DOCUMENTO TÉCNICO

Rango de captación del micrófono: factores que influyen en el rendimiento del audio

Septiembre 2026

AXIS COMMUNICATIONS

Índice

1	Introducción	3
2	Rango de grabación física	3
3	Grabaciones en distintos entornos	5
4	Resultados de las grabaciones obtenidas en distintos entornos	7
1	Apéndice	9

1 Introducción

Rara vez resulta sencillo establecer un alcance máximo definitivo para los dispositivos de grabación de audio, ya que su rendimiento depende del entorno acústico. Aunque las especificaciones técnicas actúan como punto de referencia, el rango de la grabación real depende de una compleja interacción de factores ambientales y acústicos que varían de una instalación a otra.

Este informe técnico analiza los principales factores que determinan la distancia a la que un micrófono puede captar el sonido de forma fiable, con un enfoque sobre la inteligibilidad del habla como referencia principal. Entre estos factores se incluyen el nivel de ruido de fondo, las características espectrales de dicho ruido, el tiempo de reverberación, la directividad del micrófono y el ruido propio tanto del micrófono como del dispositivo de grabación.

Con el fin de establecer un punto de referencia práctico, empleamos la voz humana como fuente sonora de referencia. El ruido se mide con un filtro de ponderación A para lograr una mejor correlación entre el valor medido y el nivel de ruido percibido. La relación señal-ruido (SNR) de la grabación de voz se calcula como la diferencia entre el nivel de la voz y el nivel de ruido ponderado A, expresada en decibelios *(véase el anexo para obtener más información sobre la SNR)*. Las pruebas empíricas demuestran que la inteligibilidad del habla se percibe hasta una relación señal-ruido (SNR) de aproximadamente 0 dB (es decir, cuando el ruido y el sonido tienen la misma intensidad), siempre que el ruido tenga una respuesta en frecuencia relativamente plana dentro del rango de frecuencias del habla.

Al comprender estas variables, tanto los instaladores como los diseñadores de sistemas pueden adoptar medidas mejor fundamentadas a la hora de seleccionar y disponer dispositivos de audio en situaciones reales.

2 Rango de grabación física

La capacidad de grabación de un micrófono depende, por lo general, del ruido de fondo del entorno. Los principales factores que limitan el alcance de una grabación física son:

- **Nivel de presión del sonido:** no todos los sonidos recorren la misma distancia. Por ejemplo, un disparo genera un nivel de presión del sonido extremadamente alto que puede oírse a distancias de hasta un kilómetro. Sin embargo, la voz humana es más suave y se atenúa más rápidamente a medida que aumenta la distancia. Para calcular un alcance de captación realista, primero debe definirse el tipo de sonido que se necesita capturar, de modo que el nivel de la fuente establezca el límite máximo de lo que es posible realizar en la grabación.
- **Ruido de fondo:** cuanto más ruidoso es el entorno, más difícil resulta para un micrófono aislar y captar el sonido deseado. El ruido de fondo varía de un entorno a otro y compite directamente con el sonido que se desea escuchar. Esto incluye el ruido del viento, del tráfico, de la maquinaria y de la multitud, entre otros. Cuanto mayor sea el ruido de fondo en la localización del micrófono, menor será el alcance efectivo de captación. Esto explica por qué un mismo micrófono puede funcionar de forma muy diferente en un pasillo silencioso en comparación con un cruce muy transitado.
- **Perfil espectral del ruido:** al igual que el volumen del ruido es importante, también lo son las frecuencias que ocupa. El habla humana se sitúa principalmente en el rango de frecuencias medias, aproximadamente entre 300 Hz y 3 400 Hz. Si el ruido de fondo se concentra en frecuencias muy bajas o muy altas, es menos probable que interfiera en la inteligibilidad del habla. No obstante, si el ruido se solapa con el mismo rango de frecuencias que el sonido de interés, resultará difícil distinguir el habla del ruido, incluso con niveles de ruido moderados.
- **Tiempo de reverberación en la localización del micrófono:** la reverberación es la persistencia del sonido una vez que la fuente original ha dejado de sonar, provocada por su reflexión en paredes, techos, suelos y otras superficies. Un tiempo de reverberación prolongado puede reducir la inteligibilidad del habla, incluso en situaciones en las que el nivel sonoro general y la relación señal-ruido parezcan aceptables. En espacios con mucha reverberación, como grandes salas vacías o estancias con suelo de baldosas, el sonido de la voz llega al micrófono acompañado de un sinfín de reflexiones de la sala, provocando que el sonido resulte confuso. Esto conlleva que tanto a los oyentes como a los sistemas de grabación les resulte difícil distinguir claramente las palabras por separado.

- **Directividad de la configuración de los micrófonos:** los micrófonos varían en cuanto a su sensibilidad a los sonidos procedentes de distintas direcciones. Un micrófono omnidireccional capta el sonido por igual desde todas las direcciones, lo que también significa que capta el ruido de fondo y la reverberación procedentes de todas las direcciones. Una configuración direccional, como varios micrófonos configurados para la formación de haces, tiene un enfoque en una dirección concreta, lo que reduce de forma eficaz la influencia del ruido y la reverberación procedentes de otros ángulos. Esto puede ampliar considerablemente el alcance de captación útil en entornos ruidosos o con reverberación, pero solo en la dirección hacia la que está orientado el conjunto de micrófonos.
- **Ruido propio de los micrófonos y los dispositivos de grabación:** todos los micrófonos y dispositivos de grabación de audio generan una pequeña cantidad de ruido eléctrico interno, conocido como ruido propio. Esto establece un límite mínimo absoluto de lo que el sistema puede capturar. Por muy silencioso que sea el entorno, los sonidos que están por debajo del umbral de ruido propio simplemente no se pueden grabar con claridad. A la hora de comparar micrófonos, un nivel de ruido propio más bajo suele indicar que el micrófono es capaz de captar mejor los sonidos más bajos, lo que puede traducirse en un alcance de captación mayor si el nivel de ruido y el tiempo de reverberación son suficientemente bajos.

Para simplificar el análisis, podemos partir de un escenario idealizado en el que no haya ruido de fondo, ni reflexiones, ni reverberación, y utilizar la voz humana como fuente de sonido de referencia. En estas circunstancias, la pregunta clave es: ¿Qué relación señal-ruido (SNR) deberíamos utilizar como umbral para un rango de captación aceptable?

Las pruebas empíricas ofrecen una respuesta clara. Cuando aplicamos un filtro de ponderación A al ruido del sistema y comparamos el nivel medio cuadrático (RMS) del ruido con el nivel medio RMS del habla, la inteligibilidad del habla solo empieza a disminuir cuando la relación señal-ruido (SNR) desciende por debajo de 0 dB. Por lo tanto, utilizamos una relación señal-ruido (SNR) de 0 dB como criterio límite para la grabación de la voz humana. Esto es así siempre y cuando el ruido tenga una respuesta en frecuencia relativamente plana, similar a la del ruido blanco, y se encuentre dentro del rango de frecuencias del habla; esto suele ser así en el caso del ruido propio del micrófono.

Por ejemplo, un micrófono digital puede tener una sensibilidad de -36 dBFSrms *(véase el anexo para obtener información sobre dBFSrms)* cuando se reproduce un tono de 1 kHz a 94 dBspl (una señal sinusoidal digitalizada a escala completa se define aquí como 0 dBFSrms, según las fichas de datos de los micrófonos) *(véase el anexo para obtener información sobre dBspl)*. Esto se aplica a muchos productos de Axis que incorporan micrófonos MEMS digitales cuando la ganancia del micrófono está ajustada a 0 dB. Cuando se coloca este micrófono en un entorno muy silencioso, donde el nivel de ruido de fondo es inferior a 10 dBspl(A) *(véase el anexo para obtener información sobre los dBspl(A))*, el ruido interno del micrófono pasa a ser predominante. En este ejemplo, la ficha de datos del micrófono especifica una salida digital de -105 dBFSrms(A). La relación señal-ruido (SNR) del micrófono es la diferencia entre el nivel de referencia y el ruido de fondo: –36 – (–105) = 69 dB.

Convertimos este nivel de ruido eléctrico de fondo en un nivel de ruido acústico equivalente:
94 – 69 = 25 dBspl(A).

Seguidamente, calculamos la distancia máxima a la que podemos realizar la grabación de una voz humana con una relación señal-ruido (SNR) de 0 dB. Según las normas de la UIT, la voz humana tiene un nivel medio nominal de voz (ASL) de aproximadamente 57 dBspl a 1 m. En espacios libres, sin objetos que reflejen el sonido, el nivel de presión del sonido disminuye con la distancia según la ley de la inversa del cuadrado:

$$dB = 20\log\frac{reference\ distance}{test\ distance}$$

.

Si la distancia de referencia es de 1 m, para alcanzar una relación señal-ruido (SNR) de 0 dB en este micrófono, el nivel del habla debe reducirse a 25 dBspl(A); esto supone una reducción de 32 dB con respecto al nivel de referencia de 1 m. Al reorganizar la fórmula, obtenemos:

$$test\ distance = \frac{reference\ distance}{10^{\frac{-(57-noise\ level)}{20}}} = \frac{1}{10^{\frac{-32}{20}}} = 39.8\text{m}$$

.

Esto significa que, en espacios libres, sin ruido de fondo, sin viento y sin reverberación, este micrófono puede, en teoría, grabar voz inteligible a una distancia aproximada de 40 m. Este es el alcance máximo teórico de captación.

Sin embargo, los factores ambientales tienen un mayor impacto en el nivel de ruido que el ruido de fondo del micrófono interno. Una oficina tranquila fuera del horario laboral, con los sistemas de ventilación aún en funcionamiento, presenta un nivel de ruido de fondo aproximado de entre 28 y 32 dBspl. Esto por sí solo puede reducir el alcance efectivo de captación en aproximadamente un 50 %, lo que reduce el alcance máximo a unos 20 m. Si realizamos una grabación al aire libre en un entorno urbano, se producirá una combinación de ruidos: el ruido del viento, el ruido del tráfico, los ruidos provocados por la actividad humana, etcétera. Cada nivel y cada espectro de ruido tendrán su propio rango de captación correspondiente cuando se utilice un micrófono omnidireccional. En este caso, el alcance de captación puede ser de entre 2 y 4 m como máximo si el ruido del tráfico es intenso.

Una consideración práctica es que las grabaciones de voz humana realizadas en el umbral de 0 dBSNR o cerca de este valor requieren una ganancia digital significativa para que sean audibles, ya que el nivel de la voz grabada es tan bajo que, de no ser así, quedaría por debajo de nuestro propio umbral auditivo durante la reproducción. En los ejemplos de grabación recogidos a continuación, los archivos de audio se han amplificado en +40 dB para compensar el nivel de audio de la voz lejana.

3 Grabaciones en distintos entornos

Esta sección detalla las localizaciones de grabación, acompañadas de capturas de pantalla de las grabaciones de vídeo y muestras de voz captadas a distintas distancias.

Ejemplo 1: La localización es al aire libre, con ruido del tráfico procedente de una autopista. El nivel de ruido de fondo oscila entre 62 y 68 dBspl[A] y también se percibe algo de ruido del viento.



Figura 3.1 *La localización de la grabación está próxima a una autopista (véase el punto rojo).*



Figura 3.2 *En la localización de grabación, la distancia máxima de grabación es de 4 m respecto a una cámara corporal de Axis.*

Preferiblemente, utilice auriculares para escuchar la grabación.

Ejemplo 2: Un pasillo silencioso con un nivel de ruido de fondo de 32 dBspl[A]. En este caso, el pasillo actúa como una larga guía para las ondas acústicas. El sonido se propaga por el pasillo rebotando en las superficies reflectantes y llega a los micrófonos con más intensidad de la que tendría en un espacio libre y a la misma distancia.



Figura 3.3 *La distancia máxima de grabación es de 27 m.*

Preferiblemente, utilice auriculares para escuchar la grabación.

Ejemplo 3: Un vestíbulo amplio con un nivel de ruido de fondo de 42 dBspl[A]. Aquí, las paredes cuentan con materiales de absorción acústica situados detrás de los paneles difusores, logrando que desaparezcan la mayoría de las reflexiones tempranas. Desde una distancia mayor, el sonido se propaga por el canal acústico existente entre el suelo y el techo reflectante.



Figura 3.4 *A una distancia máxima de grabación de 40 m, apenas se ve al orador, pero sí se le oye.*

Preferiblemente, utilice auriculares para escuchar la grabación.

4 Resultados de las grabaciones obtenidas en distintos entornos

Este análisis tiene un enfoque sobre la captación de voz. Los sonidos más fuertes, como los disparos, se oirán desde una distancia mayor. Analizamos y medimos cada frase de las grabaciones de audio, y calculamos la relación señal-ruido (SNR) de cada frase grabada.

La relación señal-ruido(SNR) se calculó como: + ruido - ruido de fondo. Una forma más correcta consistiría en calcular el nivel de voz - y el nivel de ruido, pero no es posible medir el nivel de voz sin el ruido. Esto significa que, cuando el nivel del habla se aproxima al nivel de ruido, la precisión de la relación señal-ruido (SNR) disminuye.

Localización y ruido de fondo:	En el exterior, con ruido del tráfico: 61-68 dBspl[A]	En un pasillo silencioso, 32 dBspl[A]	En un vestíbulo amplio: 42 dBspl[A]	Espacio libre, sin reflexiones, silencio absoluto
Distancia (m)	**SNR [dB] (estimado)**	**SNR [dB] (estimado)**	**SNR (estimado en [dB])**	**SNR teórico = ruido del micrófono a nivel de voz [dB]**
1	2,4	29,4	21,5	32,0

2	1,0	25,6	17,2	26,0
4	0,5	23,4	14,2	20,0
8	0,4	18,9	11,2	13,9
16	No audible	17,5	6,3	7,9
24	No audible	14,8	5,4	4,4
32	No audible	El pasillo es demasiado corto	4,6	1,9
40	No audible	El pasillo es demasiado corto	4,7	0,0

En teoría, el nivel del habla debería reducirse en -6 dB al duplicarse la distancia. Esto también significaría que la relación señal-ruido (SNR) debería disminuir en consecuencia. En esta tabla podemos ver que la relación señal-ruido (SNR) es mejor de lo esperado en el pasillo y en el vestíbulo. El motivo es que las reflexiones en la sala aumentaron los niveles de presión del sonido. Especialmente en el pasillo, las paredes, el suelo y el techo actúan como una guía de ondas para las ondas acústicas. Esto provoca que el nivel de señal y la relación señal-ruido sean unos 10 dB mejores que la relación señal-ruido teórica en espacios libres, cuando se realiza la prueba a una distancia de 24 m, en un pasillo.

Siempre que el ruido de fondo supere el ruido propio acústico equivalente de un único micrófono omnidireccional, el rango de captación dependerá exclusivamente del ruido de fondo ambiente y no del micrófono.

Apéndice 1 Apéndice

- **Nivel de presión del sonido en decibelios (dBspl):** el nivel de presión del sonido expresado en decibelios (dB) con respecto a 20 µPa. A la hora de medir los niveles de ruido y de voz, suelen emplearse los niveles de RMS integrados en el tiempo. A veces se le denomina "nivel de potencia".

$$RMS = \sqrt{\frac{1}{T2 - T1}\int_{T_1}^{T2}[f(t)]^2\,dt}$$

  Los dos tiempos de integración más habituales son el "RMS rápido", con una ventana temporal de 125 ms, y el "RMS lento", con una ventana temporal de 1 segundo. Para el análisis del ruido de fondo, el nivel suele promediarse a lo largo de un periodo de tiempo más prolongado.

- **Nivel de presión del sonido ponderado mediante el filtro de ponderación A (dBspl[A]):** a veces se expresa como dB(A). Es el valor ponderado A del nivel en dBspl. La ponderación A es un modo de filtrar la señal del micrófono para mejorar la correspondencia entre la audición humana y el nivel de presión del sonido medido. Esto se debe a que el oído humano tiene una sensibilidad reducida a las frecuencias bajas, por debajo de los 500 Hz, y a las frecuencias altas, por encima de los 10 kHz.

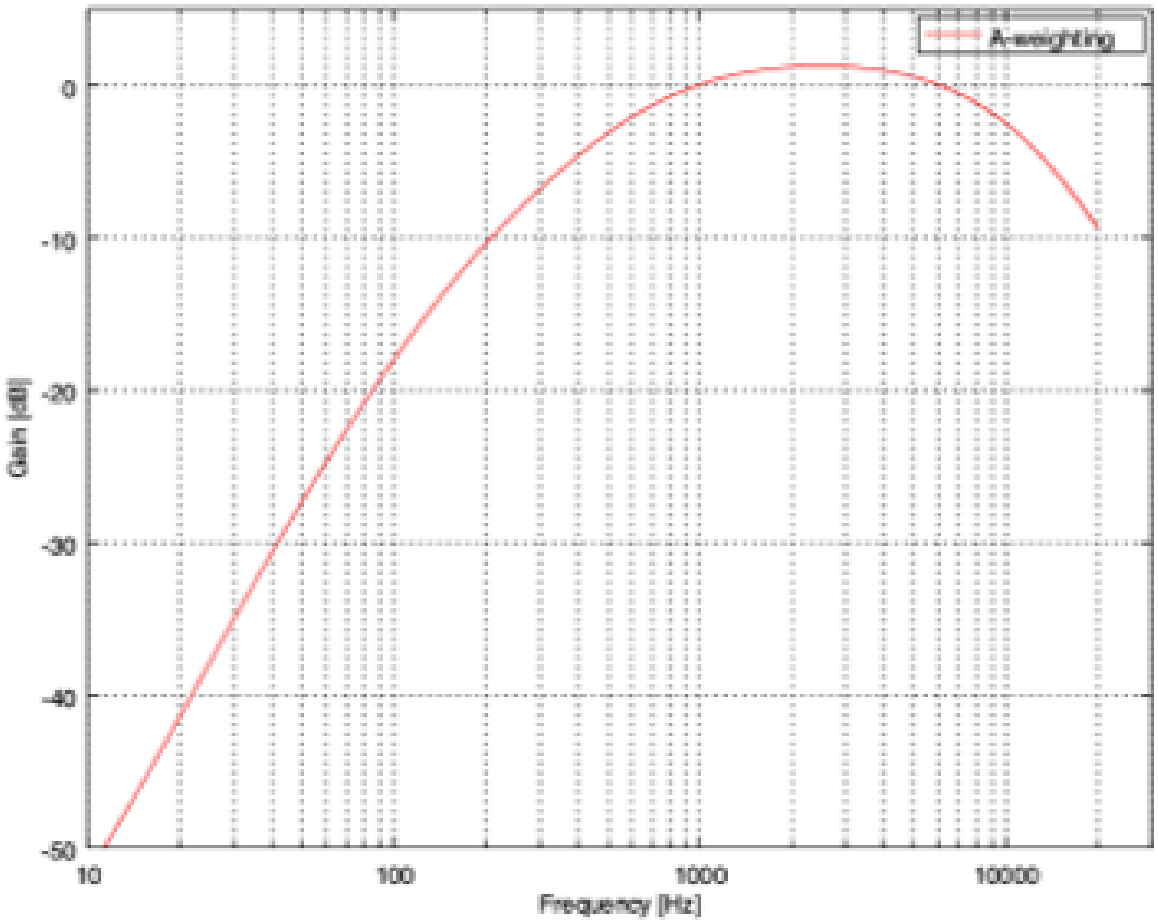


Figura .1 *Filtro de ponderación A.*

- **El pico en escala completa en decibelios (dBFSpeak/dBFSp)** se refiere a la amplitud de una señal de audio digital respecto a la representación digital máxima disponible (escala completa). Una onda sinusoidal con amplitud digital máxima sin recorte tiene un valor pico de 0 dBFSp.

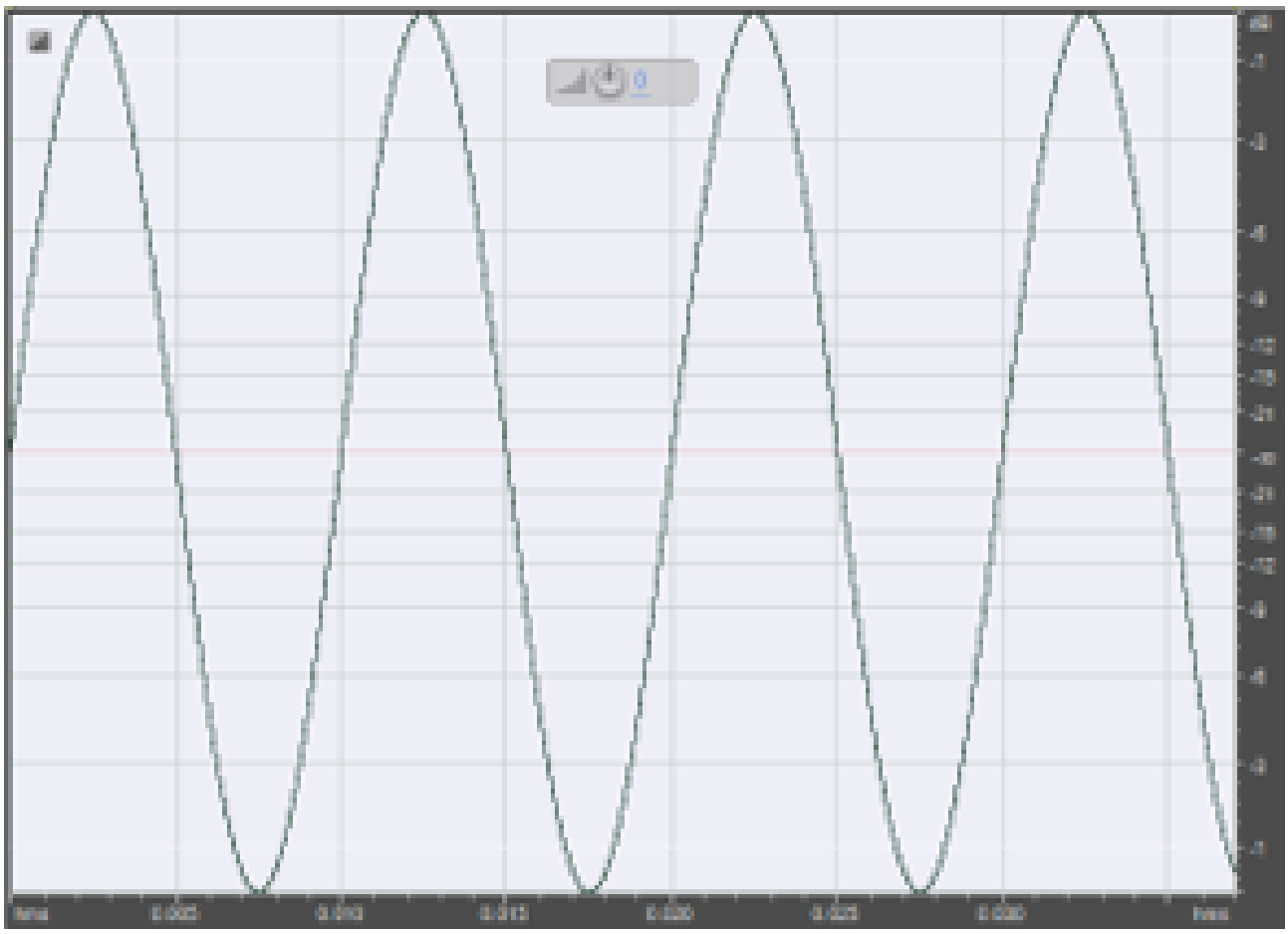


Figura .2 *Un ejemplo de una señal con una amplitud de 0 dBFSpeak.*

- **dBFSrms** hace referencia al nivel de audio digital de una señal digital. Una señal sinusoidal a escala real con 0 dBFSpeak puede tener distintos niveles de RMS dependiendo de cómo se defina la unidad. En las fichas de datos de los micrófonos, una señal sinusoidal a escala completa se define como 0 dBFSrms; por ese mismo motivo empleamos esta definición en este documento técnico. En general, la otra definición de una señal sinusoidal a escala completa —-3 dBFSrms— es más correcta desde el punto de vista matemático y más coherente con la forma en que se definen los niveles de RMS y los niveles de pico en el dominio analógico. Según esta otra definición, una onda cuadrada a plena escala equivale a 0 dBFSrms.
- **La sensibilidad de un micrófono** es su nivel de salida cuando la entrada acústica es un tono sinusoidal de 1 kHz con 1 Pa, 0 dB a o 94 dBspl (tres formas de expresar el mismo nivel de presión del sonido).
- **Relación señal-ruido (SNR): cuando** los niveles de audio se expresan en dB, esta relación se calcula como la diferencia entre el nivel de la señal y el nivel de ruido.
- **La relación señal-ruido (SNR) del micrófono** se calcula como la diferencia en el nivel de salida entre un tono de prueba de 1 kHz a 94 dBspl y el ruido propio del micrófono, cuando este se encuentra en silencio absoluto. La relación señal-ruido (SNR) típica de un micrófono MEMS digital es de 69 dBspl[A].
- **En** este informe técnico, la relación señal-ruido (SNR) del habla es una comparación entre el habla humana y el nivel de ruido ponderado A cuando no hay habla. La relación señal-ruido (SNR) se mide en decibelios. Empleamos una relación señal-ruido (SNR) del habla de 0 dB como definición de la SNR mínima aceptable a la hora de calcular el alcance teórico de captación del habla. En este contexto, el ruido puede provenir del ruido ambiente y del ruido eléctrico, así como del ruido propio del micrófono.
- **Nivel de ruido del micrófono**: el nivel de ruido digital de un micrófono se mide en dBFSrms. Por lo general, los mejores micrófonos digitales tienen un ruido de fondo teórico de -105 dBFSrms[A].
- **El ruido acústico equivalente de un micrófono** se puede calcular como el nivel de prueba de 94 dBspl menos la relación señal-ruido (SNR) del micrófono. Un ejemplo habitual de esto podría ser 94 – 69 = 25 dBspl[A].

Acerca de Axis Communications

Axis contribuye a crear un mundo más inteligente y seguro mejorando la seguridad, la operatividad de las empresas y la inteligencia empresarial. Como líder del sector y empresa especializada en tecnología de redes, Axis ofrece videovigilancia, control de acceso, intercomunicadores y soluciones de audio. Su valor se multiplica gracias a las aplicaciones inteligentes de analítica y una formación de primer nivel.

Axis cuenta aproximadamente con 5.000 empleados especializados en más de 50 países y proporciona soluciones a sus clientes en colaboración con sus socios de tecnología e integración de sistemas. Axis fue fundada en 1984 y su sede central se encuentra en Lund (Suecia).aboutaxis_text2

© 2026 — Axis Communications AB. AXIS COMMUNICATIONS, AXIS, ARTPEC y VAPIX son marcas comerciales registradas de Axis AB en diferentes jurisdicciones. Todas las demás marcas comerciales son propiedad de sus respectivos titulares.Nos reservamos el derecho de introducir modificaciones sin previo aviso.