



WHITEPAPER

Enabling emotion detection in AI using high-SNR MEMS devices

Dr. Gunar Lorenz, Senior Director Technical Marketing Sensors, Infineon Technologies AG

Table of contents

Abstract	3
1 The human experience of sound	4
2 Sound transmission and the role of devices	6
3 Why high-quality sound matters	8
4 Sound as AI's gateway to understanding	9
4.1 Diarization	9
4.2 Advanced algorithms and the hardware connection	10
5 Infineon MEMS microphones – Performance that stands out	13
5.1 Infineon's MEMS solutions for diverse needs	14
6 Conclusion	15
References	16

Abstract

In 1940, Disney released an animated movie called Fantasia, with the soundtrack being a major attraction. The intent of the movie makers was to have a sound system that would make the audience feel as though they were standing on the podium with the orchestra. To that end, they devised an elaborate system consisting of 33 microphones with 8 sound recorders and a bunch of engineers, taking 42 days and a final recording film that ran 147,000 meters in length.

Today, recording something similar might take nothing more than a flagship phone.

Ever since audio recording began almost a century ago, the quest for perfect audio has continued. Levels of noise such as static and crackling which were seen as inevitable at one point are now virtually eliminated. The only justification for experiencing noise in any form of audio today is that it was added intentionally, most likely for artistic and/or nostalgic purposes. Today's audio is expected to be clean and crystal clear, irrespective of the device with which the recording occurs. And in today's age of almost-invisible earphones, the saying "size doesn't matter" is not really applicable. Size, in fact, does matter, but in reverse. The smaller, the better. More discrete; more valuable.

So how did we get here? How did the pursuit of reducing product size while improving quality begin and what is its current state? What does the future look like, for this pursuit, especially in the age of AI? This whitepaper answers these questions.

1 The human experience of sound

Sound is central to how we connect and communicate, and there is more to sound than just hearing words. Our voices convey emotions, intent, and context through tone, pitch, and rhythm. In fact, research indicates that actual words make up only about 7% of our communication. Vocal tone accounts for 38%, and body language makes up the remaining 55% [1]. Humans naturally tune out irrelevant noises and focus on meaningful sounds, especially speech, because it conveys crucial social and emotional information.

In this age of long-distance connectivity, high-quality audio recording is therefore vital in daily interactions to capture nuances to understand not just what we are saying, but how we are feeling as we say it. Ultimately, sound is about more than listening; it is about understanding. The clearer the audio, the more natural and meaningful our interactions—whether with each other or with intelligent machines.

Technically, sound is a combination of two key factors – frequency and sound pressure level (SPL). SPL, what we call as volume in everyday language, measured in decibels (dB_{SPL}) indicates the “amount” of sound in the listener’s immediate vicinity. The more the dB_{SPL} value, the higher is the sound output available for perception. For context, an SPL of 100 dB_{SPL} is comparable to a very loud noise from a lawn mower or helicopter.

The lowest point in the SPL range (0 dB) is associated with a sound pressure oscillation of 20 μPa , which represents roughly the hearing threshold at 1 kHz of a young, healthy individual with optimal hearing. All human sounds related to speech fall into the frequency band of 100 Hz and 8 kHz. The corresponding human hearing threshold according to ISO 226:2023 [2] is shown in Figure 1.

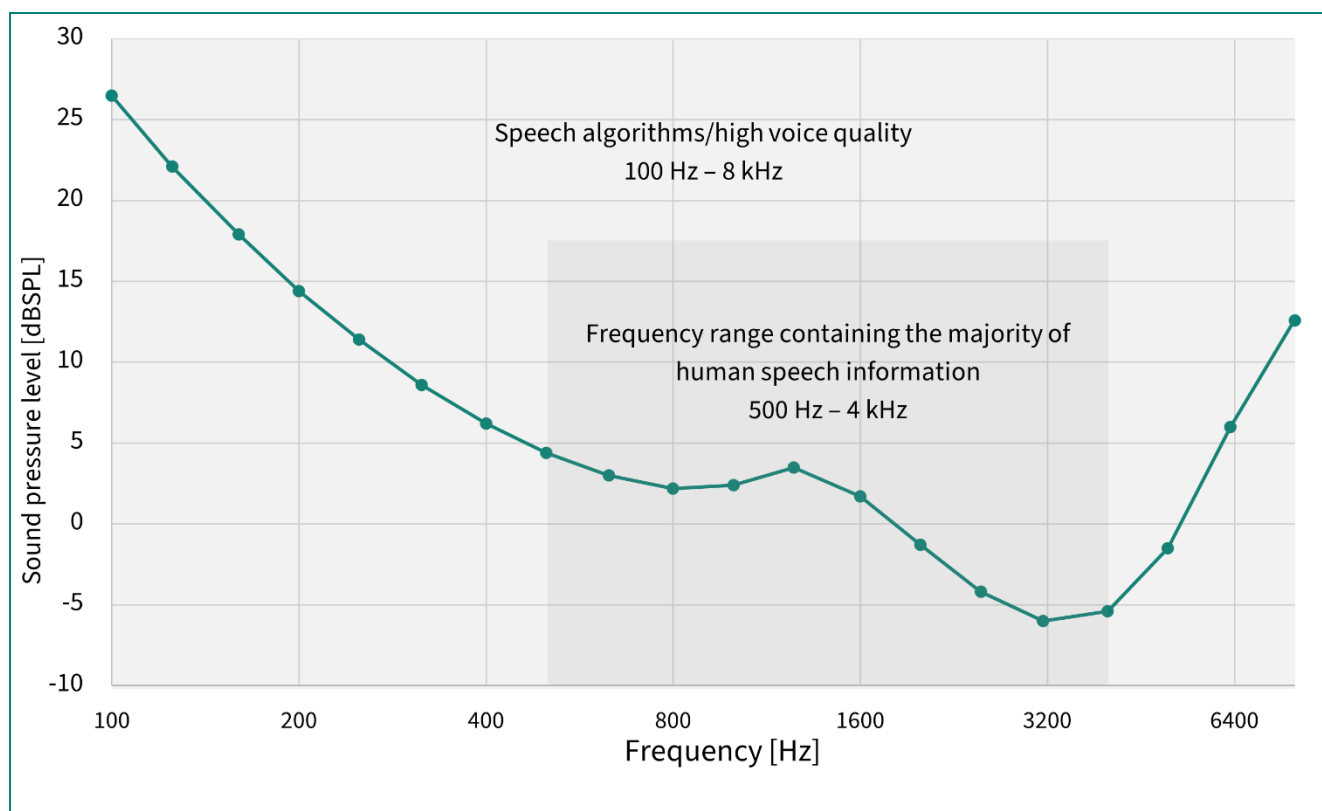


Figure 1 Threshold of hearing: The level of sound at which a person gives 50% of correct detection responses on repeated trials, according to ISO 226:2023

Although an individual human’s perception of sound relies equally on frequency and volume, most humans are particularly sensitive to frequencies between 500 Hz and 6 kHz—the range crucial for clearly understanding speech. Even within this range, specific frequencies are of particular importance. For example, the 1 kHz frequency is the range where the sensitivity of the human ear is at its highest. A 1 kHz sine sound wave is often used for testing audio parameters, such as the sensitivity of microphones. Most music is within the 5 kHz to 10 kHz range, and so any speech

within this range becomes comparatively difficult to comprehend, all the more due to the presence of sibilance (words with hissing sounds like ship, chip, zip, etc.). In fact, sibilance in the range of 6 kHz to 8 kHz has a noticeably detrimental effect on speech intelligibility.

All of this is generalized data, which needs to be adjusted to counter the reduction in loss of hearing that occurs naturally as humans age. Figure 2 shows how the SPL (volume) needs to increase with age for similar auditory perceptions. Measurements are as per ISO 7029:2017 [2] and are made under monaural earphone conditions, meaning one earpiece in one ear, with the other ear free to hear all ambient noises. Note that a similar graph exists for females which shows a slightly lower degradation in comparison.

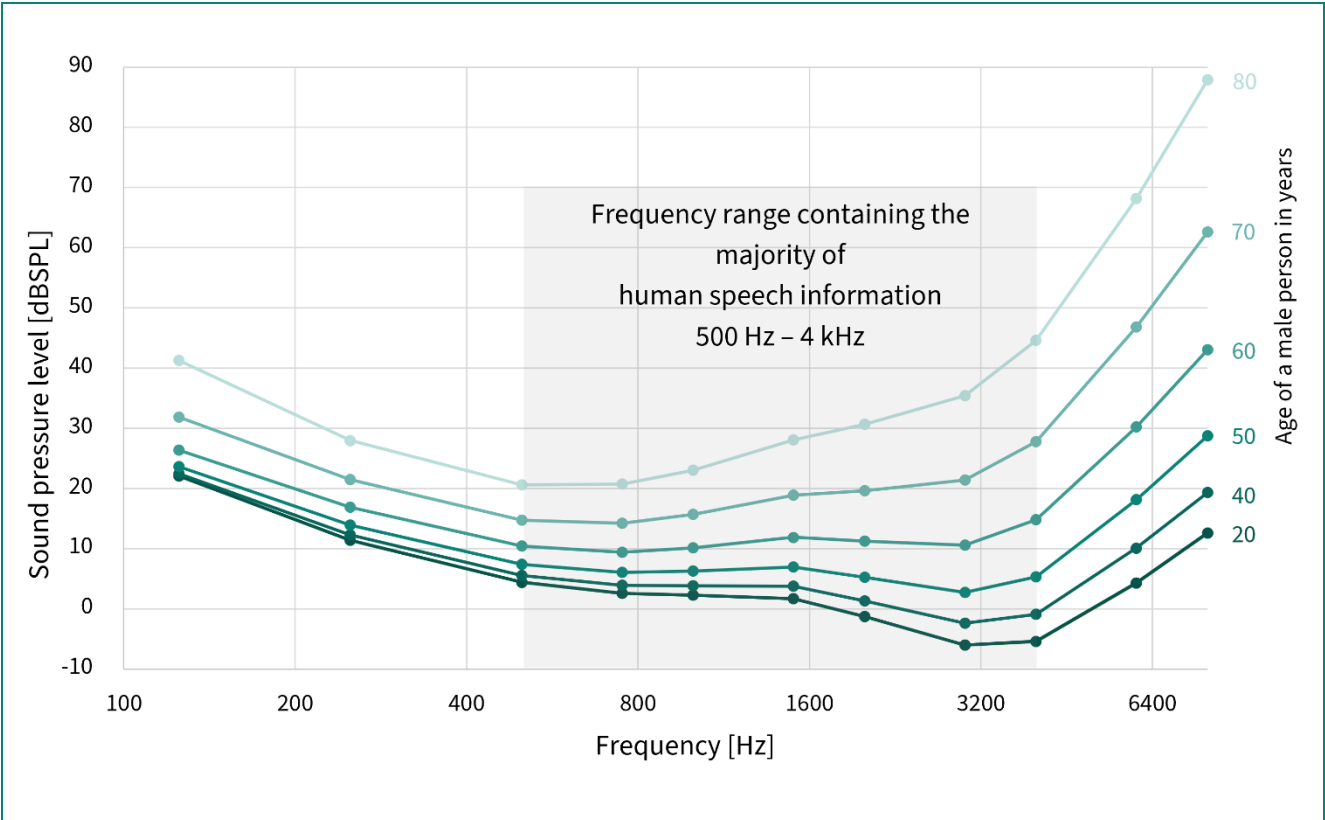


Figure 2 Hearing threshold degradation of an ontologically normal human male with age

2 Sound transmission and the role of devices

Now that we have explored how critical clear audio is to human interactions, especially for elderly people, it is important to consider how this audio travels from our voices to the listener—or even to an AI assistant. The clarity we rely on every day depends largely on the technology within our devices, particularly microphones.

Microphones act like the ears of our devices. They transform physical sound vibrations into electrical signals that can be easily processed and transmitted. As we have seen, capturing the subtle nuances of speech—like tone, emotion, and intent—is key to effective communication. Therefore, the performance of microphones plays a significant role in preserving these important details during transmission.

The quality of the audio recordings produced by a MEMS microphone depends on its dynamic range. The upper limit of the dynamic range is defined by the acoustic overload point (AOP), which defines the microphone's distortion performance at high SPLs. The self-noise of the microphone limits its dynamic range at the lower end of the spectrum. One of the most influential factors in the preservation of important details during transmission is the signal-to-noise ratio (SNR). Basically, the SNR describes how much unwanted self-noise of a microphone is carried through in the wanted audio signal. A microphone with a higher SNR is expected to carry the recorded audio more clearly by effectively minimizing noise added by the microphone itself and amplifying the intended sounds. For devices we use daily—such as smartphones or earbuds—achieving high SNR is essential to clearly capture low volume conversations without losing critical details or context.

However, we are now in the age of AI, in which there is a lot of human-machine verbal communication. Although the traditional measure of a microphone's quality is SNR, it is somewhat misleading for the purposes of audio comprehension by AI, as SNR implies how humans perceive an audio signal using A-weighting.

AI perceives and processes audio very differently than humans.

If the intended recipient of the recorded signal is an AI, a related microphone parameter, called equivalent noise level (ENL), is a more appropriate way to specify performance, as it ignores the human perception element of recorded sound. ENL refers to the signal produced by the microphone in the absence of an external sound source. ENL indicates the sound pressure level (SPL) that will create the same voltage as the self-noise from the microphone. The microphone's ENL across frequency can be considered the closest match to a microphone's hearing threshold. It should be noted that this is a highly simplified assumption, as there are usually numerous other components in the audio chain, including sound channels, additional water protection, and the audio processing chain. [Figure 3](#) shows a visual representation of the ENL curves of two MEMS microphones in comparison with the human hearing threshold.

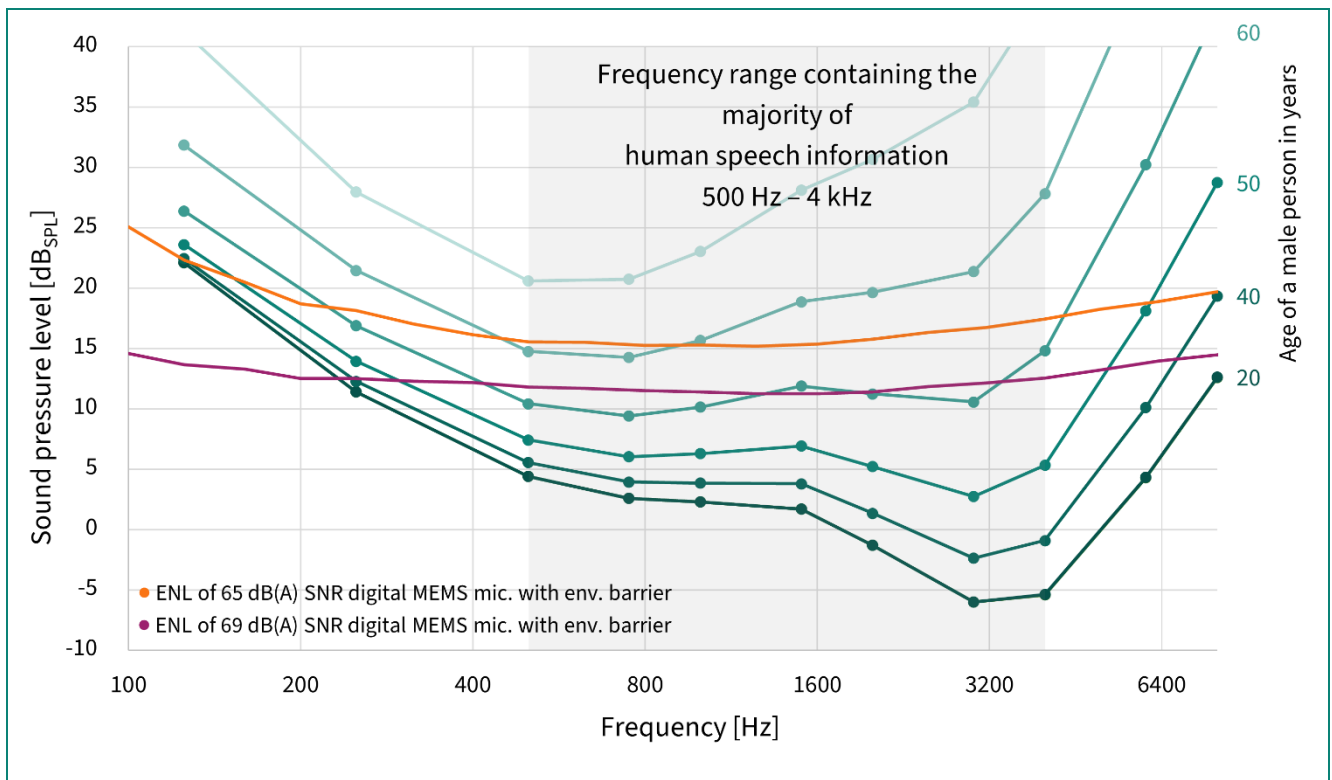


Figure 3 Acoustic 1/3 octave self-noise (ENL) of a mid-range and a high-end MEMS microphone compared to the hearing threshold of a typical male person

The orange line depicts the ENL curve of a 65 dB(A) SNR microphone with an integrated environmental barrier for dust protection. The corresponding MEMS microphone is currently used in several high-end smartphones from a variety of vendors. The purple line below illustrates the ENL curve of Infineon’s latest high-end digital microphone, which features an innovative environmental barrier that protects against particles and humidity. This microphone represents the current state of the art and was released in 2024 in a high-end tablet. 2024 also saw the release of many smartphones having microphones with comparable performances. It is worth noting that reducing the microphone self-noise by 5–10 dB_{SPL} represents a significant achievement, particularly when considering the logarithmic scale of sound pressure. It should also be noted that this assessment only considers the self-noise of the MEMS microphone and does not take into account the additional noise sources of the audio chain, which will further reduce the overall performance.

The current limitations of MEMS microphone technology are most apparent in the frequency range that contains the majority of human speech information (500 Hz to 4 kHz). Even the most sophisticated MEMS microphones currently on the market are only capable of capturing sound at a dB_{SPL} level comparable to that of a 60-year-old male.

The next phase of development of these high-end microphones will improve their ability to discern low sound pressure levels when compared to the human ear, in particular, the 2 kHz range. This range is of paramount importance for ensuring high levels of sound intelligibility for human listeners. The long-term goal from Infineon is to reduce the discrepancy of between 12–17 dB_{SPL} between the capabilities of a young person’s hearing and Infineon’s state-of-the-art microphone.

3 Why high-quality sound matters

As discussed, sound is more than just hearing—it is about understanding. High-quality audio is no more a luxury but a necessity for clear and effective communication, especially in our technology-driven world. But why exactly does sound quality matter so much?

Imagine sitting in a busy café trying to have an important conversation. The background noise makes it difficult to hear clearly, and soon you are misinterpreting words or asking your companion to repeat themselves constantly. This frustration is not just an inconvenience; it can significantly impact the depth and quality of our interactions.

Moreover, high-quality sound has profound implications in specialized applications, such as assisting elderly individuals. Many older adults experience mild hearing loss, making everyday interactions challenging, especially in noisy environments. For them, audio clarity is not merely about convenience but also about maintaining meaningful social connections and quality of life.

This need for audio clarity becomes even more critical when we consider AI-driven communication.

AI has quietly transformed our daily routines in ways we might not even notice. From asking Siri about the weather to having Alexa play our favorite playlist, voice-based interactions are becoming second nature. This trend is constantly accelerating. AI's presence in our everyday lives is growing rapidly, impacting everything from productivity to how we communicate with loved ones.

When AI interacts with people, it has no visual cues or body language to rely on, making audio quality paramount. High-quality sound directly enhances AI's ability to interpret human speech accurately, helping it to recognize not just words, but the emotional and contextual layers beneath them. For instance, a personal assistant powered by AI might detect urgency or anxiety in someone's voice, prompting quicker responses or more empathetic interactions. There are already tools available on the market like Sesame AI [4] that are making inroads into emotional comprehension by AI.

One area seeing significant benefits from AI advancements is the support provided to the elderly. Many seniors experience hearing difficulties, especially in noisy environments, making even simple conversations stressful. AI-enabled devices equipped with high-quality audio inputs can offer crucial assistance by clearly capturing and amplifying conversations, making communication easier and more accessible. Imagine AI assistants that can recognize when a user struggles to hear, automatically adjusting volume or clarity to accommodate their needs.

Similarly, AI-driven live translation is breaking down language barriers in personal and professional contexts. However, these translations are only as effective as the audio clarity allows. High-quality audio ensures that AI accurately distinguishes speakers, identifies emotional cues, and translates messages without losing context or nuance.

Beyond these everyday conveniences, AI-driven audio capabilities could also significantly impact fields like telemedicine, education, and emotional support. Imagine healthcare applications where AI assistants accurately interpret patient voices to detect subtle emotional changes, enabling doctors to intervene proactively. Or educational tools that recognize a student's confusion or engagement simply by analyzing their vocal tone, adjusting lessons accordingly.

But, as mentioned, AI perceives and processes audio very differently than humans.

4 Sound as AI’s gateway to understanding

Think of AI’s audio capabilities as steps on a ladder, each representing a progressively deeper level of understanding. At the base level, AI recognizes simple commands—this is the “wake word” detection stage. Saying “Hey Siri” or “Alexa” instantly activates your assistant, allowing basic interactions without much complexity.

Moving one step up, we have speech recognition, where AI converts spoken words into text. This level helps devices perform straightforward tasks, like setting alarms or sending text messages. But, as we have seen, human conversations are not merely exchanges of words—they involve emotions, intent, and multiple speakers, often simultaneously.

This is where speaker diarization, the next step, comes in. Diarization helps AI understand “who spoke when”, crucial in group conversations and translation scenarios. Infineon’s research illustrates the importance of diarization clearly: without it, conversations can feel confusing, like watching a play without knowing who is speaking each line. Diarization brings clarity, context, and structure, significantly improving the accuracy and usefulness of AI systems.

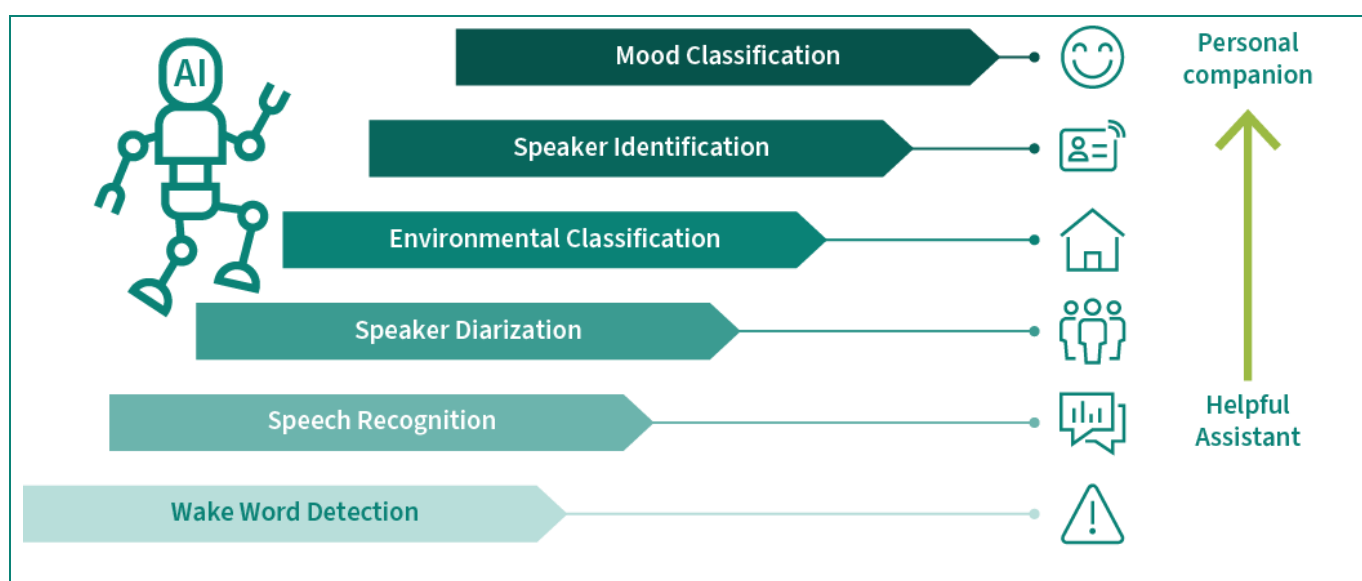


Figure 4 AI capability ladder

Higher levels on this AI ladder include recognizing individual speakers (speaker identification), understanding environmental contexts, and even detecting emotional nuances. Each step upwards depends heavily on sound quality—clearer audio means more accurate AI interpretations.

4.1 Diarization

To validate all the AI hype from the major phone manufacturers last year and to evaluate the latest advances in live translation, a then-latest flagship smartphone was used to follow a Korean conversation for translating to English. The English text appeared on the screen with a slight delay but the speed of the translations were impressive, nonetheless. However, on occasion, the meaning of the original text was not fully captured. The real challenge arose when additional people joined the Korean conversation, as the English translation appeared on the screen without any indication of who said when. The multiple and sometimes overlapping speakers, in combination with the small delay between the spoken and the translated word, made it difficult to follow the conversation.

Of the several live translation apps available then, none of them added IDs or labels to indicate who was speaking. In AI terminology, understanding “who spoke when” is called speaker diarization. Without it, live translation feels like reading a Shakespeare play without knowing who’s who – you cannot really understand the meaning. Have a look at the text on the left in [Figure 5](#).

<p>What an AI gets without diarization:</p> <p>Have not saints' lips, and holy palmers too?– Ay, pilgrim, lips that they must use in prayer. O, then, dear saint, let lips do what hands do; Saints do not move, though grant for prayers' sake. Then move not, while my prayer's effect I take. Thus from my lips, by yours, my sin is purged. Then have my lips the sin that they have took. Sin from thy lips? O trespass sweetly urged! Give me my sin again. You kiss by the book.</p>	<p>What an AI gets with diarization:</p> <p>Speaker 1: Have not saints lips, and holy palmers too? Speaker 2: Ay, pilgrim, lips that they must use in prayer. Speaker 1: O, then, dear saint, let lips do what hands do; They pray, grant thou, lest faith turn to despair. Speaker 2: Saints do not move, though grant for prayers' sake. Speaker 1: Then move not, while my prayer's effect I take. Thus from my lips, by yours, my sin is purged. Speaker 2: Then have my lips the sin that they have took. Speaker 1: Sin from thy lips? O trespass sweetly urged! Give me my sin again. Speaker 2: You kiss by the book.</p>	<p>Shakespeare's original intent:</p> <p>Romeo: Have not saints lips, and holy palmers too? Juliet: Ay, pilgrim, lips that they must use in prayer. Romeo: O, then, dear saint, let lips do what hands do; They pray, grant thou, lest faith turn to despair. Juliet: Saints do not move, though grant for prayers' sake. Romeo: Then move not, while my prayer's effect I take. Thus from my lips, by yours, my sin is purged. Juliet: Then have my lips the sin that they have took. Romeo: Sin from thy lips? O trespass sweetly urged! Give me my sin again. Juliet: You kiss by the book.</p>
--	---	--

Figure 5 Diarization on a classic dialog from Shakespeare's *Romeo and Juliet* [5]

That is the extent of what an AI would get from a smartphone recording of a beautifully acted Shakespeare play. With proper diarization (text in the middle in Figure 5), the AI gets a clear indication that the recorded audio contains a dialog involving two people. Based on the content, and the fact that two people are talking, AI might guess the romantic nature of the conversation. What is still not possible for AI, even with proper diarization, is determining the sex and the age of the two people involved.

4.2 Advanced algorithms and the hardware connection

Clear audio is essential, but how exactly do advanced algorithms use this audio to achieve tasks like speaker diarization? In other words, how does AI understand sound, when compared with humans? Just as the human brain breaks down the audio into smaller pieces of recognizable patterns and maps it to existing knowledge, AI uses algorithms to interpret audio signals that it receives.

We at Infineon initially tried a publicly available tool to measure the error in identifying speakers based on an audio file of a conversation. Eventually we used NVIDIA's NeMo framework [11] to provide a clear example of how these sophisticated algorithms work—and why hardware quality is vital.

From the Nvidia website [6]: "Speaker diarization is the process of segmenting audio recordings by speaker labels and aims to answer the question "who spoke when?"

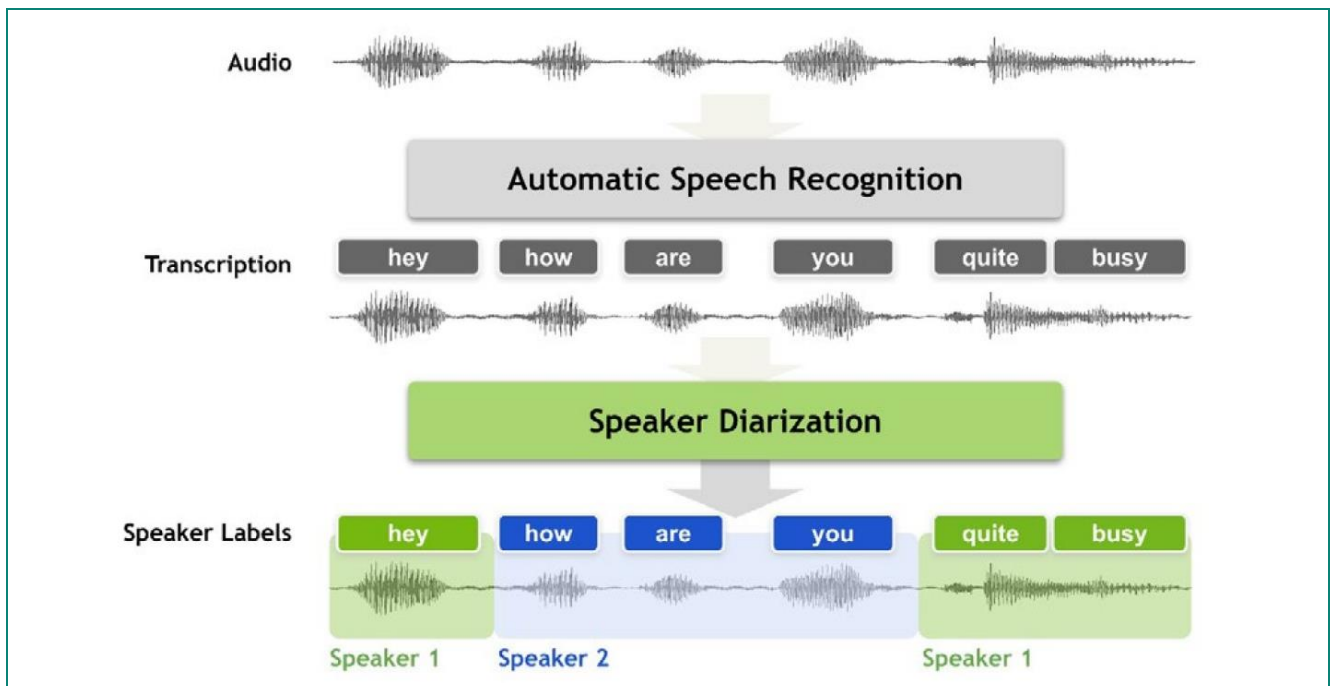


Figure 6 Diarization – a visual explanation from Nvidia [6]

The algorithm must perform a series of steps before decoding the input speech to determine and attach speaker labels appropriately. The first step is to detect voice activity in the first place – the human equivalent of listening to a sound. This depends almost entirely on the hardware setup and its capabilities to extract speech signals even in noisy surroundings. However, our current experiments were conducted in controlled environments to begin with, starting off with comparatively simpler settings so that the effect of each additional parameter could be clearly and unambiguously understood. For instance, our two test speakers did not particularly stress NeMo, as they added a small pause before speaking each time. Neither of the speakers spoke at the same time. Then, for the second test we added background noise by performing the same conversation in our very busy Infineon cafeteria.

For our first test setup we used Infineon’s microphone evaluation board Audiohub Nano [7] which allows to record audio with two different MEMS microphones at the same time. The first microphone is roughly the same as one can find in most high-end mobile phones with an SNR of 65 dB(A). For the other microphone, as a high-end reference for our test, we used our latest digital XENSIV™ microphone IM72D128V [8] with an SNR of 72 dB(A). Figure 7 shows the audio signals of the recorded conversation from the first test, represented as recorded audio streams of the two microphones, analyzed with NeMo. Note that the two microphones were used for performance comparison against each other, with both speakers speaking into both microphones. The setup was not intended to have one speaker per microphone, but to have both voices recorded by both microphones, to observe which microphone had a higher ability to detect multiple speakers, as demonstrated in the results below.

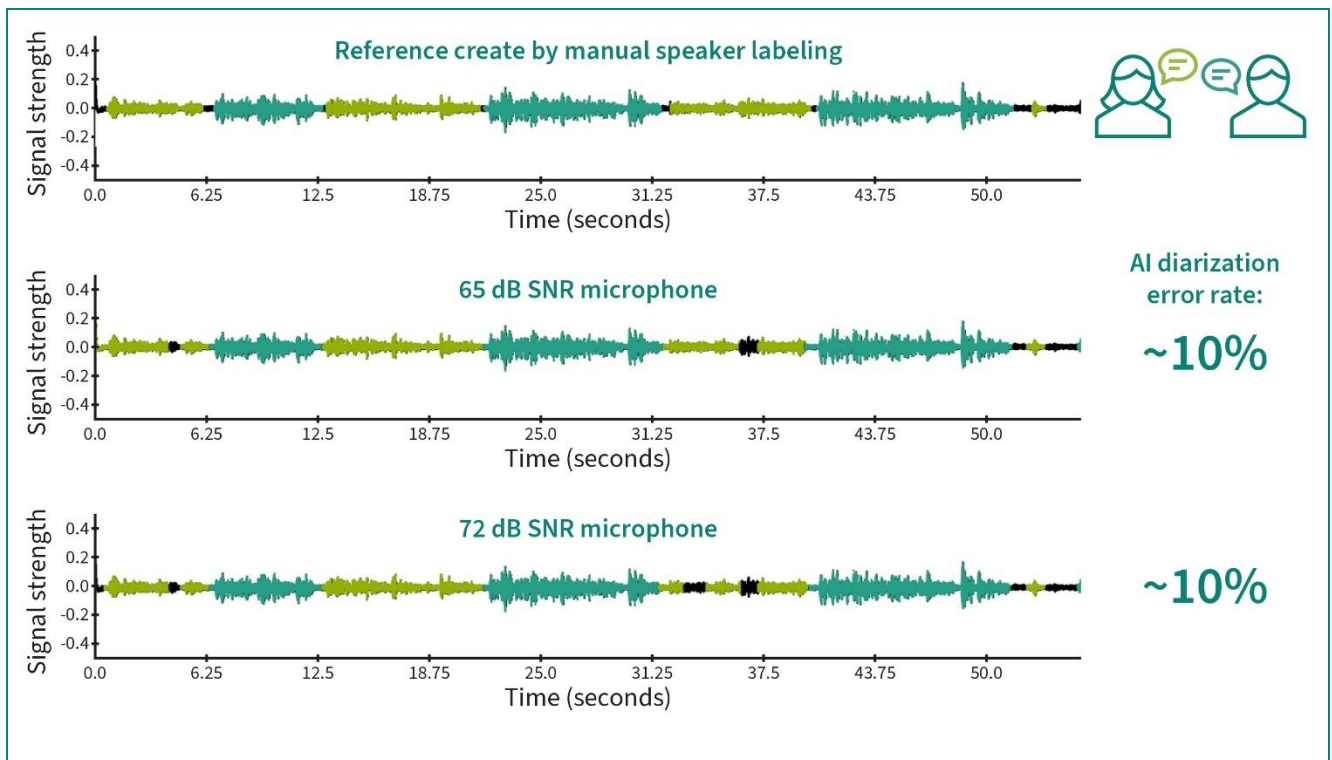


Figure 7 Diarization accuracy levels by NeMo using Infineon’s Audiohub Nano [7] evaluation board

The two colors in the waveforms represent the two speakers, with the first speaker’s audio shown in green and the second speaker’s in blue. The reference was provided to NeMo by manual labeling. Each time a speaker spoke, a specific button was pressed. The graph in the middle shows the NeMo’s automatic speaker labeling using the standard smartphone microphone whereas the bottom graph shows NeMo’s speaker labeling using the high SNR microphone from Infineon. Both graphs show that NeMo correctly identified speaker 1 and 2 for the most part expect for a few “black” sequences in the middle, which it could not attribute to either of the speakers. For both microphones, NeMo calculated a fairly low diarization error of about 10%. Thus, regardless of the used microphone NeMo could successfully identify both speakers in 90% of the time, in surroundings that were not noisy.

However, the presence of background noise, as performed in our second test in the noisy cafeteria, had starkly different results. With background noise, NeMo struggled to even detect the second speaker when the recoding was done with the medium SNR microphone. The diarization error rate was as high as ~70%. With the high SNR microphone in the same environment, the diarization error rate was dramatically reduced to about ~22%.

The conclusion is clear – what seems to work well in a quiet environment does not work well in noisy environments.

These findings clearly show that advanced algorithms rely heavily on hardware quality. Even the best algorithms like NVIDIA’s NeMo need high-quality audio to deliver accurate, reliable results. Hardware is not just complementary—it is fundamental to the effectiveness of AI-driven audio applications. This is all the more so for the elderly with their natural hearing loss, and it becomes the responsibility of the hardware-software combination to offset this loss by understanding accurately the persons speaking as well as their ambience.

5 Infineon MEMS microphones – Performance that stands out

With so many MEMS microphones available, what sets Infineon apart? It boils down to a longstanding commitment to audio clarity and innovative design, now being tailored specifically for modern AI and consumer electronics, as well as specialized applications such as hearing aids. Infineon is the market leader for silicon microphones (SiMiC) with a >45% market share since 2021 [9], with ~3.5 billion MEMS sold in that year, and more than 200 million packaged microphones sold.

A unique strength of Infineon's MEMS microphones is their ability to achieve exceptionally high SNR ratings while maintaining extremely compact form factors and still maintaining the same cost optimization standards, as shown in Table 1. This provides Infineon MEMS devices a very wide dynamic range for stable handling of low volume sounds as well as high peak signals, best suited for devices requiring active noise cancellation (ANC).

Table 1 High SNR in small packages

SNR	Package size	Product name
73 dB (analog)	4 mm × 3 mm × 1.2 mm	IM73A135
72 dB (digital)	4 mm × 3 mm × 1.2 mm	IM72D128V
69 dB (digital)	3.5 mm × 2.65 mm × 1.0 mm	IM69D128S

Additionally, some of Infineon’s microphones feature specialized protective barriers against environmental factors like dust, particles, and humidity—elements that could otherwise degrade audio quality over time—with IP57 rated devices. These innovations reduce the need for waterproof meshes and ensure that devices maintain consistent audio performance throughout their lifespan, even under demanding daily usage conditions.

With wireless devices becoming the norm, Infineon devices operate with different power modes including ultra-low power for extracting the maximum battery life, particularly useful for in-ear devices, TWS, and hearing aids.

5.1 Infineon's MEMS solutions for diverse needs

Infineon offers a range of MEMS microphones specifically designed to address the growing demands of audio clarity and AI applications across various consumer devices, from mid-range smartphones to premium earbuds and tablets.

Table 2 Products with features fine-tuned for specific applications [10]

Application	Use cases	Specific requirements	Products
Truly wireless stereo (TWS) earbuds	<ul style="list-style-type: none">– Voice calling– ANC	<ul style="list-style-type: none">– Small or medium form factor– Low power consumption	<ul style="list-style-type: none">– IM69D128S– IM70A135– IM68A130
Headphones	<ul style="list-style-type: none">– Voice calling– ANC	<ul style="list-style-type: none">– Low power consumption– Highest SNR	<ul style="list-style-type: none">– IM73A135– IM72D128– IM69D128S
Laptops and tablets	<ul style="list-style-type: none">– Conference calling– Audio recording	<ul style="list-style-type: none">– Highest SNR– Smallest form factor for thinnest bezels	<ul style="list-style-type: none">– IM70D122– IM73D122
Conferencing systems	<ul style="list-style-type: none">– Far field voice pickup	<ul style="list-style-type: none">– Array mics with up to 32 mics per system– Beamforming and source separation	<ul style="list-style-type: none">– IM73D122– IM72D128– IM73A135
Home IoTs, smart speakers	<ul style="list-style-type: none">– Voice assistance– Far field voice pickup	<ul style="list-style-type: none">– Good phase and sensitivity matching– Multi-mic configuration– High AOP levels	<ul style="list-style-type: none">– IM72D128– IM70A135– IM69D128S

In short, Infineon's MEMS microphones are more than a one-size-fits-all approach; they are carefully designed solutions tailored to diverse consumer and AI needs. Whether for everyday communication or demanding professional applications, Infineon delivers clarity, reliability, and performance, enabling seamless audio interactions across the board.

6 Conclusion

As explored throughout this paper, clear audio is not just a convenience — it is crucial for meaningful human interactions and effective AI communication. From the subtle emotional cues we rely on daily to sophisticated AI capabilities like speaker diarization and real-time translations, audio clarity directly shapes our experiences.

Infineon's MEMS microphone solutions sit at the heart of this audio evolution, delivering high-quality sound-capture across diverse applications. Whether enhancing everyday smartphone conversations, facilitating accurate translations in noisy settings, or improving AI-driven emotional and contextual understanding, Infineon's microphones help bridge the gap between human and machine communication.

But what does the future hold? As AI continues to evolve, devices and applications relying on audio clarity will multiply and diversify even further. From advanced healthcare monitoring and personalized education to deeper emotional interactions with virtual assistants, high-quality sound capture will be increasingly critical. Product hardware will end up becoming the limiting factor that influences the performance of the AI assistant, so products that have not been designed with the long-term future in mind will experience a faster end of life.

To stay ahead of this trend, Infineon continues to innovate, constantly improving MEMS technology to ensure that audio interactions remain clear, intuitive, and effective, regardless of environment or context. Learn more about our advanced offerings, innovative technologies, and the future of audio at www.infineon.com/mems.

References

- [1] Albert Mehrabian: *Silent Messages: Implicit Communication of Emotions and Attitudes* (2nd Edition); Belmont; Wadsworth; 1981; [Available online](#)
- [2] International Standards Organization: *ISO 226:2023; Acoustics — Normal equal-loudness-level contours* (Edition 3); 2023; [Available online](#)
- [3] International Standards Organization: *ISO 7029:2017; Acoustics — Statistical distribution of hearing thresholds related to age and gender* (Edition 3); 2017; [Available online](#)
- [4] Sesame AI Inc: Conversational voice demo; [Available online](#)
- [5] Gunar Lorenz, Infineon Technologies AG: *Speaker Diarization: The Evolution of AI Understanding with XENSIV™ MEMS microphones*, 2025-03-20; [Available online](#)
- [6] NVIDIA Corporation: *NVIDIA NeMo Framework User Guide – Speaker Diarization*; [Available online](#)
- [7] Infineon Technologies AG: *EVAL AHNB DIGITALV01 – Infineon Audiohub Nano*; [Available online](#)
- [8] Infineon Technologies AG: *IM72D128V – Ultra-low power digital XENSIV™ MEMS microphone*; [Available online](#)
- [9] Nora Houlman, Omdia (Informa TechTarget): *MEMS Microphone Report – 2024 Analysis*, 2024-10-18; [Available online](#)
- [10] Infineon Technologies AG: *MEMS microphones for consumer*; [Available online](#)
- [11] NVIDIA Corporation: *NVIDIA NeMo*; [Available online](#)

Published by
Infineon Technologies AG
Am Campeon 1-15, 85579 Neubiberg
Germany

© 2025 Infineon Technologies AG.
All rights reserved.

Public

Version: V1.0_EN
Date: 04/2025



Stay connected!



Scan the QR code and explore our offerings
www.infineon.com

Please note!

This Document is for information purposes only and any information given herein shall in no event be regarded as a warranty, guarantee or description of any functionality, conditions and/or quality of our products or any suitability for a particular purpose. With regard to the technical specifications of our products, we kindly ask you to refer to the relevant product data sheets provided by us. Our customers and their technical departments are required to evaluate the suitability of our products for the intended application.

We reserve the right to change this document and/or the information given herein at any time.

Additional information

For further information on technologies, our products, the application of our products, delivery terms and conditions and/or prices, please contact your nearest Infineon Technologies office (www.infineon.com).

Warnings

Due to technical requirements, our products may contain dangerous substances. For information on the types in question, please contact your nearest Infineon Technologies office.

Except as otherwise explicitly approved by us in a written document signed by authorized representatives of Infineon Technologies, our products may not be used in any life-endangering applications, including but not limited to medical, nuclear, military, life-critical or any other applications where a failure of the product or any consequences of the use thereof can result in personal injury.