

ARTICLE

High-performance MEMS microphones for conversational AI: unlocking new potential for voice-enabled assistants

Edition: v1.0
Date: 10/2023
www.infineon.com/mems

Table of contents

Abstract 3
1 Devices and applications 4
2 Trends shaping the future of conversational AI 5
3 Challenges to widespread adoption of voice-enabled assistants 6
4 How high SNR MEMS microphones enhance user experience 7
4.1 Improved speech recognition 7
4.2 Noise reduction and far-field voice pickup 7
4.3 Contextual understanding and multimodal interaction 8
5 Conclusion 9
6 High SNR MEMS microphone from Infineon 10
References 11

Abstract

Conversational AI is a rapidly-evolving field of machine learning that aims to make human-machine interactions intuitive and natural. It uses advanced algorithms and technology to interpret natural language input and enables machines to respond in a human-like manner. By integrating conversational AI frameworks into tools and systems, users can interact with machines using natural language commands. These intelligent systems are designed to understand intent and context, remember user preferences, and engage in meaningful conversations.

This article focuses primarily on conversational AI that interprets and responds to spoken words rather than written text, as voice-enabled applications gain prominence in our daily lives. We discuss some of the advancements that are shaping the growing market for conversational AI and challenges to the widespread adoption of voice-enabled assistants. One crucial factor in enhancing the user experience for voice-enabled applications is the development of voice user interfaces (VUI). To enable accurate speech recognition and improve the overall audio quality, high Signal-to-Noise Ratio (SNR) MEMS (Mirco-Electro-Mechanical Systems) microphones are emerging as a critical component.

These high-performance silicon microphones, with their compact size and high sensitivity, facilitate more precise sound capture, background noise filtering and ensure clearer audio input for conversational AI systems. In this article we examine how the integration of high SNR MEMS microphones in voice-enabled applications holds tremendous potential for enhancing speech recognition accuracy and enables more seamless and natural human-machine interactions.

1 Devices and applications

Conversational AI has become an integral part of numerous devices and applications available today, transforming the way we interact with technology in various settings. Here are some familiar applications that heavily rely on conversational AI technology.

- **Smart speakers** – A smart speaker is a type of standalone speaker with an integrated voice-enabled assistant that responds to user requests. Well-known speakers on the market include Google Home with the Google Assistant, Amazon Echo with Alexa, and Apple HomePod with Siri.
- **Voice-enabled car systems** – Cars with integrated voice-activated assistants help drivers keep their hands on the wheel and eyes on the road. Drivers can control music playback, navigation systems, and climate control without searching for buttons or navigating menus.
- **Smart home systems** – Smart home systems are a convenient way to operate household controls using natural language commands. Common devices that incorporate conversational AI include things like lighting, thermostats, and security systems.
- **Smart conference systems** – Smart conference systems are productivity tools that use conversational AI to transcribe and translate meetings. These systems often integrate voice-enabled assistants for administrative tasks, like scheduling, identifying action items, and compiling meeting minutes.

2 Trends shaping the future of conversational AI

The market for applications and devices with integrated conversational AI has increased rapidly over the last few years, with adoption accelerated by COVID-19. The voice assistant market is projected to grow at a CAGR (Compound Annual Growth Rate) of 33.5% between 2023 and 2030, a demand shaped by increasing efficiency and by advancements in conversational AI.[1] These are some of the trends today that drive the growth of this technology.

- **Improved speech recognition algorithms** – As conversational AI becomes more widespread, the dataset for speech recognition grows, meaning that speech recognition algorithms become better at recognizing words, phrases, and how they are spoken by real people. This also means that speech recognition technologies can better identify languages, accents, and dialects.[2]
- **Advancements in natural language processing** – Natural language processing is the mechanism by which conversational AI interprets what the user is asking for. The increased sophistication of natural language processing algorithms improves the accuracy and personalization of conversational AI, making it more intuitive and reliable.[3]
- **Increased use of speech-enabled devices** – The continued integration of speech-enabled operation for devices and applications serves to grow the demand for conversational AI, fueling further advancements in the field. As the technology improves, virtual assistants are expected to handle increasingly complex tasks and provide better output. The number of businesses that use voice-enabled applications is projected to rise as conversational AI continues to improve workplace efficiency.[4]

[1] Vantage Market Research. "Voice Assistants Market Size, Share & Trends Analysis Report by 2030". May 2023. Accessed 7 July 2023 from https://www.linkedin.com/pulse/voice-assistants-market-size-share-trends-analysis-report-hancock/

[2] Murf Resources. "Future of AI in Speech Recognition." April 2023. Accessed 18 June 2023 from https://murf.ai/resources/future-of-ai-in-speech-recognition/

[3] Schmelzer, Ronald. "Natural language processing drives conversational AI trends." *TechTarget.* June 2019. Accessed 18 June 2023 from https://www.techtarget.com/searchenterpriseai/feature/Natural-language-processing-drives-conversational-AI-trends

[4] GlobeNewswire. "Global Conversational AI Market Report 2023: Increasing Demand for AI-Powered Customer Support Services Boosts Growth." April 2023. Accessed 18 June 2023 from https://www.globenewswire.com/en/news-release/2023/04/17/2648259/28124/en/Global-Conversational-AI-Market-Report-2023-Increasing-Demand-for-AI-Powered-Customer-Support-Services-Boosts-Growth.html

3 Challenges to widespread adoption of voice-enabled assistants

Speech recognition and natural language processing technology is rapidly progressing, and there is clear market demand for advanced conversational AI systems. In spite of these advancements, users still encounter frustrations that may hinder the widespread use of voice-enabled assistants. Many of the challenges to the adoption of this technology are related to data privacy, with users concerned about the security of voice data stored in the cloud and the possibility that devices may record private conversations via passive listening.

Other frustrations may result when users interact with assistants. Voice-enabled assistants are integrated into nearly every new operating system and device, yet they notoriously confuse homophones, misunderstand accents, and require extremely precise pronunciation. These assistants struggle in environments with any background noise and often have trouble understanding users with speech disorders. These are speech recognition issues that may result from inferior onboard microphones in integrated devices.[5]

The voice user interface (VUI) is a critical component of conversational AI technology like voice-enabled assistants. A user interacts with an assistant by speaking to the VUI. An effective voice-enabled assistant, and therefore an effective VUI, must accurately hear and understand voice commands. A failure to understand the user can result in a limited and frustrating user experience.

[5] Zetlin, Minda. “Here’s Why Alexa (and Siri and Google) Still Don’t Understand You as Well as They Should”. Inc. December 2022. Accessed 19 June 2023 from https://www.inc.com/minda-zetlin/heres-why-alexa-and-siri-google-still-dont-understand-you-as-well-as-they-should.html

4 How high SNR MEMS microphones enhance user experience

While users might avoid some of the issues of misunderstood speech by speaking clearly and directly to voice-enabled assistants, avoiding noisy environments, and giving only simple commands, these practices limit the potential of conversational AI and defeat user expectations for natural, conversational interaction with voice-enabled assistants.

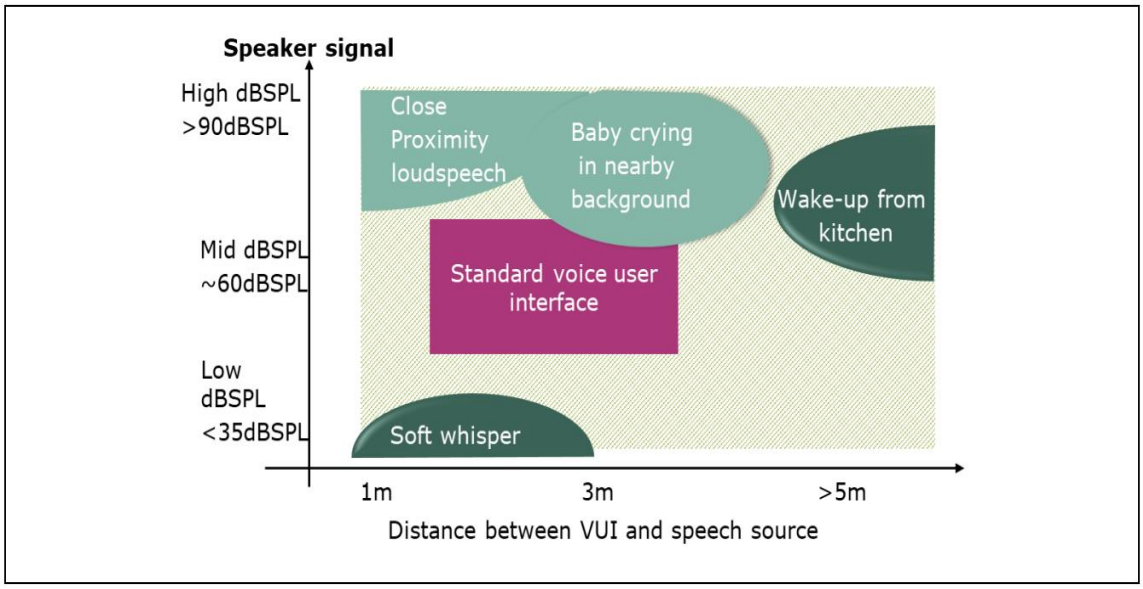
A demonstrated solution to this problem is to enhance audio capture at the VUI. High SNR MEMS microphones are designed to support the capture of clear audio in imperfect environments, and are effective for enhanced speech recognition, far-field voice pickup, contextual understanding, and multimodal systems that interpret both audio and visual input—keys to many of the challenges that impede the adoption of voice-enabled assistants.

4.1 Improved speech recognition

High SNR MEMS microphones capture clear, accurate audio signals, which sets the basis for improved performance of speech recognition algorithms. MEMS microphones capture voices through background noise, meaning that the voice-enabled assistant has better comprehension of user commands and queries. A microphone that provides a better quality input signal also improves the accuracy of the assistant's interpretation.[6] Because MEMS microphones are better equipped to handle the real-world sound environments where users will query voice-enabled assistants, these microphones can improve the overall user experience and efficiency of voice-based interactions.

4.2 Noise reduction and far-field voice pickup

A high SNR enables MEMS microphones to clearly capture voice commands. SNR refers to the difference between the desired sound that the microphone should pick up and the noise produced by the microphone itself, so a high SNR is able to capture more of the desired signal. The high SNR combined with high sensitivity enables far-field voice pickup, allowing users to interact with voice assistants from a distance or in noisy environments.[7]



Standard VUI use cases varying in speech signal levels and distance to the device | Source: Infineon

[6] Infineon. "Why you need high performance, ultra-high SNR MEMS microphones". Accessed 19 June 2023 from https://www.infineon.com/dgdl/Infineon-AN547_Why+you+need+high+performance+ultra-high+SNR+microphones+-AN-v01_01-EN.pdf?fileId=5546d4626102d35a01612d1e2afd6ad3

[7] Infineon. "Why you need high performance, ultra-high SNR MEMS microphones". Accessed 19 June 2023 from https://www.infineon.com/dgdl/Infineon-AN547_Why+you+need+high+performance+ultra-high+SNR+microphones+-AN-v01_01-EN.pdf?fileId=5546d4626102d35a01612d1e2afd6ad3

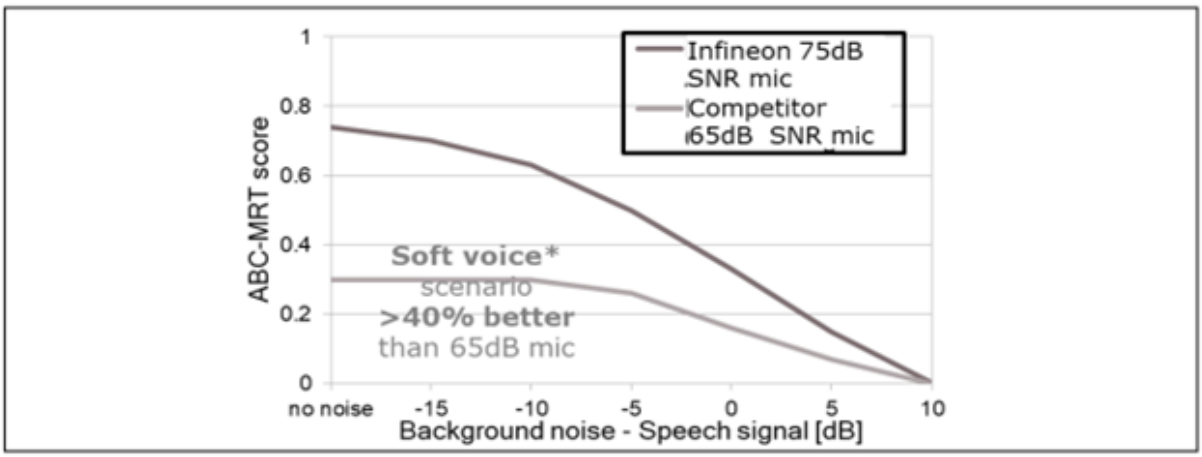


Chart showing improved performance score for high SNR microphone in a whisper/soft voice scenario | Source: Infineon

Active noise filtering and far-field voice pickup enhance the usability of voice assistants in various noisy scenarios like smart homes, conference rooms, customer support systems, and in public areas. A study performed by Infineon shows that high SNR MEMS microphones with 75dB SNR can capture audio 40% better than standard microphones, like the ones used by commercial voice-enabled assistants.[8]

4.3 Contextual understanding and multimodal interaction

VUIs with high SNR MEMS microphones also have the benefit of capturing context cues from the user's voice, such as tone and emphasis. This contextual understanding enables the voice assistant to infer user intent and provide more accurate and personalized responses.

This improved performance also opens the possibilities for multimodal interaction. For example, combining VUI and high SNR MEMS microphones with facial recognition models can enable users to interact with devices using both voice commands and facial expressions, further improving the voice-enabled assistant's understanding of the user's meaning.[9]

[8] Infineon. "Value of high-SNR microphones in Voice User Interface". Accessed 19 June 2023 from https://www.infineon.com/dgdl/Infineon-Value+of+high+SNR+microphones+in+Voice+user+Interface-ApplicationNotes-v01_01-EN.pdf?fileId=5546d46269e1c019016a78d976d852fd

[9] Ahmad, Majeed. "How MEMS Microphones Aid Sound Detection and Keyword Recognition in Voice-Activated Designs". *DigiKey*. Accessed 19 June 2023 from https://www.digikey.com/en/articles/how-mems-microphones-aid-sound-detection

5 Conclusion

High SNR MEMS microphones are essential for the improved effectiveness of conversational AI models used in VUIs. They enhance speech recognition accuracy, enable noise reduction and far-field voice pickup, support contextual understanding, and enable multimodal interactions. These microphones deliver clear audio capture by ensuring optimal performance even in noisy environments. High SNR MEMS microphones provide more reliable interactions with virtual assistants for a better user experience.

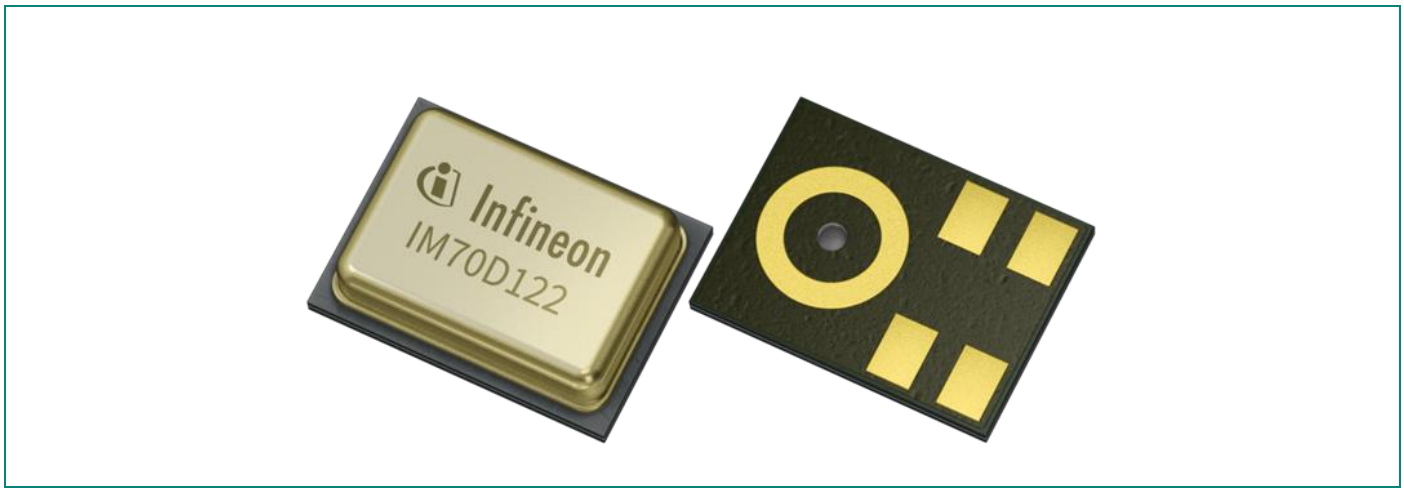
Furthermore, advancements in high SNR MEMS microphone technology hold great potential for the continued improvement and dependability of voice-enabled assistants. Ongoing developments in microphone sensitivity, signal processing, and noise cancellation techniques will further enhance the performance of conversational AI systems. As high SNR MEMS microphones continue to improve, we can expect significant advancements in human-machine interactions, unlocking new possibilities for voice-based technology.

The future of conversational AI is promising. Innovations in speech recognition, contextual awareness, and training models mean that voice-enabled assistants will be capable of handling more complex commands and conversations. Advanced algorithms coupled with superior microphones mean that users can look forward to a more comfortable and intuitive experience with voice-enabled assistants.

6 High SNR MEMS microphone from Infineon

Infineon's XENSIV™ MEMS microphones feature high SNR and low distortions even at high sound pressure levels, as well as part-to-part phase and sensitivity matching, flat frequency response with low frequency roll-off, and ultra-low group delay. Combined with selectable power modes and small package size, Infineon XENSIV™ MEMS microphones are a good match for devices with integrated conversational AI.

For more information on Infineon's best-in-class XENSIV™ MEMS microphones, we invite you to explore: www.infineon.com/mems



Infineon's high-performance digital XENSIV™ MEMS microphone (IM70D122) enables high-quality audio capture for laptop and tablet applications | Source: Infineon

References

[1] Vantage Market Research. “Voice Assistants Market Size, Share & Trends Analysis Report by 2030”. May 2023. Accessed 7 July 2023 from https://www.linkedin.com/pulse/voice-assistants-market-size-share-trends-analysis-report-hancock/

[2] Murf Resources. “Future of AI in Speech Recognition.” April 2023. Accessed 18 June 2023 from https://murf.ai/resources/future-of-ai-in-speech-recognition/

[3] Schmelzer, Ronald. “Natural language processing drives conversational AI trends.” *TechTarget.* June 2019. Accessed 18 June 2023 from https://www.techtarget.com/searchenterpriseai/feature/Natural-language-processing-drives-conversational-AI-trends

[4] GlobeNewswire. “Global Conversational AI Market Report 2023: Increasing Demand for AI-Powered Customer Support Services Boosts Growth.” April 2023. Accessed 18 June 2023 from https://www.globenewswire.com/en/news-release/2023/04/17/2648259/28124/en/Global-Conversational-AI-Market-Report-2023-Increasing-Demand-for-AI-Powered-Customer-Support-Services-Boosts-Growth.html…

[5] Zetlin, Minda. “Here’s Why Alexa (and Siri and Google) Still Don’t Understand You as Well as They Should”. Inc. December 2022. Accessed 19 June 2023 from https://www.inc.com/minda-zetlin/heres-why-alexa-and-siri-google-still-dont-understand-you-as-well-as-they-should.html

[6] Infineon. “Why you need high performance, ultra-high SNR MEMS microphones”. Accessed 19 June 2023 from https://www.infineon.com/dgdl/Infineon-AN547_Why+you+need+high+performance+ultra-high+SNR+microphones+-AN-v01_01-EN.pdf?fileId=5546d4626102d35a01612d1e2afd6ad3

[7] Infineon. “Why you need high performance, ultra-high SNR MEMS microphones”. Accessed 19 June 2023 from https://www.infineon.com/dgdl/Infineon-AN547_Why+you+need+high+performance+ultra-high+SNR+microphones+-AN-v01_01-EN.pdf?fileId=5546d4626102d35a01612d1e2afd6ad3

[8] Infineon. “Value of high-SNR microphones in Voice User Interface”. Accessed 19 June 2023 from https://www.infineon.com/dgdl/Infineon-Value+of+high+SNR+microphones+in+Voice+user+Interface-ApplicationNotes-v01_01-EN.pdf?fileId=5546d46269e1c019016a78d976d852fd

[9] Ahmad, Majeed. “How MEMS Microphones Aid Sound Detection and Keyword Recognition in Voice-Activated Designs”. *DigiKey.* Accessed 19 June 2023 from https://www.digikey.com/en/articles/how-mems-microphones-aid-sound-detection

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

© 2023 Infineon Technologies AG.
All rights reserved.

Public

Date: 10/2023

Stay connected!

Scan QR code and explore offering
www.infineon.com/mems

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.