

HOW KARDOME'S SPEECH ENHANCEMENT TECHNOLOGY SURPASSES IN-CAR SOUND QUALITY STANDARDS AT 120 KM/H

We collaborated with HEAD acoustics GmbH, a global leader in sound and vibration analysis, to study speech recognition and quality in a car, comparing traditional beamforming technology to Kardome's noise reduction and source separation algorithm.

THE CHALLENGE

Interfering signals, such as simultaneous speakers and background noise, negatively affect voice interaction performance

Voice interaction is one of the most critical functions performed by infotainment systems in cars. We use voice interactions to make hands-free phone calls and give commands to voice assistants to book restaurant reservations, play music, or get directions.

In-car use of voice interactions is a crucial driver of voice assistant adoption. As the fully connected car comes to fruition in the coming years, we may use voice interactions to get critical updates about our vehicle or even drive it.

Modern car infotainment systems use sophisticated microphones to focus on the desired speaker.

However, interfering signals, such as car radios, simultaneous speakers, and road noise still challenge state-of-the-art systems.



These noisy interferences deteriorate the performance of in-car voice assistants and reduce the quality and intelligibility of speech when using a car's hands-free telephone system.

To address these challenges, Kardome developed an innovative algorithm that spatially focuses on the desired speaker in acoustically challenging environments, such as vehicles.

KARDOME'S TECHNOLOGY

Kardome's AI-driven speech enhancement technology gives clear, real-time voice interaction and high-quality speech performance

Unlike other algorithms, which mainly use directional beamformers that cannot accurately capture individual speakers, Kardome's AI-driven speech enhancement technology gives clear, real-time voice interaction and high-quality speech performance.

This enhancement focuses the sound capture on the desired source's location rather than the source's direction.

Kardome's software also reduces background noise and interfering speakers, while accurately capturing and processing the desired speech.

This Solution Brief details a study conducted in collaboration with HEAD Acoustics GmbH, a global leader in sound, vibration, audio, and speech quality analysis.

The study analyzed speech recognition and quality in a car, comparing traditional beamforming technology to Kardome's noise reduction and source separation algorithm.



ANALYZING IN-CAR SPEECH RECOGNITION RATE

The first phase of the study compares the speech recognition rate (SRR) obtained by the Google Speech to Text (GSST) engine in a car traveling at 120 km/h with two types of speech processing systems:

The first study results show that Kardome's technology can improve speech recognition for voice interaction devices in cars in any type of sound environment.

Standard Hands-Free Telephony (HFT) audio stack, and Kardome's AI-driven signal separation and noise reduction technology packed in its [Gavel Evaluation Kit](#).

We compared Speech Recognition Rate using Google's speech-to-text engine in a car traveling at 120 km/h using:



The existing, standard hands free telephony system



Kardome's AI-driven signal separation and noise reduction technology packed in its Gavel Evaluation Kit.

SRR Evaluation Setup

We installed Kardome's Gavel Evaluation Kit in the overhead compartment of a Renault Megane GRAND Tour, next to the standard HFT system's microphones.

Next, we installed four Artificial Head Measurement Systems in the car to deliver speech through a full-band-capable artificial mouth.



Fig 1: Evaluation setup with four Artificial Head Measurement Systems and Kardome's Gavel Evaluation Kit.

SRR Evaluation Setup (cont.)

We considered three scenarios:

- Sole speech by the driver
- Two simultaneous speakers, driver and co-driver
- Four simultaneous speakers

In all the scenarios, the existing in-car HFT system and Kardome's Gavel Evaluation Kit recorded the speech signals. HEAD acoustics engineers tested the speech recognition rate by comparing the GSTT output text to the driver's actual speech.



RESULTS

Fig. 2 shows the SRR performance for each of the considered scenarios. Each bin in the plot represents the SRR obtained by the two signal processing methods, HFT and Kardome, in each of the three strategies. The green bins correspond to the HFT and the blue bins to Kardome's speech clustering system.

■ Kardome ■ Existing Hands-free Telephony

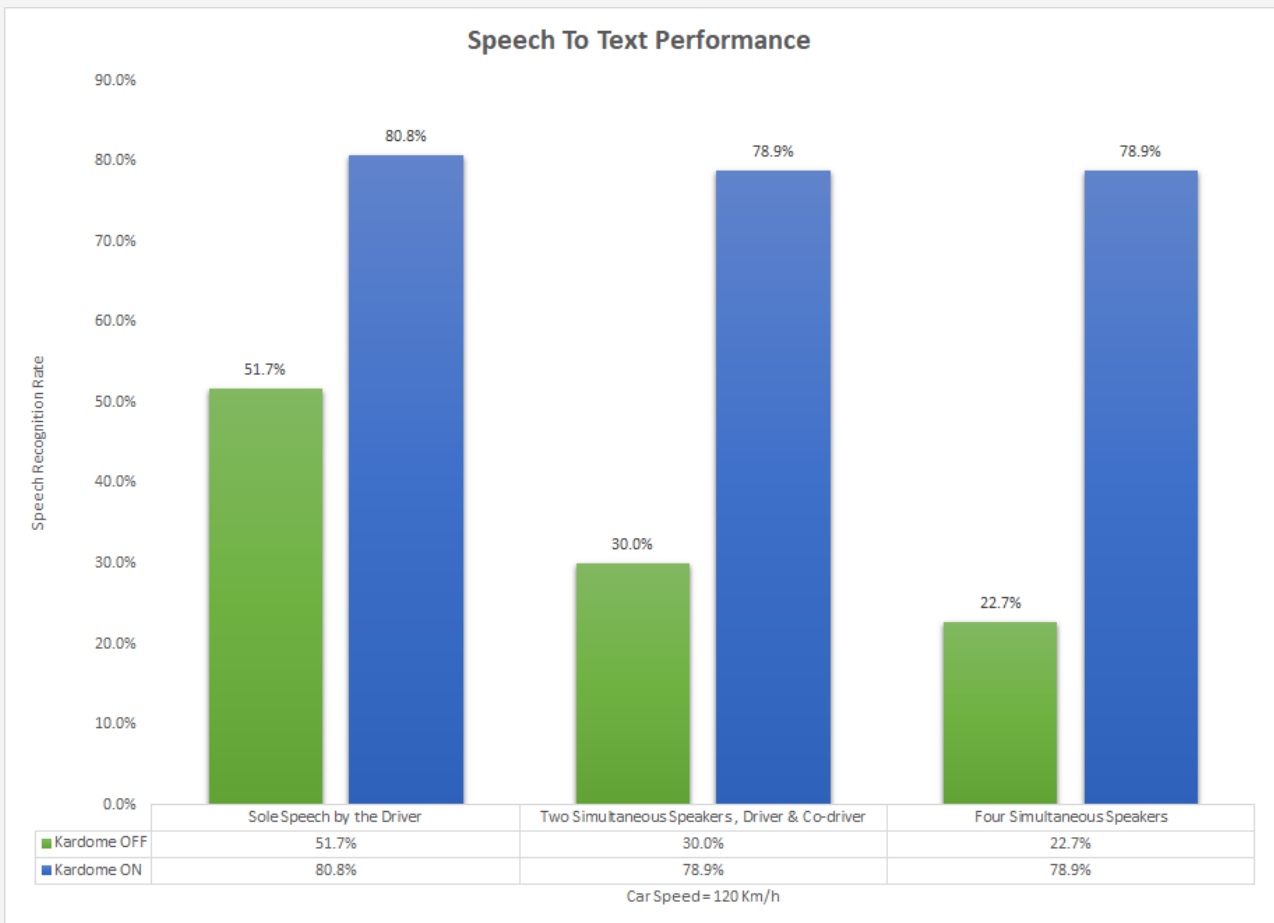


Fig 2: GSTT Speech Recognition Rates when using Kardome and standard HFT systems in the car.

RESULTS

Kardome's AI-driven signal separation and noise reduction technology significantly improved the SRR in all considered scenarios

Interfering speech signals significantly degraded the existing hands-free telephony system results as the number of speakers increased, whereas Kardome received a consistent SRR performance despite the number of speakers and interfering signals.



ANALYZING IN-CAR SPEECH QUALITY

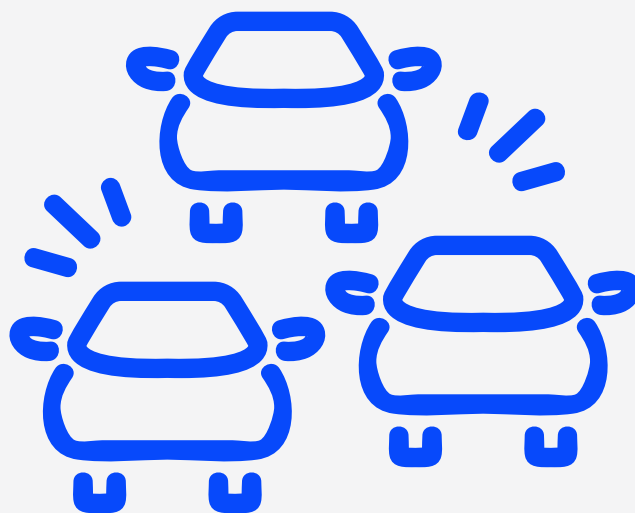
Kardome's Mallet™-based Voice User Interface surpassed the standard in all scenarios

Speech quality in many infotainment systems is acceptable, and the vast majority, if not all, comply with the International Telecommunication Union (ITU) P.1100 standard.

However, interfering speakers and background noise decimate the quality of speech, as received through the hands-free telephony system of a car.

HEAD Acoustics analyzed Kardome's impact on hands-free telephony speech quality, focusing on six scenarios.

The results revealed that the car's existing infotainment system failed to pass the standard P.1100 requirement in all simultaneous speech scenarios, while Kardome's Mallet™-based Voice User Interface surpassed the standard in all scenarios.



Background noise drastically reduces the quality of speech.

IMPLEMENTATION

Traditional speech quality tests suggested by the [ITU-T recommendation P.835](#) require a subjective determination of the mean opinion score (MOS) of the speech, noise, and overall quality of the sound.

However, in this study, we collaborated with [HEAD Acoustics GmbH](#) to evaluate the speech quality in a car using the 3Quest model (3-Fold Quality Evaluation of Speech in Telecommunications) introduced by them in the ETSI standard EG 202 396-3.

This study is an objective measurement procedure, which highly correlates with the subjective MOS.

The suggested procedure uses three scales to evaluate speech quality: speech mean opinion score (S-MOS), noise mean opinion score (N-MOS), and global mean opinion score (G-MOS).

Determination of subjective speech MOS (S-MOS)	Determination of subjective noise MOS (N-NOS)	Determination of subjective global MOS (G-MOS)
<p>Attending ONLY to the SPEECH SIGNAL, select the category which best describes the sample you just heard.</p> <p>The SPEECH SIGNAL in this sample was</p> <p>5 – NOT DISTORTED 4 – SLIGHTLY DISTORTED 3 – SOMEWHAT DISTORTED 2 – FAIRLY DISTORTED 1 – VERY DISTORTED</p>	<p>Attending ONLY to the BACKGROUND, select the category which best describes the sample you just heard.</p> <p>The BACKGROUND in this sample was</p> <p>5 – NOT NOTICEABLE 4 – SLIGHTLY NOTICEABLE 3 – NOTICEABLE BUT NOT INTRUSIVE 2 – SOMEWHAT INTRUSIVE 1 – VERY INTRUSIVE</p>	<p>Select the category which best describes the sample you just heard for purposes of everyday speech communication.</p> <p>The OVERALL SPEECH SAMPLE was</p> <p>5 – EXCELLENT 4 – GOOD 3 – FAIR 2 – POOR 1 – BAD</p>

3Quest - Instructions and cases [Application Note 3QUEST]

IMPLEMENTATION

Kardome's Voice User Interface was installed in a Renault MEGANE Grande Tour

We installed Kardome's Mallet™-based Voice User Interface in the overhead console of a Renault MEGANE Grand Tour, next to the car's standard infotainment system's microphones.

Next, HEAD Acoustic engineers installed four Artificial Head Measurement Systems in the car to deliver speech through a full-band-capable artificial mouth.



IMPLEMENTATION

HEAD acoustics engineers evaluated the 3Quest scores at each of the six scenarios

We considered six scenarios: Sole speech by the driver and simultaneous speech by the driver and the passenger, both when the car was parked, driving at 60 km/h, and traveling at 120 km/h.

The engineers of HEAD acoustics evaluated the 3Quest scores at each of the six scenarios.

The second phase of this study considers six scenarios:



Sole speech by the driver:

- 1) When parked
- 2) Driving at 60 km/h
- 3) Driving at 120 km/h



Simultaneous speech by the driver and a passenger:

- 4) When parked
- 5) Driving at 60 km/h
- 6) Driving at 120 km/h

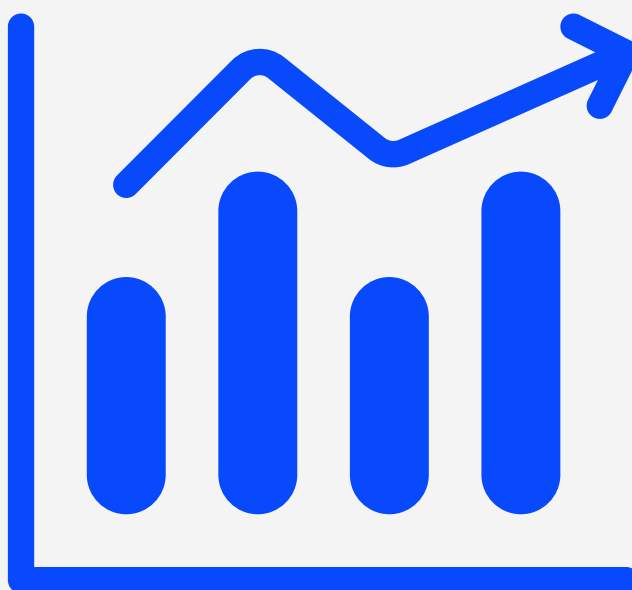
THE RESULTS

Using Kardome's Mallet™-based Voice User Interface, the car infotainment system surpassed the P. 1100 standard in all scenarios

The charts on the following page show the results. As shown, the existing infotainment system of the car failed to pass the P.1100 requirement in all simultaneous speech scenarios.

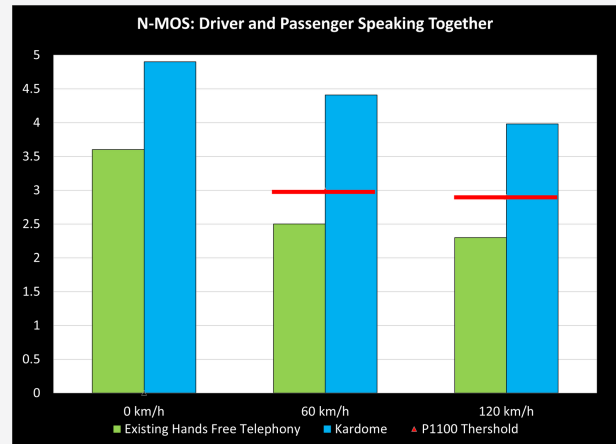
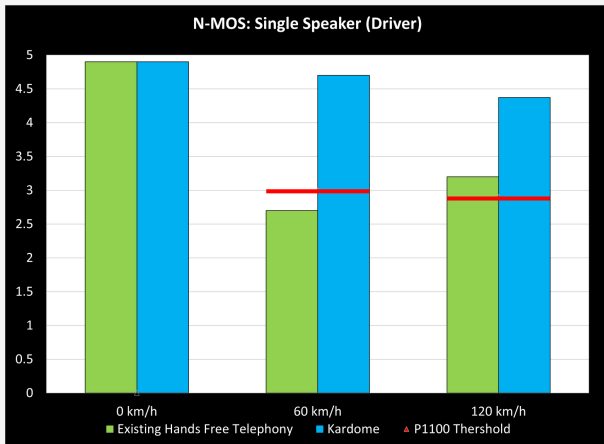
When a sole speaker is present, the infotainment passed the 120 km/h requirement but failed the 60 km/h.

Using Kardome's Mallet™-based Voice User Interface, the car infotainment system surpassed the P. 1100 standard in all scenarios.

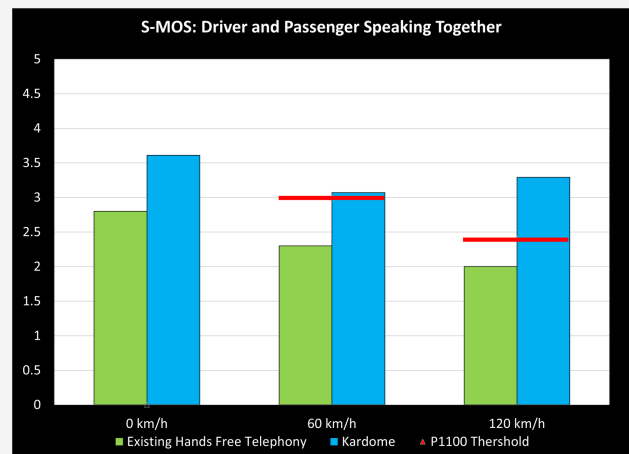
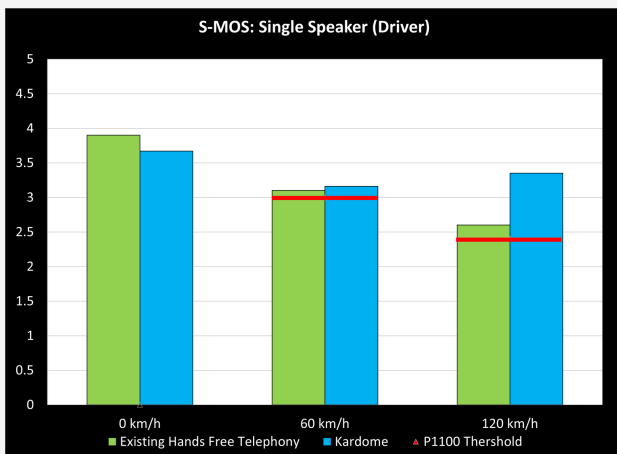


THE RESULTS

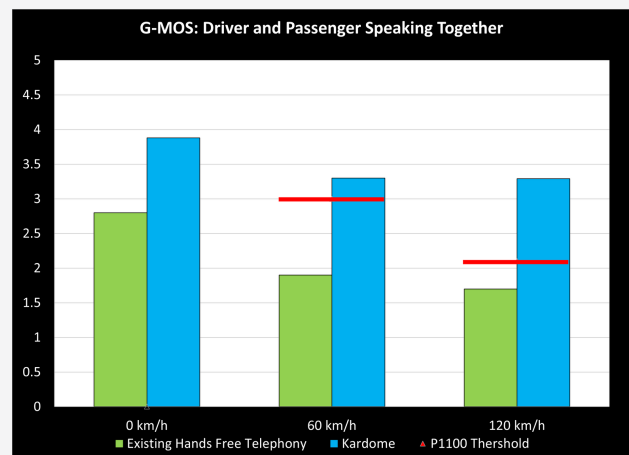
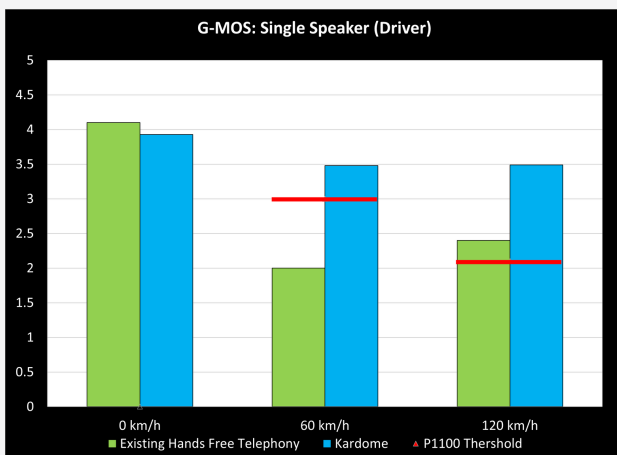
■ Kardome
 ■ Existing Hands-free Telephony
 ■ P. 1100 standard



Background Noise



Speech Signal



Overall Impression

MOVING TOWARD THE FUTURE

Car manufacturers must take advantage of voice and audio technology to position themselves as leaders in the connected car revolution

Existing, state-of-the-art voice interaction solutions cannot effectively filter and reduce background noise and interfering speakers in cars, reducing safety, customer satisfaction, and trust.

However, as this study shows, Kardome's AI-driven speech enhancement technology vastly improves speech and voice recognition quality in cars.

Car manufacturers must take advantage of voice and audio technology to position themselves as leaders in the connected car revolution.

Improving voice recognition technology in cars using Kardome's technology will help car and voice assistant manufacturers overcome inefficiencies in speech interaction.

It will also help automakers compete effectively in what will soon become a crowded marketplace for smart car voice interaction features.



NEXT STEPS

Improve Voice Interaction in Your Cars

Talk to us about how Kardome can help your company launch new customer experiences at scale, improve loyalty and retention, and prove ROI within months.

In an initial meeting with you, we might discuss:

- How speech/voice recognition fits into your existing car infotainment system
- Developing a customized, branded voice assistant
- Create integrated or branded Wake Words that work
- Developing a fully integrated voice automated connected car system

Request a demo, and we'll be in touch shortly

[Request a Demo](#)

ABOUT KARDOME

Cut Through the Noise with Kardome

Voice user interface design technology has the immense potential to open a doorway to a new age in the way humans interact with machines.

We founded Kardome to change the way people interact with machines and explore the true potential of speech recognition. Our goal is to provide a solution to user frustrations with speech recognition and voice command devices.

Kardome is committed to enabling the fundamental shift toward voice-based interaction by upholding the **Number 1 Rule of Speech Recognition: *Listen to the User, Not the Background Noise.***

Israel Office

Kardome Inc.
Shevah St 7, Tel Aviv,
Israel

Talk to Us

Email us at
Info@kardome.com

www.kardome.com

