WHITE PAPER

Microphone pickup range – factors that influence audio performance

September 2026

Table of Contents

1 Introduction

Specifying a definitive maximum pickup range for audio recording devices is rarely straightforward, as performance is linked to the acoustic environment. While technical specifications provide a baseline, the practical recording range depends on a complex interplay of environmental and acoustic factors that vary from one installation to another.

This whitepaper explores the main factors that determine how far a microphone can reliably capture sound, with a focus on speech intelligibility as the primary reference. These factors include background noise level, the spectral characteristics of that noise, reverberation time, microphone directivity, and the self-noise of both the microphone and the recording device.

To establish a practical baseline, we use the human voice as a reference sound source. Noise is measured with an A-weighting filter to get a better correlation between the measurement value and the perceived noise level. The recorded speech signal-to-noise ratio (SNR) is calculated as the difference of the voice level and the A-weighted noise level in decibels *(see appendix for SNR)*. Empirical testing shows that speech intelligibility is accepted down to an SNR of approximately 0 dB (that is, the noise and sound are equally high), provided the noise has a relatively flat frequency response within the speech frequency range.

By understanding these variables, installers and system designers can make more informed decisions when selecting and positioning audio devices in real-world scenarios.

2 Physical recording range

The recording functionality of a microphone generally depends on the environmental background noise. The main factors that limit a physical recording range are:

- **Sound pressure level:** All sounds don't travel the same distance. For example, a gunshot produces an extremely high sound pressure level and can be audible up to a kilometer away. A human voice is however quieter and fades more quickly over distance. When estimating a realistic pickup range, you need to first define the type of sound you need to capture, so that your source level can set the upper limit of what's possible to record.
- **Background noise:** The louder the environment, the harder it is for a microphone to isolate and capture the sound you want. Background noise varies from one environment to another, and competes directly with the target sound. It includes wind, traffic, machinery, and crowd noise, among others. The higher the noise floor at the microphone's location, the shorter the effective pickup range. This explains why the same microphone can perform very differently in a quiet corridor compared to a busy intersection.
- **Spectral profile of the noise:** Just as the loudness of the noise matters, the frequencies it occupies also matters. Human speech sits primarily in the mid-frequency range, roughly 300 Hz to 3,400 Hz. If background noise is concentrated in very low or very high frequencies, it's less likely to interfere with speech intelligibility. However, if the noise overlaps with the same frequency range as the target sound, it becomes difficult to distinguish speech from noise, even at moderate noise levels.
- **Reverberation time at the microphone location:** Reverberation is the persistence of sound after the original source has stopped, caused by reflections off walls, ceilings, floors, and other surfaces. Long reverberation time can degrade speech intelligibility, even in situations where the overall sound level and SNR ratio appear acceptable. In highly reverberant spaces such as large empty halls or tiled rooms, speech sound arrive to the microphone with infinite amount of room reflections making the sound blurry. This makes it difficult for both listeners and recording systems to distinguish individual words clearly.
- **Directivity of the microphone setup:** Microphones vary in how sensitive they are to sounds coming from different directions. An omnidirectional microphone picks up sound equally from all directions, which also means it picks up background noise and reverberation from all directions. A directional setup, such as multiple microphones configured for beamforming, focuses sensitivity in a specific direction which effectively reduces the influence of noise and reverberation coming from other angles. This can significantly extend the usable pickup range in noisy or reverberant environments, but only in the direction the microphone array is aimed.

- **Self-noise of microphones and recording devices:** Every microphone and audio recording device generate a small amount of internal electrical noise, known as self-noise. This sets an absolute lower limit on what the system can capture. No matter how quiet the environment is, sounds that fall below the self-noise floor simply can't be recorded cleanly. When comparing microphones, a lower self-noise specification generally means the microphone can capture quieter sounds better, which may give a longer pickup range if the noise level and the reverberation time is low enough.

To simplify the analysis, we can start with an idealized scenario where there's no background noise, no reflections, and no reverberation and use the human voice as our reference sound source. Under these conditions, the key question becomes: what signal-to-noise ratio (SNR) should we use as the threshold for an acceptable pickup range?

Empirical testing gives us a clear answer. When we apply an A-weighting filter to the system noise and compare the average Root Mean Square (RMS) level of the noise with the average RMS level of speech, speech intelligibility starts to decrease only when the SNR drops below 0 dB. We therefore use 0 dB SNR as our limiting criterion for recording a human voice. This holds true as long as the noise has a relatively flat frequency response, similar to white noise and is within the speech frequency range, which is typically the case for microphone self-noise.

For example, a digital microphone may have a sensitivity of -36 dBFSrms *(see appendix for dBFSrms)* when a 1 kHz tone plays at 94 dBspl (a digitized full scale sinusoidal is defined here as 0 dBFSrms according to microphone datasheets) *(see appendix for dBspl)*. This applies to many Axis products with built-in digital MEMS microphones when the microphone gain is set to 0 dB. When you place this microphone in a very quiet environment where the background noise level falls below 10 dBspl(A) *(see appendix for dBspl(A))*, the internal microphone noise becomes dominant. In this example, the microphone datasheet specifies a digital output of -105 dBFSrms(A). The SNR of the microphone is the difference between the reference level and the noise floor: −36 − (−105) = 69 dB.

We convert this electrical noise floor into an equivalent acoustic noise level: 94 − 69 = 25 dBspl(A).

Now we calculate the maximum distance at which we can record a human voice at 0 dB SNR. According to ITU standards, a human voice has a nominal average speech level (ASL) of approximately 57 dBspl at 1 m. In free space, without reflecting objects, sound pressure level decreases with distance according to the inverse square law:

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

.

If the reference distance is 1 m, to reach a 0 dB SNR at this microphone, the speech level must fall to 25 dBspl (A); a reduction of 32 dB from the 1 mreference level. Rearranging the formula gives us:

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

.

This means that in free space, with no background noise, no wind, and no reverberation, this microphone can theoretically record intelligible speech at approximately 40 m. This is the theoretical maximum pickup range.

However, environmental factors have bigger impact on the noise level than the internal microphone's noise floor. A quiet office after working hours, with ventilation systems still running, has a background noise level of around 28 – 32 dBspl. This alone can reduce the effective pickup range by roughly 50%, bringing the maximum down to around 20 m. If we record outside in a city environment, there will be a combination of noises: wind noise, traffic noise, noises from human activity, and so on. Each noise level and noise spectrum will have its own corresponding pickup range when dealing with an omnidirectional microphone. Here, you can have a pickup range of maximum 2 – 4 m if the traffic noise is heavy.

One practical consideration is that human voice recordings made at or near the 0 dBSNR threshold require significant digital gain to be audible because the recorded voice level is so low that it would otherwise fall below our own hearing threshold during playback. In the recording examples that follow, the audio files have been amplified by +40 dB to compensate for the distant voice level.

3 Recordings in different environments

This section details the recording locations, accompanied by screenshots from the video recordings and voice samples captured at various distances.

Example 1: Location is outdoors with traffic noise from a highway. The background noise level varies from 62 to 68 dBspl[A] and some wind noise is also present.

Figure 3.1 *The recording location is close to a highway (see the red dot).*

Figure 3.2 *In the recording location, the maximum recording distance is 4 m away from an Axis body worn camera.*

Preferably, use a pair of headphones while listening to the recording.

Example 2: A silent corridor with a background noise level of 32 dBspl[A]. Here, the corridor acts like a long acoustic wave guide. Sound travels through the corridor bouncing on the reflective surfaces and arrives to the microphones louder than it would in a free space and at the same distance.

Figure 3.3 *Maximum recording distance is 27 m.*

Preferably, use a pair of headphones while listening to the recording.

Example 3: A large lobby with a background noise level of 42 dBspl[A]. Here, the walls have acoustic absorbents behind the diffusing panels, making most of the early reflections disappear. From a longer distance, the sound travels in the acoustic channel between the floor and the reflective ceiling.

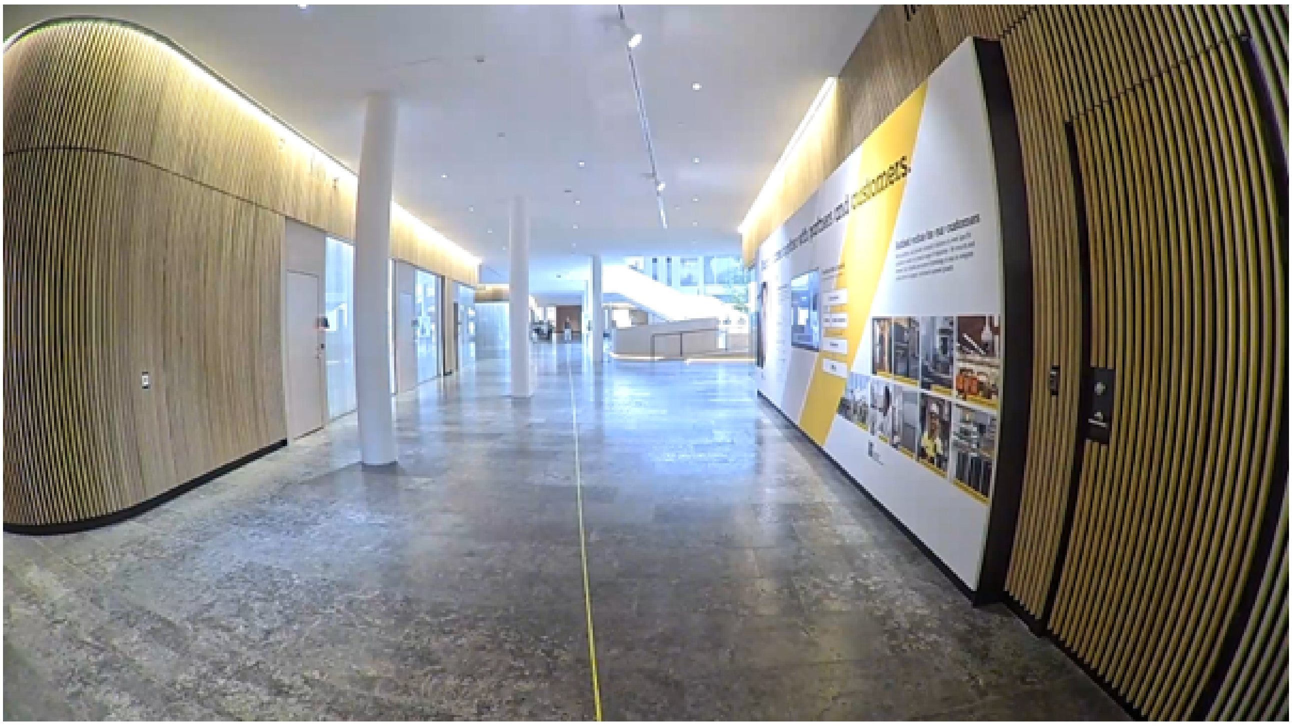

Figure 3.4 *At a maximum recording distance of 40 m. The speaker is barely visible but can still be heard.*

Preferably, use a pair of headphones while listening to the recording.

4 Measurement results from recordings in different environments

This analysis focuses on voice pickup. Stronger sounds, like gunshots, will be heard from a wider distance. We analyzed and measured each sentence from the audio recordings, and estimated signal to noise ratio (SNR) for each recorded sentence.

The SNR was estimated as speech + noise - background noise. A more correct way would be to calculate speech level - noise level,, but there is no possibility to measure the speech level without the noise. This means that when the speech level is close to the noise level, the accuracy of the SNR decreases.

Location and background noise:	Outside with traffic noise 61-68 dBspl[A]	In a silent corridor 32 dBspl[A]	In a large lobby 42 dBspl[A]	Free space, no reflections, complete silence
Distance (m)	**SNR [dB] (estimated)**	**SNR [dB] (estimated)**	**SNR ([dB] estimated)**	**Theoretical SNR = speech-level microphone noise [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	Not audible	17.5	6.3	7.9

24	Not audible	14.8	5.4	4.4
32	Not audible	Corridor too short	4.6	1.9
40	Not audible	Corridor too short	4.7	0.0

In theory, the speech level should decrease by -6 dB when the distance is doubled. This would also mean that the SNR should decrease accordingly. In this table, we can see that the SNR is better than expected in the corridor and in the lobby. The reason is that the reflections in the room increased the sound pressure levels. Especially in the corridor, the walls, floor, and the ceiling act as a waveguide for the sound. This makes the loudness and the SNR about 10 dB better than the theoretical SNR in free space, when tested at 24 m distance, in a corridor.

Whenever the background noise exceeds the acoustic equivalent self-noise of a single omnidirectional microphone, the pickup range depends entirely on the environmental background noise rather than the microphone.

Appendix 1 Appendix

- **Decibels Sound Pressure Level (dBspl):** The sound pressure level in decibels (dB) with reference to 20 µPa. When measuring noise and voice levels, time-integrated RMS levels are often used. This is sometimes referred to as a power level.

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

  The two most common integration times are "fast RMS" with a time window of 125 ms and "slow RMS" with a time window of 1 second. For noise floor analysis, the level is often averaged over a longer time.

- **Sound Pressure Level weighted by the A-weighting filter (dBspl[A]):** This is sometimes written as dB(A). It is the A-weighted value of the dBspl level. A-weighting is a way of filtering the microphone signal to get a better match between the human hearing and the measured sound pressure level. This is because the human hearing has a reduced sensitivity for low frequencies below 500 Hz and high frequencies above 10 kHz.

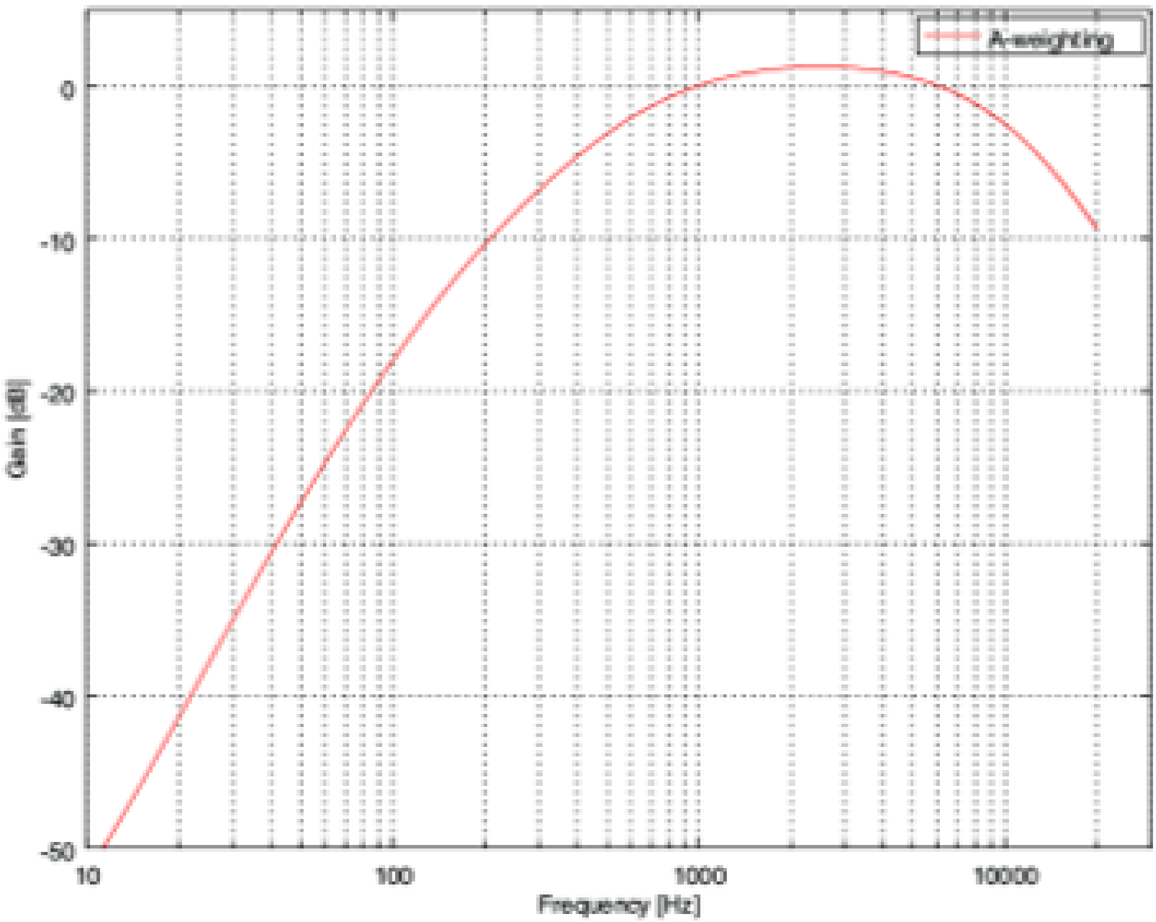

Figure .1 *A-weighting filter.*

- **Decibel Full Scale peak (dBFSpeak/dBFSp)** refers to a digital audio amplitude with reference to the maximum available digital representation (Full Scale). A sinusoidal with maximum digital amplitude without clipping, has a peak value of 0 dBFSp.

Figure .2 *An example of a signal with 0 dBFSpeak amplitude.*

- **dBFSrms** refers to a digital audio power level of a digital signal. A full scale sinusoidal with 0 dBFSpeak can have different RMS levels depending on how the unit is defined. In microphone datasheets, a full scale sinusoidal is defined as 0 dBFSrms, same reason why we use this definition in this white paper. In general, the other definition of a full scale sinusoidal being -3 dBFSrms, is more mathematically correct and more consistent to how RMS levels and peak levels are defined in the analog domain. A full scale square wave is 0 dBFSrms in this other definition.
- **Microphone sensitivity** is the output level of a microphone when the acoustic input is a 1 kHz sinusoidal tone with 1 Pa or 0 dB a or 94 dBspl (three ways of expressing the same sound pressure level.)
- **Signal-to-Noise Ratio (SNR):** When audio levels are expressed in dB, the ratio is calculated as the difference between the signal level and the noise level.
- **Microphone SNR** is calculated as the difference in output level between a test tone of 1 kHz at 94 dBspl and the microphone's self-noise, when placed in complete silence. A typical SNR for a digital MEMS microphone is 69 dBspl[A].
- **Speech SNR** in this whitepaper, is a comparison between human speech and the A-weighted noise level when no speech is present. SNR is measured in decibels. We use 0 dB speech SNR as a definition of the lowest acceptable SNR when calculating theoretical pickup range for speech. In this context, noise can come from environmental noise and electrical noise as well as microphone self-noise.
- **Microphone noise level** The digital noise level of a microphone is measured in dBFSrms. Typically, the best digital microphones have a theoretical noise floor of -105 dBFSrms[A].
- **Acoustic equivalent noise of a microphone** can be calculated as the test level of 94 dBspl – microphone SNR. A common example of this could be 94 – 69 = 25 dBspl[A].

About Axis Communications

Axis enables a smarter and safer world by improving security, safety, operational efficiency, and business intelligence. As a network technology company and industry leader, Axis offers video surveillance, access control, intercoms, and audio solutions. These are enhanced by intelligent analytics applications and supported by high-quality training.

Axis has around 5,000 dedicated employees in over 50 countries and collaborates with technology and system integration partners worldwide to deliver customer solutions. Axis was founded in 1984, and the headquarters are in Lund, Sweden.