Y. Du, "Speech Intelligibility and Quality Evaluation of Automotive Microphones Using Different Test Metrics and Their Correlation"
J. Audio Eng. Soc., vol. 71, no. 12, pp. 887–899, (2023 December).
DOI: https://doi.org/10.17743/jaes.2022.0095

Speech Intelligibility and Quality Evaluation of Automotive Microphones Using Different Test Metrics and Their Correlation

YU DU, *AES Member*
(yu.du@harman.com)

Harman International, Novi, Michigan

Speech intelligibility and speech quality (SI&SQ) of voice microphones used in automotive hands-free communication systems are affected not only by microphone acoustic characteristics but also by interactions between the microphone and the background acoustic field. Due to the complex acoustic environment inside a vehicle cabin which constantly changes with driving modes, it is not a trivial task to choose proper acoustic characteristics of voice microphones for different vehicle and cabin designs. To establish a relationship between microphone characteristics and SI&SQ performance in automotive applications, a study is conducted using three common types of automotive voice microphones. Their performance is evaluated using both subjective and objective metrics described in ANSI standards S3.2-2009, S3.5-1997 and ITU-T Recommendation P.862.2. It is found that the objective SI index (SII) and the subjective SI results correlate nearly linearly. Furthermore, when the original SII scores are weighted by the speech-to-noise ratio, the weighted SII (wSII) data also exhibit linear correlation relationship with the objectively calculated subjective mean opinion score. Because the SII/wSII calculation is significantly less complex than the SI or mean opinion score evaluation process, results from this study demonstrate that the SI/wSII may be conveniently used as a tool to guide automotive voice microphone designs and evaluations.

0 INTRODUCTION

Ever since its inception over a century ago, microphone has found uses in numerous sound-sensing applications, among which speech communication is arguably the most popular one. In an early publication, French and Steinberg gave a thorough theoretical and experimental analysis on factors governing the intelligibility of speech sounds [1]. Their results on fundamental characteristics of speech and hearing have helped guide communication system development for decades. Meanwhile, understanding and evaluating the effectiveness of speech communication under various application scenarios have remained to be popular topics in the literature [2–8].

Driven by the success of the cellular telephone industry and the wide adoption of smartphones, microphones, and particularly microphones targeted at voice/speech applications have become an integral part of modern automotive communication systems [9]. Because of physical constraints of vehicle cabin designs, automotive voice microphones are typically installed at locations around the headliner, sun visor, or steering column and are spatially separated from the talker by a significant distance. This arrangement together with safety considerations require automotive communication systems to be operated in a hands-free (HF) style so that both hands of the driver can be kept on the steering wheel. Consequently, it also means that the performance of the HF communication system is significantly affected by the acoustic environment of the vehicle, which makes the design and evaluation of the speech intelligibility (SI) and speech quality (SQ) of HF communication systems a challenging task.

Although sometimes used interchangeably in the literature, there are some distinct differences between SI and SQ. SI emphasizes on the aspect of speech understanding by listeners [1], and SQ represents the degree of general satisfaction perceived or judged by listeners based on the overall quality of the speech sound delivered through the communication system [10].

When accessing speech communication systems, SQ may be used as a more comprehensive performance measure because it includes the naturalness and intelligibility of speech sounds as well as the effort required to comprehend the speech. However, depending on specific automo-

tive applications, SI and SQ may be used individually or jointly. For example, for emergency calls whose focus is the intelligibility/clarity of the speech communication but not so much about the naturalness of the sound, the system design may give priorities to capture and transmit high-frequency rather than low-frequency components with a goal of improving intelligibility but may come at a cost of degraded sound naturalness. On the other hand, to achieve high-quality HF phone call experience, in addition to good intelligibility, it is also highly desirable to preserve the full band including the low-frequency content of the speech and minimize the intrusiveness of the background noise [11].

The effectiveness of speech communication (i.e., SI/SQ) can be assessed in many ways both subjectively and objectively. Compared to objective evaluation methods that are based on mathematical and acoustic models, conducting subjective evaluations is usually much more tedious, sometimes even practically prohibitive, because the use of human participants is required. Nevertheless, results obtained from objective methods should be correlated with and representative of those obtained from subjective methods of similar test scopes. Ultimately, the SI/SQ describes the effectiveness of the speech sound perception by human subjects.

A widely used objective method for the SI evaluation is the articulation index (AI) as described in ANSI S3.5, which has been revised and renamed to SI index (SII) since 1997 [12, 13]. The primary factor considered in the AI/SII calculation is the speech-to-noise ratio (SNR). Ebbitt used AI/SII to compare sound package designs for a passenger sedan and showed a nearly linear correlation between the AI/SII scores under driving noises represented by the engine speed at 2,000 RPM and above [14]. At low driving noise levels (i.e., low engine RPM and high SNRs), the AI/SII scores approached a theoretical limit of 1, thus losing the sensitivity for comparing engineering designs under such situations. Existing studies showed that a 0-dB SNR indicates an AI/SII score around 0.6, and the AI/SII scores approach 1 when the SNR is more than 15–20 dB [15].

Another popular objective SI evaluation method is the speech transmission index (STI) described in the standard IEC 60268-16 [16]. In addition to the SNR that is considered as the main contributing factor in the AI/SII calculation, STI also considers influences of the surrounding acoustic environment within which the talkers and listeners are located, including reverberations and echoes because they are all factors contributing to SI [1]. Because of this, STI is relatively more commonly adopted in the automotive industry to evaluate communication quality inside the cabin, such as the in-car communication between the driver and passengers seating in different rows [17]. However, in vehicles where echoes and reverberation are not significant, STI results had been shown to be very similar to AI/SII results. One benefit that comes with the AI/SII method is that its calculation process is simpler than that of the STI process [17].

For SQ evaluation, a subjective method generating a rating for the listening SQ on a five-point scale of mean opinion score (MOS) is established and described in several international standards [18–21]. Although, by definition, MOS shall be obtained with subjective testing, it may also be simulated or calculated based on objective measurement data combined with cognitive hearing models. In one study on in-car communication systems conducted by HEAD acoustics, objectively calculated MOS results were compared to subjective MOS data and good correlation between them was found at mid to high MOS values [22].

A complete communication system often starts with a microphone picking up speech sounds produced by a talker, proceeds with signal processing circuitry and software cleaning up the speech before sending it out through wired or wireless networks, and ends with a receiver (e.g., loudspeaker) playing the speech back to a listener. The effectiveness of the speech communication is obviously affected by all components involved in the complete communication system/chain. Therefore, for practical design and integration purposes, it is necessary to understand not only the overall SI/SQ performance of the entire system, but also the effects of design parameters of each key component (e.g., microphone) in the entire communication system/chain. Using the objective SII method, Du et al. studied and compared different automotive microphone designs with various frequency response shapes and directivities [23]. As a continued effort, this work extends the previous study by including both objective and subjective SI/SQ evaluation results for three automotive microphones and their correlation.

Sec. 1 below gives a brief review of three SI/SQ test methods adopted in this study. Sec. 2 describes the experimental design and how test signals were prepared and evaluated following standard procedures required by the three test methods. Experimental data are summarized in Sec. 3, followed by a correlation analysis in Sec. 4. It is shown that an approximately linear relationship between objective SII and subjective SI data can be established directly. In contrast, a highly nonlinear correlation is observed between the objective SII evaluation that focuses mainly on the intelligibility of the speech sound and the objectively calculated subjective MOS data that represents the overall speech sound quality. A weighted SII (wSII) calculation method is proposed so that an approximately linear correlation between the objective measurement based SII/wSII data and the objectively calculated MOS results can also be obtained. It is concluded in Sec. 5 that because the wSII method inherits the advantage of the SII method that has the least amount of measurement and computation costs among all three SI/SQ test metrics, it may be conveniently used to evaluate automotive voice microphone designs in the product development stage and optimize SI/SQ performance in actual application scenarios.

1 REVIEW OF SI/SQ TEST METHODS

Three SI/SQ test metrics were used in this study. They were ANSI S3.2-2009 for subjective intelligibility evaluation, ANSI S3.5-1997 for objective intelligibility evaluation, and ITU-P Recommendation P.862.2 for objective perceptual evaluation of SQ (PESQ). Note that, although

PESQ is described as an objective method, it is designed to simulate or predict the subjective MOS results obtained with human subjects. For clarity, a brief description of these three metrics is given in this section. A more comprehensive review of different SI/SQ evaluation methods can be found in reference [10].

1.1 Subjective SI Method

ANSI S3.2-2009 describes a subjective method for measuring the intelligibility of speech over communication systems using trained human talkers and listeners [24]. A minimum of five talkers and five listeners with normal hearing are required. In this study, the modified rhyme test (MRT) method described in ANSI S3.2-2009 is followed. The talker reads a list of 50 words with each word randomly chosen from a six-word list of monosyllabic English words. The talker's speech is delivered to the listener over the communication system under test. For each word read by the talker, the listener is given six word options from which to choose the correct one. The results from all listeners are summarized to calculate an SI score, normally expressed in percentage, as given in Eq. (1) [24].

$$R_A = \left(R - \frac{W}{n-1}\right)/50 \times 100\%, \tag{1}$$

where R_A is the SI score adjusted for guessing, R is the number of correct words, W is the number of incorrect words, and n is the number of alternative choices for each word ($n = 6$ for MRT). Because ANSI S3.2-2009 focuses on the intelligibility, it may overlook certain quality factors of speech sounds such as listening effort and harshness/intrusiveness of background noise accompanying the speech.

1.2 Objective SII Method

A method of calculating a single index number based on objective instrumental measurements is provided in ANSI S3.5-1997 [12]. This index number intends to be an indication of the SI performance; thus, it is called SII. Note that in the earlier version of the standard ANSI S3.5-1969, the term AI was used instead of SII [13]. To derive an SII score, it only requires measurements of the speech spectrum and the ambient noise spectrum. Because the speech level can vary significantly under different use cases and from talker to talker, reference speech levels are provided in the standard and can be used in lieu of measured speech. This makes the SII implementation very straightforward in practice. ANSI S3.5-1997 provides four ways to compute SII depending on the choice of number and size of the frequency bands. The one-third octave band method is used in this study. Eq. (2) gives the expression for the SII calculation [12].

$$SII = \sum_{i=1}^{J} I_i A_i, \tag{2}$$

where I_i is the band importance function, A_i is the band audibility function determined by the equivalent SNR in each band, and J is the total number of frequency bands. For the one-third octave method, $J = 18$ ranging from 160 to 8,000 Hz. According to the standard, the SII score is limited to a value between 0 and 1 for equivalent SNRs between –15 and +15 dB (Eq. (12) in [12]). An equivalent SNR outside this range results in an SII score of either 0 or 1.

Table 1. MOS rating scales.

Rating	1	2	3	4	5
MOS	Bad	Poor	Fair	Good	Excellent

It should be noted that the calculated SII scores based on objective measurements cannot be directly converted to subjective SI scores. Instead, the SII should be used as an indication of the fraction of speech cues (out of 1) reaching the listener through the communication system. Nevertheless, a previous study showed that a 50% sentence recognition rate reported by human subjects seemed to correspond to an SII value of approximately 0.22 [25].

1.3 Objective PESQ Method

As an objective method described in the ITU-P Recommendation P.862, the calculation of PESQ only needs the clean speech signal and the noise degraded speech signal as model inputs. Thus, no human subjects need to be involved [18]. However, the PESQ method was developed with an intention to simulate and correlate with the subjective MOS test results. Therefore, compared to the SII method, the PESQ method is more computationally intensive because it analyzes the inputs through a cognitive model to generate an objectively predicted MOS score [Mean Opinion Score–Listening Quality Objective (MOS-LQO)] representing the listening quality that would otherwise be evaluated subjectively. As mentioned earlier, by design, MOS is a rating method used in subjective testing and is given on a five-point absolute category rating scale as shown in Table 1. The PESQ method can be used to evaluate both narrowband (300–3,100 Hz) and wideband (50–7,000 Hz) signals [18, 19]. The wideband version of the P.862.2 model gives MOS-LQO scores ranging from 1.0 to 4.6 [19].

According to the ITU-T Recommendation document, the objective PESQ method was developed and benchmarked to subjective MOS listening quality scores obtained using the process described in ITU-T Recommendation P.830 [21]. The average correlation coefficient between MOS-LQO results obtained from the PESQ process and subjective MOS results was 0.935 [18]. Because of the nature of the MOS rating process as stated in ITU-T P.830 and its close correlation with the objectively calculated MOS-LQO results, the PESQ method is considered to be a more comprehensive assessment criterion in this study because it reflects not only how intelligible the speech sound is, but also the overall listening comfort level of the sound.

2 EXPERIMENTAL DESIGNS

2.1 In-Vehicle Recording Setup

Three automotive microphones with different directivity and frequency response characteristics were chosen

Table 2. Key electroacoustic specifications of three automotive microphones under test.

	Omni-ANC	Omni-HF	Uni-HF
Sensitivity (dBV)	–13.0	–9.7	–9.4
SNR [dB(A)]	65	65	62
–3 dB cutoff (Hz)	<20	390	560

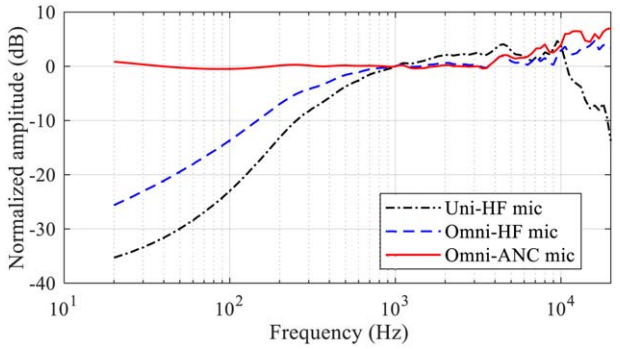


Fig. 1. Frequency responses of three automotive microphones under test, normalized at 1 kHz.

in the experiments. They were an omnidirectional microphone targeting active noise cancellation (ANC) applications (Omni-ANC), an omnidirectional microphone for HF voice applications (Omni-HF), and a unidirectional microphone for HF voice applications (Uni-HF). All three microphones used MEMS microphone elements in their designs. Their key electroacoustic specifications are compared in Table 2. Fig. 1 plots the frequency responses normalized at 1 kHz, which were obtained using the impulse response measurement in a semi-anechoic room. The main difference in the frequency response shape is that the Omni-ANC microphone has a relatively flat response from 20 Hz to 10 kHz covering the entire range that is typically considered for the wideband speech communication. In contrast, the Omni-HF and Uni-HF microphones have a rising response with gradually reduced sensitivity toward lower frequencies to help remove low-frequency noises in the background.

Because the focus of the current study is to assess the SI/SQ performance of automotive microphones with different directivity and frequency response, the microphones under test were separated from the rest of the communication system by removing the signal processing circuitry and additional noise reduction algorithms.

Fig. 2 shows the in-vehicle recording setup. The Omni-ANC, Omni-HF, and Uni-HF microphones were mounted on the driver-side sun visor together with a 0.5-in reference microphone (B&K Type 4190). Their sound port locations were kept as close as possible to ensure that they were exposed to substantially the same acoustic field. Therefore, for practical purposes, the influence that the physical proximity between the microphone bodies could have on the response of each of them was considered ignorable. The outputs from all four microphones were wired directly to and recorded simultaneously by a multi-channel data acquisition system. A mouth simulator (B&K Type 4227) was positioned in the driver's seat to play back speech signals.

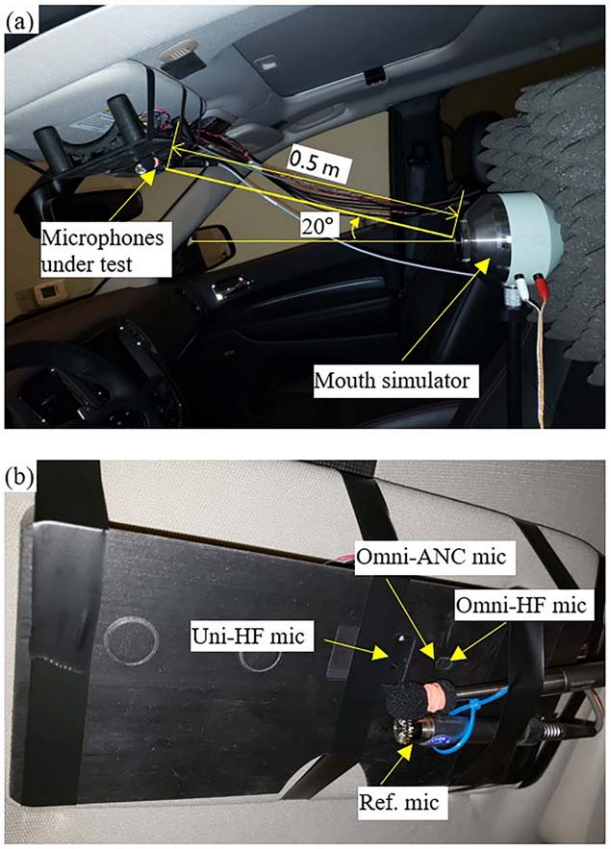


Fig. 2. In-vehicle recording setup (a) and mounting locations of the Omni-ANC, Omni-HF, Uni-HF, and reference microphones on the driver-side sun visor (b).

The orientation and mounting positions of the mouth simulator and test microphones with reference to each other resulted in a sound incident angle of about 20° coming from the mouth toward the microphones. The straight-line distance between the mouth lip-ring and the microphones was approximately 0.5 m.

The test vehicle was a full-size SUV with a 3.6-L V6 gasoline engine. Seven driving modes were considered to simulate various noise conditions commonly experienced by drivers and passengers. As summarized in Table 3, each driving mode was characterized by the driving speed, the ventilation fan setting, and the open/close status of the two front side windows (all other windows were closed). The driving noise characteristics were mainly affected by the driving speed and the window status. The defrost mode of the ventilation fan represented a highly directional noise condition as the airflow was directed to the windshield behind the microphones at the defrost setting. For stationary and idling test modes, the vehicle was parked in an indoor garage. For on-road moving test modes, the vehicle was driven on a local highway with asphalt pavement surfaces, located in the area of Novi, MI.

Note that driving mode 7 at a speed of 80 km/h with open windows while still having a HF phone call may be rare; however, it was included in this study as an extreme scenario as such a use case may practically occur, especially for people driving convertibles. Driving at such a

Table 3. Definition of driving modes used in the experiment.

Mode #	Speed	Fan setting	Window status
Mode 1	Stationary (engine off)	Off	Closed (WC)
Mode 2	Idling	Lowest defrost	Closed (WC)
Mode 3	Idling	Highest defrost	Closed (WC)
Mode 4	40 km/h	Low	Closed (WC)
Mode 5	80 km/h	Low	Closed (WC)
Mode 6	120 km/h	Low	Closed (WC)
Mode 7	80 km/h	Low	Fully open (WO)
Mode 8	Mode 7 + 8 dB increase in SNR to account for the "Lombard effect"		

relatively high speed with open windows means high background noise. When speaking in loud noise, speakers have a natural tendency to increase their vocal effort to enhance the audibility of their voice. This phenomenon is known as the "Lombard effect." The Lombard effect indicates an increase of 3 dB in the vocal effort for every 10-dB increase in the A-weighted background noise level exceeding 50 dB(A) with the maximum vocal effort increase clipped at 8 dB [26]. Mode 8 in Table 3 represents mode 7 with the Lombard effect.

2.2 Test Signal Preparation

Based on the input signals required by the SI/SQ test metrics, the clean speech (S_c), background noise (N), and noise degraded speech (S_n) need to be recorded through each microphone under test in the test vehicle. The original speech in the form of MRT words as described in ANSI S3.2-2009 was recorded with human talkers in a soundproof listening test room located in the Harman facility (Novi, MI) using the reference mic (B&K Type 4190). Following the standard process, a carrier phrase "mark the word" was added in front of each word to assist the listening test [24]. The total recording length of one MRT word list containing 50 words was approximately 4 min. The recorded original speech was then played back through the mouth simulator and recorded again through each microphone under test using the setup shown in Fig. 2 when the test vehicle was in the stationary condition (mode 1). This process generated the clean speech, S_c, being recorded through the Omni-ANC, Omni-HF and Uni-HF microphones.

In reference to the standard raised speech level of 68-dB sound pressure level (SPL) given in ANSI S3.5-1997 measured 1 m away from the lip-ring and considering the actual distance of 0.5 m between the mouth simulator and the microphones in the test setup, the mouth simulator playback volume was calibrated to be around 74-dB SPL measured by the reference microphone shown in Fig. 2(b) (i.e., a 6-dB boost due to the half distance).

Similarly, the background noise, N, was recorded through each microphone under test modes 1–7 without the speech, either in the garage or on the road. All recordings were saved in 16-bit wav format with a sampling rate of 48 kHz. Finally, the degraded speech, S_n, for each microphone under test was obtained by digitally mixing the clean speech recording and the noise recording together. Because all in-vehicle speech and noise recordings were made using the same multi-channel data acquisition system with identical gain settings, their relative levels were preserved during digital mixing.

2.3 SI/SQ Test Process

Among the three SI/SQ test metrics considered in this study, the subjective ANSI S3.2-2009 test process is the most complex and laborious one. Strictly following the ANSI S3.2-2009 standard, it requires a minimum of five trained talkers and listeners, and a minimum of three MRT lists per talker. A total of six talkers (three male, three female, ages 20–60) and five listeners (four male, one female, ages 20–50) participated in the study. Both talkers and listeners had English as their native language and were selected to be representative of typical drivers (i.e., automotive microphone users). Although audiograms were not formally tested, all subjects reported to have no known hearing impairments or difficulties in daily communication. Prior to the test, all subjects were trained to be familiar with the test procedure and the MRT words as required by the standard.

When recording the original speech containing the MRT words, talkers were directed and trained to maintain a relatively constant speech level that was comfortable to them. Depending on the talker, speech levels in the range of 65–75 dB A-weighted were measured using a handheld sound level meter positioned in front of the mouth adjacent to the recording reference microphone (B&K Type 4190). The soundproof room within which the original speech recordings were made had an A-weighted background noise floor lower than 35-dB SPL. This means that a minimum of 30-dB SNR was obtained by the original speech recordings. As described in Sec. 2.2, the original speech recordings were then used to generate the clean and degraded speech recordings (stored as wav files) for each microphone under test in the actual vehicle environment.

The recorded wav files were given to each listener so that they could play back through computers for the subsequent listening test. A computer program with a graphic user interface was used to record the listener's choices while the listener was listening to the playback of the MRT word lists. Considering a total of 10 h of listening materials generated for multiple microphones and test conditions, in-person listening test sessions involving five listeners were hard to manage. Therefore, after initial in-person training and practicing sessions on how to properly conduct the

listening test, listeners were asked to carry out the actual SI listening test at a place and time of their own choices. This allowed listeners to decide their own test schedule to avoid fatigue of long test sessions. With the above arrangement, all five listeners managed to complete the entire listening test from start to finish within a three-month period.

To remove potential impacts caused by variations in individual listening environment and listening device, each listener was directed to choose a quiet place, such as a closed-space meeting room at work, an office, or bedroom at home for the listening test. A brand new over-the-ear and closed-back headphone made by a major audio brand was given to each listener to be used as the playback device. This headphone helped achieve consistent listening environment across listeners as it further attenuated the background noise. The chosen headphone also had a neutral tonal balance feature, which was important to ensure that the frequency content of the recorded microphone signal would not be altered (at least not significantly) by the headphone during playback.

Because the ANSI S3.2-2009 standard does not require a specific listening level, listeners were directed to adjust the playback volume through the headphone to a level comfortable to each of them and was comparable to the volume they normally use when making a HF phone call in a vehicle. As a precaution of minimizing systematic biases, the order of the wav files given to each listener was randomized so that no two listeners conducted the listening tests for various test microphones and test modes in the same order.

To minimize the total number of listening trials, all six talkers and five listeners participated in a pilot test to compare the SI results obtained by following three MRT lists per talker (standard) versus one MRT list per talker. The pilot test was done with audio files prepared under test modes 2 and 3. It was reasonably doubted that effects of the number of MRT lists per talker may be buried if the SI scores were all high or all low corresponding to cases with no background and high background noises, respectively. Therefore, test modes 2 and 3 were chosen with the assumption that they were representative of typical automotive microphone use cases containing low to medium level background noises (e.g., engine idling noise, defrost fan noise).

Based on a statistical t test using SI scores of the Omni-ANC microphone, it was demonstrated that, for both test modes, no statistically significant difference was found in the SI scores between the three versus one list per talker arrangements ($\alpha = 0.05$ or 95% confidence level). The maximum difference in the calculated mean SI score was less than 0.5%. Therefore, the one (instead of three) MRT word list per talker approach was adopted in this study, which generated 30 individual SI scores for each test scenario (one microphone under one driving mode) using five listeners and six talkers. It should be pointed out that the pilot SI test result was simply used as a guideline to mitigate the amount of time and efforts required by the listening test within the scope of the current study, it was not intended to modify the standard SI test method of ANSI S3.2-2009.

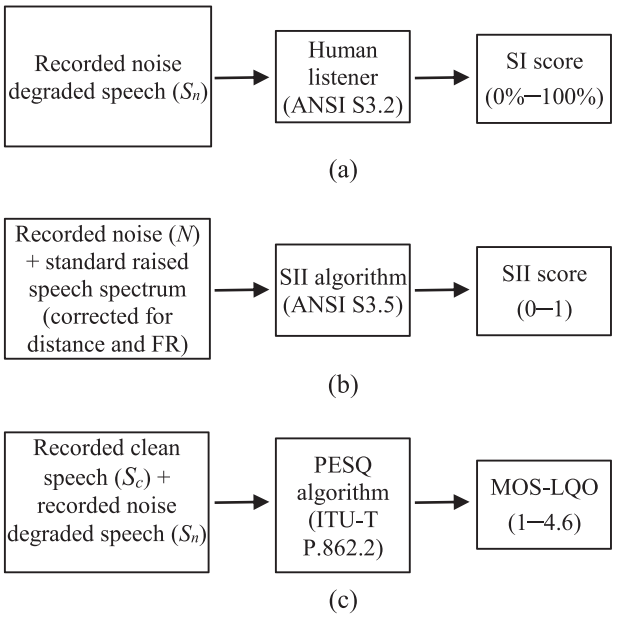


Fig. 3. Test processes, required signal inputs and generated outputs of (a) ANSI S3.2 SI evaluation, (b) ANSI S3.5 SII calculation, and (c) ITU-T P.862.2 MOS-LQO prediction.

The test processes for the objective PESQ and SII methods are rather straightforward. The PESQ prediction algorithm described in ITU-T Recommendation P.862 and P.862.2 requires the recorded clean speech and the noise degraded speech as inputs to calculate the MOS-LQO score. The objective SII calculation of the ANSI S3.5-1997 requires the recorded background noise spectrum and the clean speech spectrum to derive the SII value. To minimize the influences due to variations in talkers' speech levels, the ANSI S3.5-1997 method also allows the use of standard speech spectrum levels in the calculation. The standard spectrum level of 68-dB SPL representing the raised vocal effort was adopted in the SII calculation in this study. The standard raised speech was raised again by another 6 dB in each spectrum band to account for the shorter distance between the talker and microphone in the in-car experiment setup (0.5 vs. 1 m).

Furthermore, to account for the frequency response characteristics of the three microphones under test, the standard speech spectrum was corrected using the corresponding frequency response of each microphone given in Fig. 1. The test processes of the three SI/SQ methods, including their required inputs and generated outputs, are illustrated in Fig. 3.

3 EXPERIMENTAL RESULTS

3.1 Driving Noise Levels

The total background noise levels (20 Hz–20 kHz) of all driving modes 1–8 monitored by the reference microphone are plotted in Fig. 4. The non-weighted noise levels range from 44 dB (stationary) to slightly above 100 dB. Because of the gasoline engine noise emission characteristics, the total noise level is not always proportional to the driving speed.

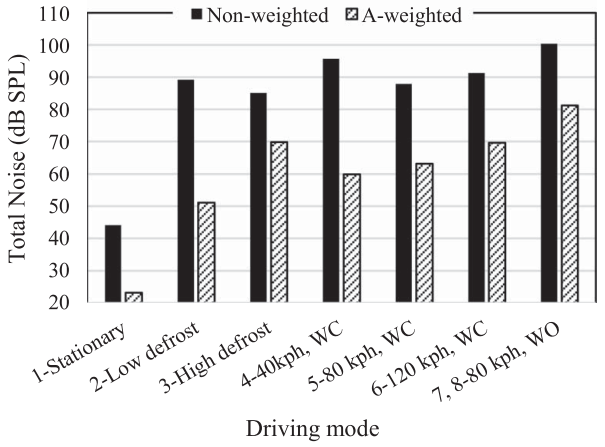


Fig. 4. Total non-weighed and A-weighted noise levels (20–20,000 Hz) of all driving modes monitored by the reference microphone using the experimental setup shown in Fig. 2.

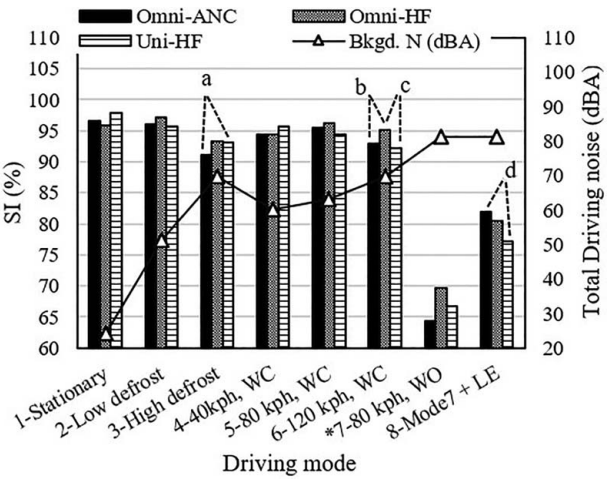


Fig. 5. Average SI scores obtained using the ANSI S3.2-2009 process for the Omni-ANC, Omni-HF, and Uni-HF microphones under various driving modes of a gasoline-engine SUV (* indicates data obtained with only one listener).

It seems that, with all windows being closed, the driving noise level is higher at 40-km/h speed than at 80 and 120 km/h. The same trend is also observed for the low defrost fan mode versus the high defrost fan mode. On the other hand, it is seen that the A-weighted total noise level is proportional to the driving speed and the defrost fan speed. Differences in the trends shown by the non-weighted and A-weighted noise levels indicate that the test vehicle with a gasoline engine had more low-frequency noise at low engine/fan speed than at higher speed. Among all driving modes, mode 7 (and mode 8 because they only differ in speech level due to the Lombard effect) of 80-km/h speed with two open front windows has the highest noise level, in both non-weighted and A-weighted ratings.

3.2 Subjective SI Test Results

Fig. 5 summarizes the average SI scores obtained using the ANSI S3.2-2009 process for the Omni-ANC, Omni-HF, and Uni-HF microphones under various driving modes. It should be noted that, because of high wind noise and low SNR when driving at 80 km/h with open front windows, the listeners reported that it was very uncomfortable to listen to the noise degraded speech for driving mode 7. The data point included in Fig. 5 corresponding to mode 7 was only obtained from one subject who managed to complete the evaluation. Consequently, mode 7 was not considered in the statistical analysis on the SI results.

In practice, talkers tend to increase the vocal effort in noisy environment owing to the Lombard effect, which is considered in the eighth driving mode by increasing the mode 7 SNR by an additional 8 dB [26]. For all three microphones and majority driving modes (excluding modes 7 and 8) considered in this study, the SI scores were at least 91% (Omni-ANC microphone under mode 3). At higher driving speed with open windows (mode 7), the resulted SI scores were the lowest (<70%). This was clearly attributed to the elevated wind noise induced by opening the two front windows. When considering the Lombard effect (mode 8), the SI scores were improved to above 75% for all three microphones because of the increased SNR.

According to existing literature results, an SI score at 75%, 91%, and 97% indicates the minimum, normal, and high intelligibility requirement for practical communication, respectively [10]. This seems to imply that, regardless the directivity and frequency response characteristics, all three common automotive microphone designs considered in this study can deliver normal to high intelligibility performance under typical driving scenarios when windows are closed. Even with excessive wind noise, the SI scores still met the minimum requirement of 75% when factoring the contribution from the naturally occurring Lombard effect. Obviously, this conclusion may be only valid when the focus is on the speech understanding part of the communication as the ANSI S3.2-2009 method does not necessarily consider factors like listening effort and intrusiveness of the background noise.

A closer examination on Fig. 5 further shows that SI scores of three microphones under the same driving mode present small differences. These differences are presumably associated with the microphone directivity and frequency response characteristics, which can be analyzed through a statistical study.

With six talkers and five listeners, the SI test resulted 30 individual SI scores for each test case (one microphone under one driving mode), which enabled SI performance comparison with statistical tools. A statistical *t*-test was performed between all combinations of any two out of three microphones under each driving mode with a 95% confidence level. As indicated in Fig. 5 and summarized by Table 4, among all non-stationary driving modes 2–6 and 8 (mode 7 was not included in the *t*-test) consisting of a total of 18 test cases, SI scores can only be stated statistically significantly different for the following four scenarios:

a. When the background noise level is highly directional and is relatively high [mode 3, 70 dB(A)], the unidirectional microphone resulted in 2% higher SI score than the Omni-directional ANC microphone.

Table 4. Summary of four test cases indicated in Fig. 5 under which the differences in SI scores are statistically significant with the 95% confidence level ($\alpha = 0.05$, r-crit = 2.0017).

	Driving mode and noise characteristics	Conclusion	$\lvert t \text{ stat} \rvert$ / p value
a.	Mode 3; high level, directional	Uni-HF > Omni-ANC	2.009 / 0.049
b.	Mode 6; high level, non–wind turbulence	Omni-HF > Omni-ANC	2.2369 / 0.029
c.	Mode 6; high level, non–wind turbulence	Omni-HF > Uni-HF	3.1819 / 0.002
d.	Mode 8; high level, wind turbulence	Omni-ANC > Uni-HF	2.2213 / 0.030

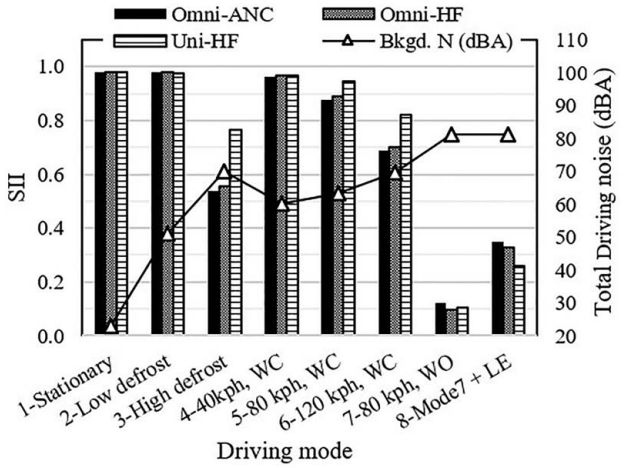


Fig. 6. SII scores obtained using the ANSI S3.5-1997 process for the Omni-ANC, Omni-HF, and Uni-HF microphones under various driving modes of a gasoline-engine SUV.

b. Under high driving noise that is not induced by wind turbulence [mode 6, 70 dB(A)], the omnidirectional HF microphone with a rising frequency response shape resulted in a better SI score than the unidirectional microphone.

c. Under high driving noise that is not induced by wind turbulence [mode 6, 70 dB(A)], the omnidirectional HF microphone with a rising frequency response shape also resulted in a better SI score than the omnidirectional ANC microphone with a flat frequency response shape.

d. When high wind turbulence noise presents [mode 8, 81 dB(A)], the unidirectional microphone has an SI score that is statistically significantly lower than the score obtained by the omnidirectional ANC microphone.

3.3 Objective SII Test Results

Fig. 6 plots the SII scores obtained using the ANSI S3.5-1997 process. At relatively low background noise levels (e.g., <60 dB(A) under modes 1, 2, and 4), the SII scores for all three microphones under test are nearly perfect (i.e., close to 1). Comparing modes 4, 5, and 6 that all have basically non–wind turbulence–induced driving noise (windows closed), it is seen that as the driving speed (and thus driving noise) increases, the SII scores reduce steadily for all three microphones. It is also seen that the unidirectional HF microphone performs the best compared to the other two microphone types. This is again clearly the case for mode 3 when high-level directional noise presents. Lastly, for the open window driving mode (mode 8) with high-level noise mostly induced by wind turbulence, the Omni-ANC microphone seems to give the best SII score, whereas the Uni-HF mic gives the worst. This observation is consistent with what is shown by the SI result in Fig. 5 for the same driving mode 8.

Note that SII scores for all three microphones under mode 7 are about 0.1 or lower because of high driving noise (i.e., low SNR). As previously mentioned in Sec. 1.2, existing literature results stated that an SII value of 0.22 indicated roughly 50% sentence recognition rate by human subjects [25]. This implies that all three microphones under test delivered nearly unintelligible speech under mode 7; thus, comparing the SII scores under mode 7 is less meaningful.

Comparing the subjective SI results in Fig. 5 to the objective SII results in Fig. 6, one trend that seems to be different is the best performing microphone under modes 5 and 6 that both had medium to high-level non–wind turbulence–induced driving noise (windows closed). The subjective SI data show that the omnidirectional HF microphone with a rising frequency response shape is better than either the Omni-ANC mic or the Uni-HF mic, and the difference is even statistically significant for mode 6. On the other hand, the objective SII data indicate that the unidirectional microphone (Uni-HF) has better performance over the other two types of microphones. It is also noted that the objective SII scores of all three microphones under both driving modes 5 and 6 are all about 0.7 or better (corresponding approximately to a subjective SI value of 97%) that is considered high intelligibility already [10]. Therefore, it can be concluded that the performance trends of the three automotive microphones represented by the subjective SI and the objective SII results are consistent overall.

3.4 Objective PESQ Test Results

As explained earlier, the SI and SII methods focus mainly on how intelligible the speech sound is. They may not necessarily capture the overall listening comfort level of the speech signal, which is certainly affected by factors such as sound naturalness and background noise intrusiveness (i.e., the level of annoyance). To study the overall listening quality, the MOS-LQO scores obtained using the ITU-T P.862.2 PESQ process are compared in Fig. 7. The ITU-T P.862.2 algorithm generates a score between 1 and 4.6. In general, a MOS-LQO score of 4.6 corresponds to perfectly intelligible speech and inaudible background noise; a score of 3 may mean good intelligibility and slightly noticeable noise; and a score of 1 indicates bad intelligibility and significantly

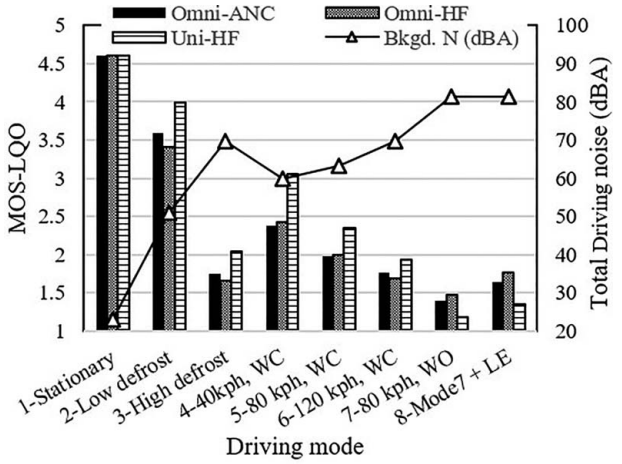


Fig. 7. MOS-LQO scores obtained using the ITU-T P.862.2 PESQ process for the Omni-ANC, Omni-HF, and Uni-HF microphones under various driving modes of a gasoline-engine SUV.

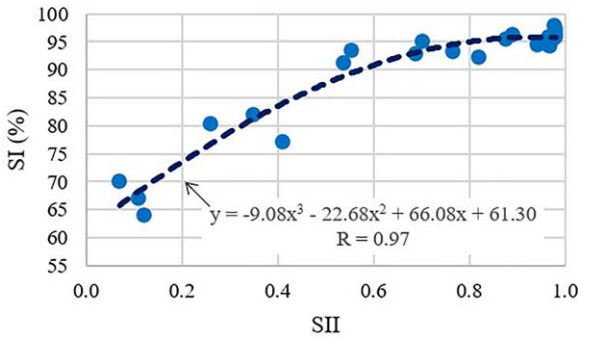


Fig. 8. Comparison of SII and SI scores with the third-order polynomial mapping.

annoying noise. Normally, a minimum MOS-LQO score of 2.5 would be required in practical applications.

It is seen from Fig. 7 that the MOS-LQO scores obtained from all three microphones are above 3 only at low noise levels (e.g., <60 dB(A) under modes 1 and 2). When the background noise level is around 60 dB(A) (mode 4), only the Uni-HF microphone has a score about 3, whereas the other two microphone scores are slightly below 2.5. For all other cases, the MOS-LQO scores are below 2.5, sometimes even well below 2 (mode 7). It should be restated that the experiment in this study was designed to only measure the microphone performance. In practice, automotive microphones are typically used with signal processing algorithms to help remove background noise from the speech. Therefore, the results in Fig. 7 may represent conservative performance estimates of automotive microphones in actual applications.

To this end, it would be interesting to compare the MOS-LQO results with the SI and SII results. For example, consistent with Figs. 5 and 6, the MOS-LQO data also demonstrate that the unidirectional microphone has the worst performance at high-level wind turbulence–induced noise (modes 7 and 8). On the other hand, for non-stationary driving modes with non–wind turbulence noise (modes 2–6), the unidirectional microphone seems to consistently perform better than the two omnidirectional microphones. This trend is not clearly shown by either the SI or SII results, which demonstrates the difference between the SI and SQ evaluation results and is presumably related with the noise factor in the speech sound.

4 CORRELATION ANALYSIS

The previous section presented SI/SQ evaluation results based on three different test metrics. These results were obtained in an actual vehicle environment for three common automotive microphone designs with different directivity and frequency response characteristics. Although some common trends in terms of how each type of microphone performs under each driving mode can be identified, a correlation analysis is needed to possibly allow one to predict performance on one measurement approach based on the data obtained from the other. Such a correlation, if demonstrated to be valid and reasonable, can be a handy tool for choosing and/or designing microphones for various automotive applications. For example, because the objective SII process (ANSI S3.5-1997) is relatively much easier to implement, an established correlation between SII and SI scores and/or between SII and MOS-LQO scores would make it possible to estimate the subjective SI and MOS performance without physically performing laborious SI and MOS tests.

4.1 Correlation Between SII and SI

In Fig. 8, the raw SII and SI scores for all test cases (three microphones and eight driving modes) are compared on one plot. The dotted line represents a third-order polynomial mapping between SII and SI scores with a Pearson correlation coefficient of 0.97. Direct comparison between SI and SII results are rare in the literature. In [27], Anderson and Kalb developed a relationship between the subjective SI scores based on the MRT test and the objective STI scores. The shape of their curve (Fig. 2 in [27]) closely resembles the dotted line in Fig. 8 below. Because it is known that the STI and SII methods would yield similar results in a vehicle environment where echoes and reverberation are not prominent [17], previous data from Anderson and Kalb cross verified the finding in the current study in terms of the SII and SI correlation.

A further examination shows that the SII-SI correlation can even be reasonably approximated using a simple linear relationship. Fig. 9 shows this linear relationship that has a correlation coefficient of 0.93 within the 95% confidence internals.

4.2 Correlation Between SII and MOS

Similarly, the scatter plot in Fig. 10 compares the SII and MOS-LQO data for all test cases. It is noted that the SII scores basically approach the theoretical upper limit of 1 when the MOS score is above 2.5. This means that the standard SII method does not have sufficient resolution to

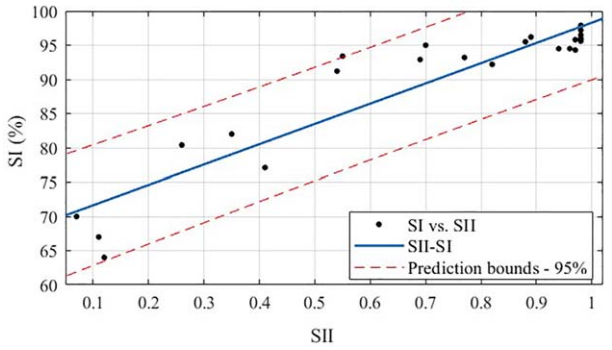


Fig. 9. Comparison of SII and SI scores with the linear mapping and 95% confidence intervals.

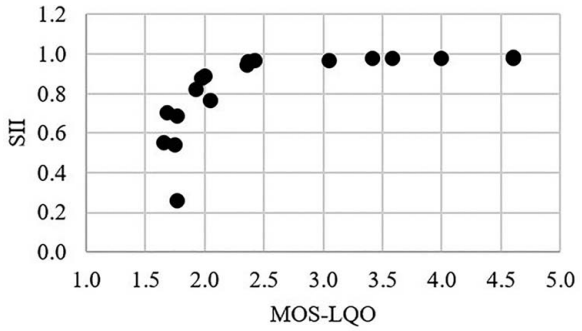


Fig. 10. Comparison of SII and MOS-LQO scores showing a nonlinear correlation.

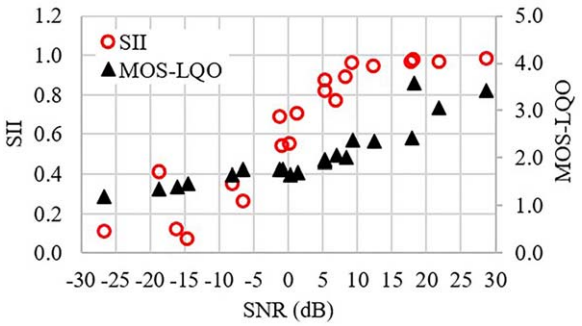


Fig. 11. Comparison of SII and MOS-LQO scores versus SNR.

predict the MOS score. In other words, a simple polynomial mapping between SII and MOS does not seem to exist.

Recall that the MOS-LQO results obtained from the PESQ process include factors like noise annoyance that are not necessarily captured by the SI and SII methods described in standards ANSI S3.2 and S3.5. This point may be further illustrated when plotting the SII and MOS-LQO data versus the SNR as shown in Fig. 11. The SII value changes rapidly when SNR is between –8 and +10 dB. Inspired by this observation, a linear correlation may be reconstructed between the MOS-LQO and wSII results by adding an SNR-dependent weighting function to the original SII data.

Referring to Fig. 11 again, to regain an approximately linear correlation, the goal is to reshape the original SII curve

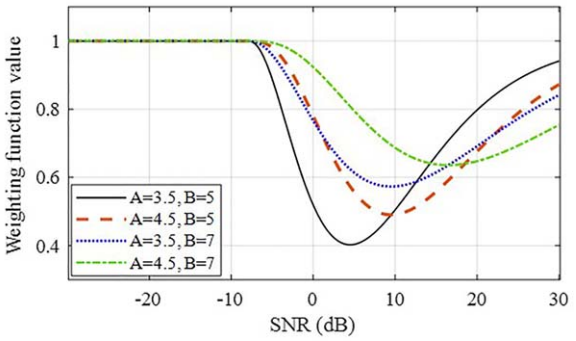


Fig. 12. Values of the weighting function calculated with $c = 12.24$, $s = 8$, and with various values for parameters A and B.

through certain mathematical weighting so that it closely matches the trend of the MOS-LQO curve. Following this concept, the weighting function should de-emphasize the original SII values in the SNR range centered around 5–10 dB and leave the SII values untouched when the SNR is higher or lower than approximately 20 or –10 dB, respectively. Obviously, this weighting function has to be a nonlinear function of SNR. Such a weighting function may be developed using the Gamma probability density function (PDF).

In Eq. (3), the weighting function $f(snr)$ is expressed as a function of the unweighted speech to noise ratio (snr). is the Gamma PDF with input variable snr and shape parameters A and B. Because Gamma PDF is 0 for negative input variable values, to account for negative SNRs that could practically occur as shown in Fig. 11, the input snr value is shifted by a positive s dB in the Gamma PDF. The output of the Gamma PDF is multiplied by a scalar factor c, then subtracted from 1. The mathematical expression of Eq. (3) ensures that values are always between 0 and 1, and its minimum value (i.e., maximum value of the Gamma PDF) occurs at the SNR where the original SII needs to be de-emphasized the most. The weighted SII value, wSII, can then be obtained by multiplying the weighting function to the original SII value as expressed in Eq. (4).

$$f(snr) = 1 - c * \Gamma(snr + s,\ A,\ B) \tag{3}$$

$$wSII(snr) = f(snr) * SII(snr) \tag{4}$$

From the derivation of the weighting function, effects of parameter s and c are self-explanatory. Parameters A and B, both non-negative, are parameters used in the standard Gamma PDF. They determine the shape of the Gamma PDF curve. Fig. 12 plots the weighting function values calculated using Eq. (3) with different A and B values but fixed values of s and c.

In Fig. 13, the wSII derived from Eqs. (3) and (4) using the set of parameter values of $A = 3.5$, $B = 5$, $c = 12.24$, and $s = 8$ is plotted against the corresponding MOS-LQO scores (scattered dots). The linear correlation prediction between the wSII and MOS-LQO scores is denoted by the solid line. It is seen that, after transforming the original SII to the weighted wSII, all data points fall within the 95%

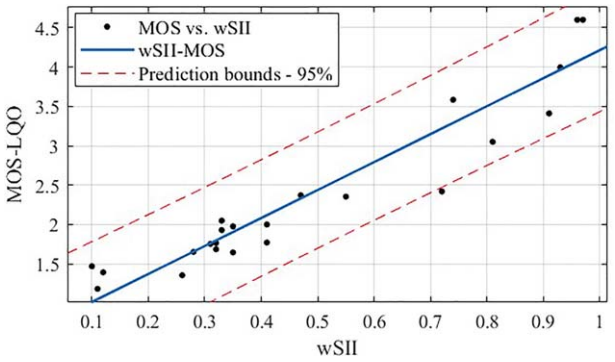


Fig. 13. Comparison of wSII and MOS-LQO scores with the linear mapping and 95% confidence intervals. Values of the weighting function calculated with $A = 3.5$, $B = 5$, $c = 12.24$, and $s = 8$.

confidence intervals of the linear correlation model. Results in Fig. 13 imply that the newly proposed wSII concept may be used as an effective indicator of the MOS score. This is similar to the correlation shown in Fig. 9 between the original SII and the SI data.

It is also noted that because parameters A and B values affect the shape of the weighting function, they can consequently affect how the derived wSII values correlate with MOS-LQO scores. Thus, a sensitivity analysis on weighting function parameter values should be conducted in the future. Furthermore, although the current study included eight driving modes that were considered representative of common automotive voice microphone use scenarios, it was also limited to just one type of test vehicle. It would be prudent to supplement this study with data obtained from a variety of vehicle types and possibly additional driving modes to further validate the correlation characteristics between wSII and MOS-LQO results in future studies.

5 CONCLUSION

Automotive voice microphones usually come with designs of different directivities and frequency response shapes. Direct measurement data showing how different microphone types interact with various noise conditions of automotive driving modes (i.e., use cases) are rare to find in the literature. Additionally, the overall SI/SQ performance of automotive communication systems is usually evaluated on the system level where the microphone is integrated with additional signal processing algorithms together. This makes it difficult to identify the microphone influences on the SI/SQ performance perceived by end users.

For years, automotive manufacturers and microphone suppliers have followed general rules for choosing microphone designs to meet certain requirements. For example, unidirectional microphones or microphones with a rising frequency response are almost always recommended for automotive HF voice applications. A unidirectional microphone gives higher SNR than an omnidirectional microphone because of its ability to filter out off-axis noise in the background, and a microphone with a rising frequency response shape helps remove low-frequency noise in the vehicle environment. This recommendation was established mainly for the purpose of improving the SNR and thus may overlook the frequency response shape and directivity influences on the SI/SQ performance under various driving conditions. When the SI is considered together with the SQ that also includes factors like sound naturalness and noise intrusiveness, those general rules may be not sufficient to decide which type of microphone would result in preferred performance under a specific driving mode. This problem could be further intensified considering that the driving mode changes constantly when travelling in a car.

Based on three common SI/SQ evaluation methods, this study measured the performance of three typical automotive microphone designs under different driving modes using a gasoline-engine SUV. It was shown that there is no "one-size-fits-all" microphone design that can deliver optimal SI/SQ performance for all possible driving conditions. In general, an omnidirectional microphone with a flat frequency response shape (i.e., Omni-ANC) is preferred at low noise levels because it provides sufficient SI and conserves the naturalness of the sound. Under driving conditions with high wind-induced noise, use of unidirectional microphones should be avoided. On the other hand, under most driving conditions with non–wind-related noise, Uni-HFs seem to give the best SI/SQ performance, followed by the Omni-HF with a rising frequency response shape.

Furthermore, it was verified that the objective SII results are fundamentally linearly correlated with the subjective SI results in automotive voice microphone applications. On the contrary, a simple linear correlation between the original SII and the objective MOS-LQO results did not seem to exist. However, it was demonstrated that, when a nonlinear weighting function based on the SNR was multiplied to the original SII, the resulting wSII was approximately linearly correlated with the MOS-LQO. As the objective SII/wSII calculation process only requires the measured background noise spectrum, its practical implementation is much simpler and less costly than the subjective SI or MOS test. Supported by findings from the correlation analysis, it is concluded that the SII/wSII method can be conveniently used to evaluate and compare microphone designs for automotive voice applications. It may be even plausible to build computation algorithms of SII/wSII into hardware and/or software designs of automotive voice microphones so that their acoustic characteristics may be adaptively adjusted in real-time based on the background noise conditions.

Lastly, although the mathematical expression of the nonlinear weighting function and its parameter values proposed in this study for the wSII calculation seemed to work well to reconstruct a linear correlation between the wSII and MOS-LQO results, it is recognized that the correlation characteristics between wSII and MOS-LQO data should be investigated further using a variety of vehicle types and additional driving modes.

6 REFERENCES

[1] N. R. French and J. C. Steinberg, "Factors Governing the Intelligibility of Speech Sounds," *J. Acoust.*

Soc. Am., vol. 19, no. 1, pp. 90–119 (1947 Jan.). https://doi.org/10.1121/1.1916407.

[2] J. Wouters, L. Litiere, and A. Wieringen, "Speech Intelligibility in Noisy Environments with One- and Two-Microphone Hearing Aids," *Int. J. Audiology*, vol. 38, no. 2, pp. 91–98 (1999 Jul.). https://doi.org/10.3109/00206099909073008.

[3] K. S. Rhebergen and N. J. Versfeid, "A Speech Intelligibility Index-Based Approach to Predict the Speech Reception Threshold for Sentences in Fluctuating Noise for Normal-Hearing Listeners," *J. Acoust. Soc. Am.*, vol. 117, no. 4, pp. 2181–2192 (2005 Apr.). https://doi.org/10.1121/1.1861713.

[4] K. S. Rhebergen, J. Lyzenga, W. A. Dreschler, and J. M. Festen, "Modeling Speech Intelligibility in Quiet and Noise in Listeners With Normal and Impaired Hearing," *J. Acoust. Soc. Am.*, vol. 127, no. 3, pp. 1570–1582 (2010 Mar.). https://doi.org/10.1121/1.3291000.

[5] A. J. Brammer, G. Yu, E. R.Â Bernstein, et al., "Intelligibility of Speech Corrupted by Nonlinear Distortion," in *Proceedings of the 10th International Congress on Noise as a Public Health Problem (ICBEN)*, vol. 24, pp. 287–292 (London, UK) (2011 Jul.).

[6] M. Blue-Terry, M. McBride, and T. Letowski, "Effects of Speech Intelligibility on the Callsign Acquisition Test (CAT) and Modified Rhyme Test (MRT) Presented in Noise," *Arch. Acoust.*, vol. 37, no. 2, pp. 199–203 (2012 Jan.). https://doi.org/10.2478/v10168-012-0026-3.

[7] S. Maeda, K. Kobayashi, H. Nakatani, and A. Nakatani, "Comparison of Speech Intelligibility Between Normal Headsets and Bone Conduction Hearing Devices at Call Center," in *Proceedings of the 43rd International Congress on Noise Control Engineering (INTER-NOISE)*, paper 323 (Melbourne, Australia) (2014 Nov.).

[8] N. Mansour, M. Marschall, T. May, A. Westermann, and T. Dau, "Speech Intelligibility in a Realistic Virtual Sound Environment," *J. Acoust. Soc. Am.*, vol. 149, no. 4, pp. 2791–2901 (2021 Apr.). https://doi.org/10.1121/10.0004779.

[9] R. Frodl, "New Generation of Active Microphones for Handsfree Communication and Voice Recognition," *SAE Tech. Rep. 1999-01-0487* (1999 Mar.).

[10] T. R. Letowski and A. A. Scharine, "Correlational Analysis of Speech Intelligibility Tests and Metrics for Speech Transmission," *ARL-TR-8227* (2017 Dec.). https://apps.dtic.mil/sti/pdfs/AD1043203.pdf.

[11] H. W. Gierlich, "Speech Quality Evaluation in Vehicles," in *Proceedings of the Jahrestagung für Akustik (DAGA)*, pp. 52–55 (Kiel, Germany) (2017 Mar.).

[12] American National Standard Institute, "Methods for Calculation of the Speech Intelligibility Index," *ANSI S3.5-1997* (1997 Jun.).

[13] American National Standard Institute, "Methods for Calculation of the Articulation Index," *ANSI S3.5-1969* (1969 Jan.).

[14] G. Ebbitt, "The Assessment of Speech Intelligibility in Vehicles," in *Proceedings of the NOISE-CON*, pp. 334–337 (Portland, ME) (2001 Oct.).

[15] C.G. Herrera, P.F. Donoso-Garcia, and E.B. Medeiros, "Intelligibility in Low-Cost Automotive Audio Systems," *J. Audio Eng. Soc.*, vol. 64, no. 5, pp. 320–331 (2016 May). https://doi.org/10.17743/jaes.2016.0010.

[16] International Electrotechnical Commission, "Sound System Equipment - Part 16: Objective Rating if Speech Intelligibility by Speech Transmission Index," *IEC 60268–16* (2020 Sep.).

[17] G. L. Ebbitt and T. M. Remtema, "Automotive Speech Intelligibility Measurements," *Sound Vibr.*, vol. 51, no. 6, pp. 6–11 (2017 Jun.).

[18] ITU, "Perceptual Evaluation of Speech Quality (PESQ): An Objective Method for End-to-End Speech Quality Assessment of Narrow-Band Telephone Networks and Speech Codecs," *ITU-T Recommendation P.862* (2001 Feb.).

[19] ITU, "Wideband Extension to Recommendation P.862 for the Assessment of Wideband Telephone Networks and Speech Codecs," *ITU-T Recommendation P.862.2* (2007 Nov.).

[20] ETSI, "Speech Processing, Transmission and Quality Aspects (STQ); Speech Quality Performance in the Presence of Background Noise, Part 3: Background Noise Transmission - Objective Test Methods," *ETSI EG 202 396-3* (2008 Nov.).

[21] ITU, "Method for Objective and Subjective Assessment of Quality: Subjective Performance Assessment of Telephone-Band and Wideband Digital CODECS," *ITU-T Recommendation P.830* (1996 Feb.).

[22] U. Muesch, F. Kettler and M. Lepage, "Speech Quality Tests of In-Car Communication Systems," in *Proceedings of the ITG-Fachbericht 236: Sprachkommunikation*, pp. 1–4 (Braunschweig, Germany) (2012 Sep.).

[23] V. Dobos, B. Vargan, and R. Yang Y. Du, "Automotive Microphone Performance: From Specification to User Experience," in *Proceedings of the AES International Conference on Automotive Audio* (2019 Sep.), paper 5.

[24] American National Standard Institute, "Method for Measuring the Intelligibility of Speech Over Communication Systems," *ANSI S3.2-2009* (2009 May).

[25] J. Lyzenga and K. Rhebergen. "Auditory Model for the Speech Audiogram," presented at the *2nd Workshop on Speech in Noise: Intelligibility and Quality* (Amsterdam Netherlands) (2010 Jan.).

[26] ITU, "Telephone Transmission Quality, Telephone Installations, Local Line Networks: Wideband Hands-Free Communication in Motor Vehicles," *ITU-T Recommendation P.1110* (2017 Mar.).

[27] B. W. Anderson and J. T. Kalb, "English Verification of the STI Method for Estimating Speech Intelligibility of a Communications Channel," *J. Acoust. Soc. Am.*, vol. 81, pp. 1982–1985 (1987 Jun.). https://doi.org/10.1121/1.394764.

THE AUTHOR



Yu Du

Yu Du is a Distinguished Engineer in the Automotive Embedded Audio Group at Harman International. He received his B.S. and M.S. degrees in Vehicle Engineering from Tsinghua University (Beijing, China) and subsequently a Ph.D. degree in Mechanical Engineering from Virginia Tech (Blacksburg, VA) in 2003. Throughout his career, Dr. Du has worked on numerous R&D projects in various fields of engineering acoustics including vehicle NVH, structural vibration and noise control, and MEMS acoustic transducers. His current work at Harman focuses on advanced research and development of innovative sensing technologies for automotive voice communication, noise management, and other emerging applications related to sound and vibration. Dr. Du is a member of the Audio Engineering Society and the Acoustical Society of America. He acted as the Papers Chair for the AES 2022 International Conference on Automotive Audio, which was held in Dearborn, MI, in June 2022.