentre totalmente actualizada al respecto, por lo que se les insta encarecidamente a consultar la base de datos sobre patentes de la TSB.

© UIT 2003

Reservados todos los derechos. Ninguna parte de esta publicación puede reproducirse por ningún procedimiento sin previa autorización escrita por parte de la UIT.

## ÍNDICE

	<b>Página</b>
1 Alcance .....	1
2 Referencias .....	1
3 Términos y definiciones .....	2
4 Procesamiento de eco acústico .....	3
4.1 Componentes analógicos .....	3
4.2 Unidades funcionales de un compensador de eco acústico .....	4
4.3 Interacción entre el control de eco del terminal y el equipo de red de procesamiento de las señales .....	4
4.4 Retardo .....	5
4.5 Especificaciones de control de eco acústico.....	5
5 Reducción de ruido .....	8
5.1 Componentes analógicos .....	8
5.2 Unidades funcionales de un sistema de reducción de ruido .....	9
5.3 Interacción entre el sistema de reducción de ruido en el terminal y el equipo de procesamiento de señales en la red .....	10
5.4 Tiempo de procesamiento de reducción de ruido .....	10
5.5 Especificaciones del sistema de reducción de ruido.....	10



## Recomendación UIT-T P.330

### Dispositivos de procesamiento de señales vocales para mejorar la acústica

#### 1 Alcance

En esta Recomendación se describen las características de transmisión genéricas, la calidad de funcionamiento y los principios de prueba de los dispositivos de procesamiento de señales vocales para mejorar la acústica (SPDA, *speech processing devices for acoustic enhancement*) que se utilizan en los terminales, sin importar el tipo de aplicación.

Un dispositivo de procesamiento de señales vocales para mejorar la acústica se define como cualquier función de procesamiento de señales integrada en los terminales que permite mejorar las señales vocales. Estas funciones incluyen el control del eco acústico y la reducción de ruido. La supresión de la reverberación y cualquier procesamiento de señales avanzado para la captación y la restitución multicanal quedan en estudio.

La finalidad de esta Recomendación es definir un marco de trabajo para especificar restricciones de calidad de funcionamiento de los terminales que incluyen SPDA y, cuando proceda, definir las pruebas que se puedan llevar a cabo en esos terminales para verificar que se cumplen las restricciones. Esta Recomendación comprende las características genéricas aplicables a terminales tanto analógicos como digitales. Los requisitos que se aplican sólo a los terminales manos libres se describen en la Rec. UIT-T P.340 [13].

Los métodos de prueba apropiados para los parámetros definidos en esta Recomendación se describen en la Rec. UIT-T P.502 [15].

La utilización del simulador de cabeza y torso (HATS, *head and torso simulator*) para fines de prueba se describe en la Rec. UIT-T P.581 [16].

#### 2 Referencias

Las siguientes Recomendaciones del UIT-T y otras referencias 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 otras referencias son objeto de revisiones por lo que se preconiza que los usuarios de esta Recomendación investiguen la posibilidad de aplicar las ediciones más recientes de las Recomendaciones y otras referencias citadas a continuación. Se publica periódicamente una lista de las Recomendaciones UIT-T actualmente vigentes. En esta Recomendación, la referencia a un documento, en tanto que autónomo, no le otorga el rango de una Recomendación.

- [1] Recomendación UIT-T G.114 (2003), *Tiempo de transmisión en un sentido*.
- [2] Recomendación UIT-T G.121 (1993), *Índices de sonoridad de sistemas nacionales*.
- [3] Recomendación UIT-T G.122 (1993), *Influencia de los sistemas nacionales en la estabilidad y el eco para la persona que habla en las conexiones internacionales*.
- [4] Recomendación UIT-T G.131 (1996), *Control del eco para el hablante*.
- [5] Recomendación UIT-T G.167 (1993), *Controladores de eco acústico*.
- [6] Recomendación UIT-T G.168 (2002), *Compensadores de eco de redes digitales*.
- [7] CCITT Recomendación G.223 (1988), *Hipótesis para el cálculo del ruido en los circuitos ficticios de referencia para telefonía*.

- [8] Recomendación UIT-T P.10 (1998), *Vocabulario de términos sobre calidad de transmisión telefónica y aparatos telefónicos.*
- [9] Recomendación UIT-T P.50 (1999), *Voces artificiales.*
- [10] Recomendación UIT-T P.51 (1996), *Boca artificial.*
- [11] Recomendación UIT-T P.78 (1996), *Método de prueba subjetivo para determinar índices de sonoridad de acuerdo con la Recomendación P.76.*
- [12] Recomendación UIT-T P.79 (1999), *Cálculo de índices de sonoridad de los aparatos telefónicos.*
- [13] Recomendación UIT-T P.340 (2000), *Características de transmisión y parámetros de calidad vocal de los aparatos telefónicos manos libres.*
- [14] Recomendación UIT-T P.501 (2000), *Señales de prueba para utilización en telefonometría.*
- [15] Recomendación UIT-T P.502 (2000), *Métodos de prueba objetivos para los sistemas de comunicación vocal con señales de prueba complejas.*
- [16] Recomendación UIT-T P.581 (2000), *Uso del simulador de cabeza y torso para la prueba de terminales manos libres.*
- [17] Recomendación UIT-T P.800 (1996), *Métodos de determinación subjetiva de la calidad de transmisión.*
- [18] Recomendación UIT-T P.832 (2000), *Evaluación subjetiva de la calidad de funcionamiento de los terminales manos libres.*
- [19] Manual sobre Telefonometría del UIT-T, 1992.
- [20] Recomendación UIT-T G.161 (2002), *Aspectos de la interacción de equipos de red de procesamiento de señal.*
- [21] Recomendación UIT-T G.108.2 (2003), *Asuntos relativos a la planificación de la transmisión con compensadores de eco.*

### 3 Términos y definiciones

Se utilizan las definiciones pertinentes dadas en [8], así como las siguientes:

**3.1 dispositivo de procesamiento de señales vocales para mejorar la acústica (SPDA, *speech processing device for acoustic enhancement*):** Se define como cualquier función de procesamiento de señales integrada en los terminales, que permite mejorar las señales vocales.

**3.2 eco acústico (AE, *acoustic echo*):** El eco acústico es la señal retardada y reflejada, que resulta del trayecto acústico entre el auricular/altavoz y el micrófono de un dispositivo portátil o de un terminal manos libres.

**3.3 compensador de eco acústico (AEC, *acoustic echo canceller*):** Dispositivo que reduce el nivel de eco acústico, con efectos despreciables sobre la voz de los usuarios local y distante. Para poder seguir la variación del trayecto de eco acústico, el control de eco acústico suele implementarse mediante la identificación adaptativa de la respuesta al impulso en el trayecto de eco acústico.

**3.4 controlador de atenuación:** Dispositivo que reduce el nivel de eco acústico introduciendo atenuaciones variables en las señales de audio recibidas y/o transmitidas.

**3.5 procesador no lineal (NLP, *non-linear processor*):** Dispositivo que reduce el eco residual no cancelado por el control de eco acústico. El NLP utiliza procesamiento no lineal para suprimir el eco a un nivel no percibido por el individuo en el extremo distante de la conversación. Un limitador de amplitud en el centro es un ejemplo de un dispositivo de esta clase.

**3.6 dispositivo suplementario de control de silbido:** Dispositivo que modifica algunas características de las señales transmitidas y/o recibidas para mejorar el margen de estabilidad del terminal. Esta función suele implementarse mediante un procesador armónico. Para evitar perturbaciones en la red, conviene no utilizar estos dispositivos en los terminales que han de utilizarse en conexiones con compensadores de eco eléctrico de red conformes a la Rec. UIT-T G.168, ya que éstos no pueden funcionar adecuadamente en trayectos de eco variables en el tiempo (por ejemplo, desplazamiento de frecuencia).

**3.7 ruido de fondo:** El ruido de fondo se define como la señal añadida a la señal vocal deseada en el extremo próximo. El ruido de fondo se debe en gran medida a la señal acústica que detecta el micrófono distinta a la señal vocal de extremo próximo.

**3.8 ruido de confort:** Inserción de ruido pseudoaleatorio durante los periodos de silencio (señal vocal inactiva) mediante un SPDA.

**3.9 reducción de ruido (NR, *noise reduction*):** Dispositivo que reduce los efectos molestos y extenuantes del ruido de fondo. Es decir, una función NR reduce el nivel del ruido de fondo para mejorar la calidad general percibida de la señal transmitida.

**3.10 estimador de ruido:** Dispositivo que calcula una estimación de la característica del ruido de fondo que molesta. En los sistemas clásicos implementados en el dominio de frecuencias, se calcula la densidad espectral del ruido. En los sistemas que funcionan en el dominio de tiempo, el valor estimado es la autocorrelación del ruido. Generalmente, el estimador de ruido funciona durante los periodos sin voz (controlados por el detector de actividad vocal), pero existen otras posibilidades.

**3.11 filtraje de ruido:** Proceso que consiste en aplicar a la entrada (es decir, a la señal de ruido) el filtro calculado con un sistema de reducción de ruido. Este proceso se incluye en el dispositivo de reducción de ruido (NR). La aplicación del filtro se puede llevar a cabo en el dominio de tiempo a una tasa muestra por muestra (convolución) o a una tasa de bloques en el dominio de frecuencia (atenuación espectral de duración limitada).

**3.12 efecto Lombard:** Cuando los individuos se encuentran en un sitio con ruido de fondo de alto nivel, tienden a hablar más fuerte. Este comportamiento se denomina "efecto Lombard".

**3.13 detector de actividad vocal (VAD, *voice activity detector*):** Dispositivo que distingue entre periodos de silencio (señal vocal inactiva), periodos de monólogo (periodos vocales de extremo próximo o de extremo distante) y periodos de habla simultánea (actividad de señales vocales de extremo próximo y de extremo distante al mismo tiempo).

## 4 Procesamiento de eco acústico

Los componentes de un terminal que se han de utilizar para controlar el eco acústico incluyen los componentes analógicos (micrófono y altavoz), el compensador de eco acústico y los dispositivos no lineales.

### 4.1 Componentes analógicos

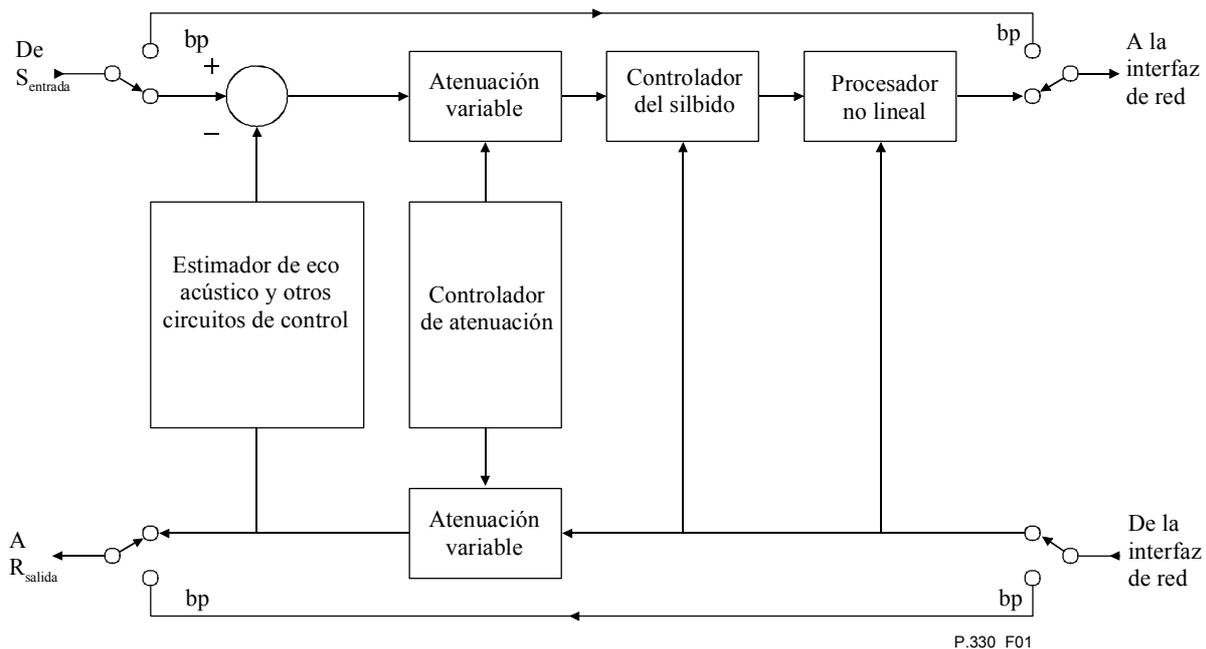
La posición y el tipo del micrófono(s) y el altavoz(es) determinan el nivel de eco acústico (acoplamiento acústico) antes de que se utilice cualquier procesamiento de señales. Sin ningún tipo de control digital, la atenuación de retorno del eco (ERL, *echo return loss*) para un sistema manos libres, por ejemplo, está entre 20 dB y -20 dB. Es decir, el nivel del eco acústico puede ser 20 dB mayor que la señal original recibida sin ningún dispositivo de control de eco digital. Se puede alterar la ERL cambiando la posición y el tipo de micrófono (cuando se utiliza un micrófono direccional) o altavoz(es).

## 4.2 Unidades funcionales de un compensador de eco acústico

Las unidades funcionales de un compensador de eco acústico son dispositivos o partes de dispositivos implementados en la unidad de procesamiento, que contribuyen a la función general de control de eco acústico. No hay ninguna restricción sobre cómo implementarlas.

Se pueden combinar las unidades funcionales para lograr una mejor calidad de funcionamiento. Éstas pueden utilizar todas las señales disponibles en el terminal (por ejemplo, las señales individuales procedentes de varios micrófonos dispuestos en un sistema acústico) y también se pueden combinar con otras funciones (por ejemplo, con la codificación vocal en subbanda) para lograr una implementación eficaz, siempre que no modifiquen las características de esas funciones cuando estén en funcionamiento.

En la figura 1 se ilustra un diagrama de bloques funcional característico de una unidad de procesamiento.



**Figura 1/P.330 – Diagrama de bloques funcional de una unidad de procesamiento típica (parte AEC) (bp indica los trayectos de puentado de señales para fines de prueba)**

## 4.3 Interacción entre el control de eco del terminal y el equipo de red de procesamiento de las señales

El control de eco acústico se efectúa más fácil y eficazmente en los terminales. Sin embargo, el equipo de red puede incluir también tratamiento de control de eco acústico. Esto conduce a condiciones de conexión en cascada de las funcionalidades de control de eco acústico. El componente que añade la red evitará cualquier degradación de la calidad general percibida.

El control de eco eléctrico se logra mediante el equipo de red descrito en la Rec. UIT-T G.168. Si la red no puede controlar adecuadamente el dispositivo de control de eco, se puede utilizar en el terminal para controlar parcialmente el eco eléctrico. Las interacciones entre ambos procesadores se describe en la Rec. UIT-T G.161. Los asuntos relativos a la planificación de la transmisión de los compensadores de eco se tratan en la Rec. UIT-T G.108.2.

## 4.4 Retardo

Los valores que se especifican más adelante corresponden al retardo adicional que puede resultar del tratamiento AEC. En cualquier caso, se debe alcanzar la conformidad con los objetivos de planificación de la transmisión.

En la Rec. UIT-T G.114 se da información general sobre los tiempos de transmisión. La Rec. UIT-T G.131 estipula las reglas para el control de eco en la red.

### 4.4.1 Retardo de procesamiento

El proceso de cancelación de eco requiere algún tiempo lo que provoca un retardo en el terminal, denominado "retardo de procesamiento".

### 4.4.2 Retardo de ida y vuelta en el trayecto de eco (EPDn, *round trip echo path delay*) – Interfaz de red

La audibilidad del eco depende del retardo de ida y vuelta en el trayecto de eco. Esto requiere el cálculo de la respuesta al impulso (trayecto de eco). Debería utilizarse ruido de banda ancha estacionario para esta medición. El terminal se puede colocar prácticamente en cualquier entorno reverberante o no, ya que el eco acústico proviene principalmente del acoplamiento directo. El ruido acústico debería cumplir con los requisitos definidos en 5.4/P.340. El AEC aprende inicialmente utilizando una señal en  $R_{\text{entrada}}$ .

## 4.5 Especificaciones de control de eco acústico

Los requisitos de la calidad de funcionamiento se describen en la cláusula 8/P.340 y los métodos de prueba en la Rec. UIT-T P.502.

### 4.5.1 Trayecto del eco acústico

Se recomienda emplear salas reales o recintos con características acústicas apropiadas. También pueden emplearse trayectos de eco simulados por dispositivos electrónicos, como reverberadores digitales con diagramas de reflexión que no varían en el tiempo, si el terminal tiene accesos internos en el lado de usuario. En este último caso, los ajustes del simulador electrónico deben corresponder a los valores recomendados para los recintos o salas reales; por otra parte, la envolvente de la respuesta impulsiva simulada debe tener una forma similar a la respuesta impulsiva del trayecto de eco real.

- Para los sistemas de teleconferencia, el tiempo de reverberación promediado en toda la banda de paso de transmisión será de 400 ms; el tiempo de reverberación en la octava más baja no será superior al doble de ese valor medio; el tiempo de reverberación en la octava más alta no será inferior a la mitad de ese valor. El volumen de una sala de prueba típica será del orden de 90 m<sup>3</sup>.
- Para los teléfonos manos libres y los videoteléfonos, el tiempo de reverberación promediado en toda la banda de paso de transmisión será de 500 ms; el tiempo de reverberación en la octava más baja no será superior al doble de ese valor medio; el tiempo de reverberación en la octava más alta no será inferior a la mitad de ese valor. El volumen de una sala de prueba típica será del orden de 50 m<sup>3</sup>.
- Para los radioteléfonos móviles puede emplearse un recinto que simule el interior de un vehículo; también puede utilizarse un vehículo real. El tiempo de reverberación medio típico es de 60 ms. El volumen del recinto será del orden de 2,5 m<sup>3</sup>.

NOTA – Se recomienda no utilizar salas demasiado largas (longitud  $\gg$  anchura, altura) ni salas con techo demasiado bajo (altura  $\ll$  longitud, anchura) y, si es posible, tampoco salas cuyas dimensiones laterales sean casi idénticas.

Deberán evitarse las grandes superficies lisas y paralelas limitadoras de la sala y las zonas cuya superficie provoque la reflexión del sonido de banda ancha, en particular las superficies de las

paredes a la altura media de la sala (aproximadamente, de 0,8 a 1,8 m por encima del suelo), ya que pueden provocar ecos múltiples y perturbaciones del tipo eco en reverberación (retorno de ecos, asperezas), si el montaje de prueba está en una posición desfavorable.

La medición de la distribución local dependiente de la frecuencia de los niveles de presión acústica dentro de una sala seleccionada en situación estable puede ayudar a determinar la posición óptima del montaje de prueba.

Como regla general se recomienda que la distancia mínima entre el montaje de prueba y las superficies limitadoras de la sala sea de 1 m, con independencia de las propiedades acústicas de esas superficies. Así se pueden evitar las perturbaciones debidas a las reflexiones iniciales y a la subida del nivel de presión acústica que pueden producirse localmente a bajas frecuencias. La misma recomendación es aplicable en caso de superficies de muebles de grandes dimensiones que reflejan el sonido.

#### **4.5.2 Parámetros y límites recomendados**

##### **4.5.2.1 Atenuación ponderada por acoplamiento del terminal – monólogo (TCL<sub>wst</sub>, *weighted terminal coupling loss – single-talk*)**

Atenuación ponderada entre las interfaces de red en  $R_{\text{entrada}}$  y  $S_{\text{salida}}$  durante el funcionamiento normal del AEC y mientras no llega ninguna señal del usuario local<sup>1</sup>.

Antes de cada prueba debe conectarse el terminal.

Los valores recomendados para cada tipo de terminal manos libres se encuentran en las Recomendaciones UIT-T correspondientes (por ejemplo, los del terminal manos libres de banda ancha en la Rec. UIT-T P.341 y los del terminal manos libres digital en la Rec. UIT-T P.342).

##### **4.5.2.2 Atenuación ponderada por acoplamiento del terminal – habla simultánea (TCL<sub>wdt</sub>, *weighted terminal coupling loss – double-talk*)**

Atenuación ponderada entre las interfaces de red en  $R_{\text{entrada}}$  y  $S_{\text{salida}}$  durante el funcionamiento normal del AEC y cuando el usuario normal y el usuario del extremo distante están activos simultáneamente<sup>1</sup>.

Los valores recomendados para cada tipo de terminal manos libres se encuentran en la cláusula 8/P.340.

##### **4.5.2.3 Atenuación de la voz recibida durante el habla simultánea (Ardt, *received speech attenuation during double-talk*)**

Atenuación (en el punto  $R_{\text{salida}}$ ) de la señal recibida, introducida por el AEC durante el habla simultánea.

La respuesta de frecuencia en el lado de recepción durante el habla simultánea debería ser idealmente la misma que durante las condiciones de monólogo. Sin embargo, en la práctica, no siempre será posible implementar una compensación del eco que proporcione suficiente atenuación de eco durante el habla simultánea, sin modificar la respuesta de frecuencia.

##### **4.5.2.4 Atenuación de la voz transmitida durante el habla simultánea (Asdt, *sent speech attenuation during double-talk*)**

Atenuación (en el punto  $S_{\text{salida}}$ ) de la señal enviada, introducida por el AEC durante el habla simultánea.

---

<sup>1</sup> La ponderación se realiza de acuerdo con la regla especificada en la Rec. UIT-T G.122 (Cálculo del índice de sonoridad del eco para el hablante). Hay que tener precaución para evitar el posible enmascaramiento de los efectos de las oscilaciones parásitas por la ponderación (en estudio).

La respuesta de frecuencia en el lado de transmisión durante el habla simultánea debería ser idealmente la misma que durante las condiciones de monólogo. Sin embargo, en la práctica no siempre será posible implementar una compensación del eco que proporcione suficiente atenuación de eco durante el habla simultánea, sin modificar la respuesta de frecuencia.

#### **4.5.2.5 Distorsión de la voz recibida durante el habla simultánea ( $D_{rdt}$ , *received speech distortion during double-talk*)**

Distorsión no lineal total de la señal en el punto  $R_{salida}$  que puede producir el AEC durante los periodos de habla simultánea.

La distorsión adicional en  $R_{salida}$ , en comparación con la situación de monólogo, debe ser pequeña en todas las aplicaciones.

#### **4.5.2.6 Distorsión de la voz transmitida durante el habla simultánea ( $D_{sdt}$ , *sent speech distortion during double-talk*)**

Distorsión total no lineal de la señal en el punto  $S_{salida}$  que puede producir el AEC durante los periodos de habla simultánea.

La distorsión adicional en  $S_{salida}$ , en comparación con la situación de monólogo debe ser pequeña en todas las aplicaciones.

#### **4.5.2.7 Tiempo de establecimiento – monólogo ( $TR_{st}$ , *build-up time – single-talk*)**

Intervalo de tiempo entre el comienzo de la señal recibida (análogamente para la señal transmitida) y el momento en que la atenuación en el trayecto de recepción (análogamente para el trayecto de emisión) alcanza el valor de [3] dB. A estos efectos, el extremo opuesto permanece en silencio.

##### **4.5.2.7.1 Lado receptor ( $TR_{st-r}$ , *receive side*)**

En todas las aplicaciones, el valor de  $TR_{st-r}$  no deberá superar [20 ms].

##### **4.5.2.7.2 Lado emisor ( $TR_{st-s}$ , *send side*)**

En todas las aplicaciones, el valor de  $TR_{st-s}$  no deberá superar [20 ms].

#### **4.5.2.8 Tiempo de establecimiento – habla simultánea ( $TR_{dt}$ , *build-up time – double-talk*)**

Intervalo de tiempo entre el comienzo de la señal recibida (análogamente para la señal transmitida) y el momento en que la atenuación en el trayecto de recepción (análogamente para el trayecto de emisión) alcanza el valor  $A_{rdt}$  (análogamente para  $A_{sdt}$ ). A estos efectos, la señal en sentido contrario al de la transmisión se mantiene en el nivel especificado.

##### **4.5.2.8.1 Lado receptor ( $TR_{dt-r}$ , *receive side*)**

Si la atenuación es superior a 6 dB,  $TR_{dt-r}$  debe ser inferior a [20 ms].

##### **4.5.2.8.2 Lado emisor ( $TR_{dt-s}$ , *send side*)**

Si la atenuación es superior a 6 dB,  $TR_{dt-s}$  debe ser inferior a [20 ms].

#### **4.5.2.9 Tiempo de convergencia ( $T_c$ , *convergence time*)**

El tiempo de convergencia es el comprendido entre el momento de aplicación de una señal de prueba determinada al puerto  $R_{entrada}$  del terminal (tras haber desactivado y posteriormente activado todas las funciones del AEC) y el momento en que la señal de eco recibida en el puerto  $S_{salida}$  queda atenuada en un valor predefinido como mínimo. El usuario local no está activo.

#### 4.5.2.10 Tiempo de bloqueo tras el habla simultánea (THdt, *hang-over time after double-talk*)

Tiempo transcurrido entre el final de un periodo de habla simultánea y el momento en que la atenuación del eco alcanza de nuevo un valor determinado (se recibe una señal continua procedente del usuario remoto).

En todas las aplicaciones, la atenuación de la señal en  $S_{\text{salida}}$  debe ser como mínimo de [20 dB] tras  $\text{THdt} = [1]$  segundo.

#### 4.5.2.11 Atenuación de acoplamiento del terminal ponderada temporalmente – monólogo (TCLtst, *terminal coupling loss temporally weighted – single-talk*)

La atenuación de retorno del eco entre  $R_{\text{entrada}}$  y  $S_{\text{salida}}$  se mide de conformidad con el procedimiento definido para la ERLtst en la Rec. UIT-T P.502.

#### 4.5.2.12 Atenuación de retorno de eco de acoplamiento del terminal ponderada temporalmente – habla simultánea (ERLtdt, *terminal coupling loss temporally weighted echo return loss – double-talk*)

La atenuación de retorno del eco entre  $R_{\text{entrada}}$  y  $S_{\text{salida}}$  se mide de conformidad con el procedimiento definido para la ERLtdt en la Rec. UIT-T P.502.

## 5 Reducción de ruido

La finalidad principal de un sistema de reducción de ruido (NR, *noise reduction*) en un dispositivo es reducir los efectos molestos y extenuantes del ruido de fondo transmitido. Las técnicas utilizadas para reducir el ruido de fondo pueden clasificarse como: sólo analógica, sólo digital y analógica y digital combinadas.

### 5.1 Componentes analógicos

Los componentes analógicos de un sistema NR incluyen el micrófono y cualquier dispositivo analógico que conecta el micrófono al CÓDEC (convertidor analógico a digital). Existen varias técnicas que se utilizan para reducir el ruido de fondo basadas únicamente en componentes analógicos:

- a) La proximidad del micrófono a la boca del hablante es un factor primordial para determinar la relación señal/ruido (SNR, *signal to noise ratio*). Colocando el micrófono más cerca de la boca del hablante se produce una mejora de la SNR (SNRE, *SNR enhancement*) obvia pero significativa.
- b) El trayecto de la señal analógica se diseña normalmente de manera que tenga una respuesta de filtro pasa alto. Cuando el ruido tiene componentes de baja frecuencia intensos (ejemplo, ruido de automóvil), esta técnica de filtrado permite mejorar la SNR (si se mide en toda la banda). El efecto conexo es una atenuación considerable del timbre de la calidad vocal (especialmente en las voces masculinas).
- c) Los micrófonos se pueden diseñar de manera que aporten una ganancia direccional pasiva. El tipo que se utiliza con mayor frecuencia en los automóviles es un micrófono diferencial de primer orden. Este micrófono se puede diseñar con un solo transductor que emplea dos puertos. En el caso de un campo con ruido difuso y una orientación correcta del micrófono, un sistema de micrófonos diferenciales de primer orden hipercardiodes mejorará la SNR en 6 dB en comparación con un micrófono omnidireccional. Es posible la utilización de micrófonos diferenciales de orden superior. Además, también es posible utilizar sistemas de micrófonos que utilizan únicamente técnicas pasivas, pero es poco probable que se utilicen con frecuencia debido a que para que tengan una repercusión en los componentes de bajas frecuencias de la voz estos sistemas deben ser muy grandes. Estos sistemas pueden constar

hasta de 16 elementos que proporcionan una ganancia direccional de aproximadamente 20 dB en algunas frecuencias.

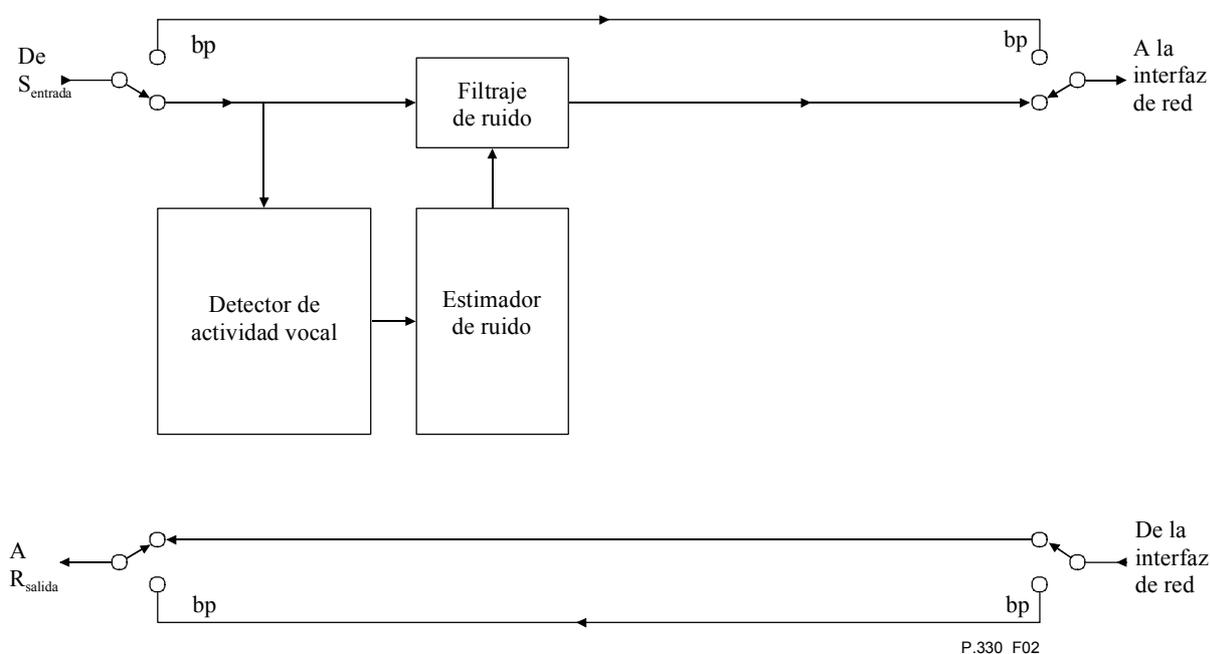
## 5.2 Unidades funcionales de un sistema de reducción de ruido

Las unidades funcionales de un sistema de reducción de ruido son dispositivos o partes de dispositivos implementados en la unidad de procesamiento, que contribuyen a la función general de reducción de ruido. No hay restricciones para su implementación.

En la figura 2 se muestra un diagrama funcional de bloques de una unidad de procesamiento típica.

Se pueden implementar dos tipos comunes de técnicas digitales de reducción de ruido en un procesador de señales digitales (DSP, *digital signal processor*) o en otro tipo de microprocesador en la terminal que utiliza un solo micrófono. Las técnicas más comunes son:

- Supresión de ruido en toda la banda: durante las pausas en la conversación, el ruido se reduce significativamente siempre que su energía se encuentre por debajo de un nivel de umbral. Durante la conversación se suprime la atenuación permitiendo que pase tanto la voz como el ruido. Esto produce un efecto de estimulación de ruido indeseable si la atenuación se fija muy alta.
- Supresión de ruido en subbanda: la señal transmitida se divide en subbandas utilizando un algoritmo de transformada rápida de Fourier (FFT, *fast Fourier transform*). En este caso, sólo se atenúan las bandas de frecuencias con ruido estacionario mientras que las bandas con señales vocales no se ven afectadas. El método ya conocido de sustracción espectral es una de esas técnicas. En la práctica, los niveles de supresión de ruido abarcan de 6 a 15 dB. El inconveniente de estas técnicas es la existencia de un compromiso entre el nivel de reducción de ruido y la distorsión de la señal vocal original. Por lo tanto, es difícil encontrar un punto de equilibrio que funcione bajo cualquier condición de ruido (SNR y tipo de ruido). Ahora bien, en condiciones de SNR reducida, la señal vocal se degrada en cierta medida si se utilizan altos niveles de supresión de ruido.



**Figura 2/P.330 – Diagrama funcional de bloques de una unidad de procesamiento típica (parte NR) (bp representa el puentado de los trayectos de señales para fines de pruebas)**

### **5.3 Interacción entre el sistema de reducción de ruido en el terminal y el equipo de procesamiento de señales en la red**

El equipo de red también puede incluir procesamiento de reducción de ruido. Esto conduce a condiciones de conexión en cascada de las funcionalidades de reducción de ruido. El componente que añade la red evitará cualquier degradación de la calidad general percibida.

### **5.4 Tiempo de procesamiento de reducción de ruido**

El tiempo de procesamiento de reducción de ruido depende en gran medida de la técnica utilizada por los filtros de ruido. Es decir, se debe alcanzar la conformidad a los objetivos de planificación de transmisión.

En la Rec. UIT-T G.114 se describe la información general relativa a los tiempos de transmisión.

### **5.5 Especificaciones del sistema de reducción de ruido**

#### **5.5.1 Entorno de ruido**

Las características acústicas del entorno de prueba se describen en 4.5.1. La repercusión en el entorno se debería tener en cuenta en el caso de captación de sonido distante, ya que se tiene que considerar la reverberación del ruido de fondo como una degradación adicional.

La difusión de señales de prueba de ruido de fondo se describe en 7.10/P.340.

Las señales de prueba de ruido de fondo deberían incluir señales reales tales como murmullo, ruido de oficinas, ruido de la calle, ruido de automóviles (motor, condiciones de manejo a diferentes velocidades) y otras señales de ruido de fondo simuladas (en función de la utilización del equipo).

Se deberían variar los niveles de las señales de prueba de ruido de fondo de modo que se obtengan SNR en la gama  $[-3 \text{ dB}, 30 \text{ dB}]$ .

NOTA – Algunas de las señales de pruebas correspondientes se describen en la Rec. UIT-T P.501 y los métodos de prueba en la Rec. UIT-T P.502. Las señales de pruebas adicionales se encuentran en estudio.

#### **5.5.2 Parámetros y límites recomendados**

Todos los parámetros especificados se deberían medir con distintos valores de SNR en la gama  $[-3 \text{ dB}, 30 \text{ dB}]$ .

Para los parámetros que consisten en la medición de una atenuación o un retardo del nivel de señal, los procedimientos de medición (métodos y estímulos) se describen en las Recomendaciones UIT-T P.501, P.502 y P.340, que se aplican con las siguientes restricciones:

- el nivel de la señal vocal debe ser al menos 10 dB mayor que el nivel de ruido,
- el ruido debe ser estacionario.

Para el resto de los casos (ruido no estacionario, valores reducidos de SNR, medición de distorsión), los métodos de prueba se encuentran en estudio.

Todos los parámetros definidos más adelante corresponden a condiciones de monólogo. Debido a las posibles interacciones entre el AEC y el procesamiento de NR integrado en el terminal, se deben considerar también los parámetros bajo condiciones de habla simultánea (en estudio).

##### **5.5.2.1 Atenuación de la voz transmitida en condiciones de silencio (*Asqc, sent speech attenuation in quiet conditions*)**

Atenuación (en el punto  $S_{\text{salida}}$ ) de la señal transmitida introducida por la NR en condiciones de silencio.

### 5.5.2.2 Distorsión de la voz transmitida en condiciones de silencio ( $D_{sq}$ , *sent speech distortion in quiet conditions*)

Distorsión no lineal total de la señal en el punto  $S_{salida}$ , producida por la NR en condiciones de silencio.

La distorsión adicional en  $S_{salida}$ , en comparación con  $S_{entrada}$ , debería ser lo más pequeña posible en todas las aplicaciones. Idealmente, la NR no debería introducir distorsión adicional.

### 5.5.2.3 Atenuación de la voz transmitida durante condiciones de ruido ( $A_{sn}$ , *sent speech attenuation during noisy conditions*)

Atenuación (en el punto  $S_{salida}$ ) de la señal transmitida introducida por la NR durante condiciones de ruido.

Idealmente, la respuesta de frecuencia en el lado de transmisión no debería modificarse cuando se activa la NR.

### 5.5.2.4 Distorsión de la voz transmitida durante condiciones de ruido ( $D_{sn}$ , *sent speech distortion during noisy conditions*)

Distorsión no lineal total de la señal en el punto  $S_{salida}$  producida por la NR durante los eventos de ruido.

La distorsión adicional en  $S_{salida}$ , en comparación con  $S_{entrada}$ , debería ser lo más pequeña posible en todas las aplicaciones. Idealmente, la NR no debería introducir distorsión adicional.

### 5.5.2.5 Tiempo de adaptación (TA, *adaptation time*)

El tiempo de adaptación es el intervalo de tiempo entre el instante en que se aplica una señal de prueba de ruido especificada en el puerto  $S_{entrada}$  del terminal (después de que se han reiniciado y rehabilitado todas las funciones de la NR), y el instante cuando la señal de prueba de ruido devuelta al puerto  $S_{salida}$  se estabiliza dentro de  $\pm 1$  dB, en comparación con el nivel de ruido reducido de largo plazo (véase la figura 3). Los usuarios local y distante no están activos.

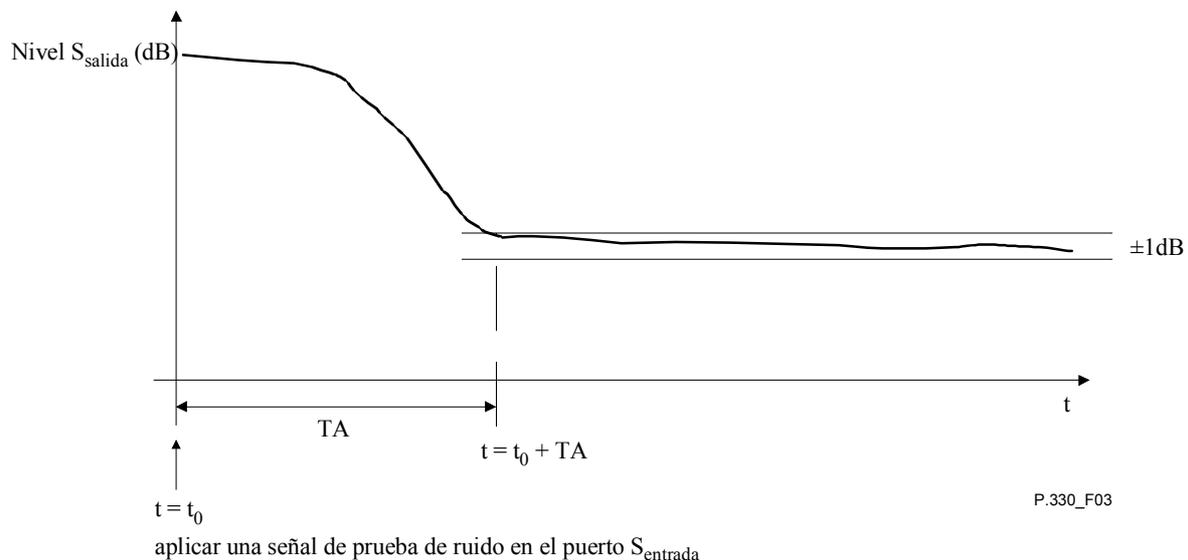


Figura 3/P.330 – Definición del tiempo de adaptación (TA)

### 5.5.2.6 Tiempo de adaptación después de una conversación ( $T_{Ase}$ , *adaptation time after speech event*)

Tiempo transcurrido entre el final de una conversación y el instante en que la atenuación del ruido recupera un valor especificado.

Para todas las aplicaciones, con altos niveles de ruido de fondo [ $-3 \text{ dB} < \text{SNR} < 15 \text{ dB}$ ], la atenuación de la señal de ruido en  $S_{\text{salida}}$  debería ser al menos [6 dB] después de  $T_{\text{Ase}} = [100]$  milisegundos.

#### **5.5.2.7 Atenuación de ruido en el terminal – sin conversación (TNAtns, *terminal noise attenuation – no speech*)**

Atenuación de ruido en el terminal entre  $S_{\text{entrada}}$  y  $S_{\text{salida}}$  introducida por la NR en la señal de ruido de fondo cuando no hay señal de conversación.

#### **5.5.2.8 Distorsión de ruido – sin conversación (Dnns, *noise distortion – no speech*)**

Distorsión no lineal total de la señal en el punto  $S_{\text{salida}}$  producida por la NR en la señal de ruido de fondo cuando no hay señal de conversación.

La distorsión en  $S_{\text{salida}}$ , en comparación con  $S_{\text{entrada}}$ , debería ser despreciable en todas las aplicaciones.

#### **5.5.2.9 Atenuación de ruido en el terminal – durante la conversación (TNAtps, *terminal noise attenuation – in the presence of speech*)**

Atenuación de ruido en el terminal entre  $S_{\text{entrada}}$  y  $S_{\text{salida}}$  introducida por la NR en la señal de ruido de fondo durante la conversación (medición de la mejora de la relación señal a ruido).

#### **5.5.2.10 Nivel de ruido de confort y concordancia de espectro – sin conversación (CNLMns y CNSMns, *comfort noise level and spectrum matching – no speech*)**

Ruido de confort en el punto  $S_{\text{salida}}$  introducido por la NR cuando no hay conversación que debería concordar en nivel y en espectro con el ruido de fondo presente en el punto  $S_{\text{entrada}}$ .

#### **5.5.2.11 Nivel de ruido de confort y concordancia de espectro – durante la conversación (CNLMps y CNSMps, *comfort noise level and spectrum matching – in the presence of speech*)**

Ruido de confort en el punto  $S_{\text{salida}}$  introducido por la NR durante la conversación que debería concordar en nivel y espectro con el ruido de fondo presente en el punto  $S_{\text{entrada}}$ .



## SERIES DE RECOMENDACIONES DEL UIT-T

Serie A	Organización del trabajo del UIT-T
Serie B	Medios de expresión: definiciones, símbolos, clasificación
Serie C	Estadísticas generales de telecomunicaciones
Serie D	Principios generales de tarificación
Serie E	Explotación general de la red, servicio telefónico, explotación del servicio y factores humanos
Serie F	Servicios de telecomunicación no telefónicos
Serie G	Sistemas y medios de transmisión, sistemas y redes digitales
Serie H	Sistemas audiovisuales y multimedia
Serie I	Red digital de servicios integrados
Serie J	Redes de cable y transmisión de programas radiofónicos y televisivos, y de otras señales multimedia
Serie K	Protección contra las interferencias
Serie L	Construcción, instalación y protección de los cables y otros elementos de planta exterior
Serie M	RGT y mantenimiento de redes: sistemas de transmisión, circuitos telefónicos, telegrafía, facsímil y circuitos arrendados internacionales
Serie N	Mantenimiento: circuitos internacionales para transmisiones radiofónicas y de televisión
Serie O	Especificaciones de los aparatos de medida
<b>Serie P</b>	<b>Calidad de transmisión telefónica, instalaciones telefónicas y redes locales</b>
Serie Q	Conmutación y señalización
Serie R	Transmisión telegráfica
Serie S	Equipos terminales para servicios de telegrafía
Serie T	Terminales para servicios de telemática
Serie U	Conmutación telegráfica
Serie V	Comunicación de datos por la red telefónica
Serie X	Redes de datos y comunicación entre sistemas abiertos
Serie Y	Infraestructura mundial de la información y aspectos del protocolo Internet
Serie Z	Lenguajes y aspectos generales de soporte lógico para sistemas de telecomunicación