Expanding the Reach of Microphones: Improving Intelligibility

Dean R. G. Anderson and
Dean G. Anderson, M.D.

ABSTRACT- The effects of microphone equivalent noise on speech intelligibility are examined. A novel solution is proposed.

Recording studio engineers understand that they cannot push the use of a small microphone beyond 30 cm without degrading the intelligibility or quality of the speech signal. Not surprisingly, when smart phone, digital assistant, in-cabin audio, and hearing aid OEMs push the use of their small microphones beyond that limit, they get degraded results. What is the source of this limitation and is there a solution?

There are many different factors that can limit the output signal quality from a microphone. The contribution of microphone equivalent noise to the output signal quality is not frequently discussed. This article will review the science describing microphone equivalent noise and how it affects speech intelligibility in a best-case model. Finally, it will discuss a potential solution to improve microphone performance at a distance and why this solution works.

Microphone Equivalent Noise

Microphone equivalent noise is the noise a microphone makes in the absence of any audio input.[1] Most microphones are pressure sensitive capacitors.[2] The size of a microphone determines its capacitance. Sound pressure variations (e.g. speech) cause the electric field in the capacitor to change.[3] The electric field in the capacitor is then turned into a voltage by an amplifier. Microphone equivalent noise is inversely proportional to the capacitance connected to the amplifier. What this means in practical terms is that the smaller a microphone is, the larger its equivalent noise will be.[4] Furthermore, microphone equivalent noise is constant and is mixed into the speech signal by the microphone. Any amplification to the output signal of a microphone will amplify both the desired signal and the microphone equivalent noise.

A straightforward way to visualize the phenomenon of microphone equivalent noise is available to most smart phone users. This is done by turning on the FFT (Fast Fourier Transform) display of a basic sound meter app while looking at the noise above 1000 Hz when no one is talking. Because most ambient noise in a room or car is below 1000 Hz, the noise seen above 1000 Hz mostly represents the equivalent noise of the smart phone's microphone.

Best-Case Model

One simple way to model the maximum theoretical efficacy of a microphone is to use a best-case model. In such a model, speech is presented to a microphone at normal speech levels, from varying distances, and without competing sound sources. ANSI Standard S3.5-1997, *Methods for Calculation of the Speech Intelligibility Index*, can be used to calculate the maximum theoretical Speech Intelligibility Index (SII) under this best-case model.[5] The SII represents the proportion of speech cues that an individual with normal hearing would discern if presented with the audio signal generated by the microphone at various distances from the speech source and inclusive of the microphone equivalent noise.

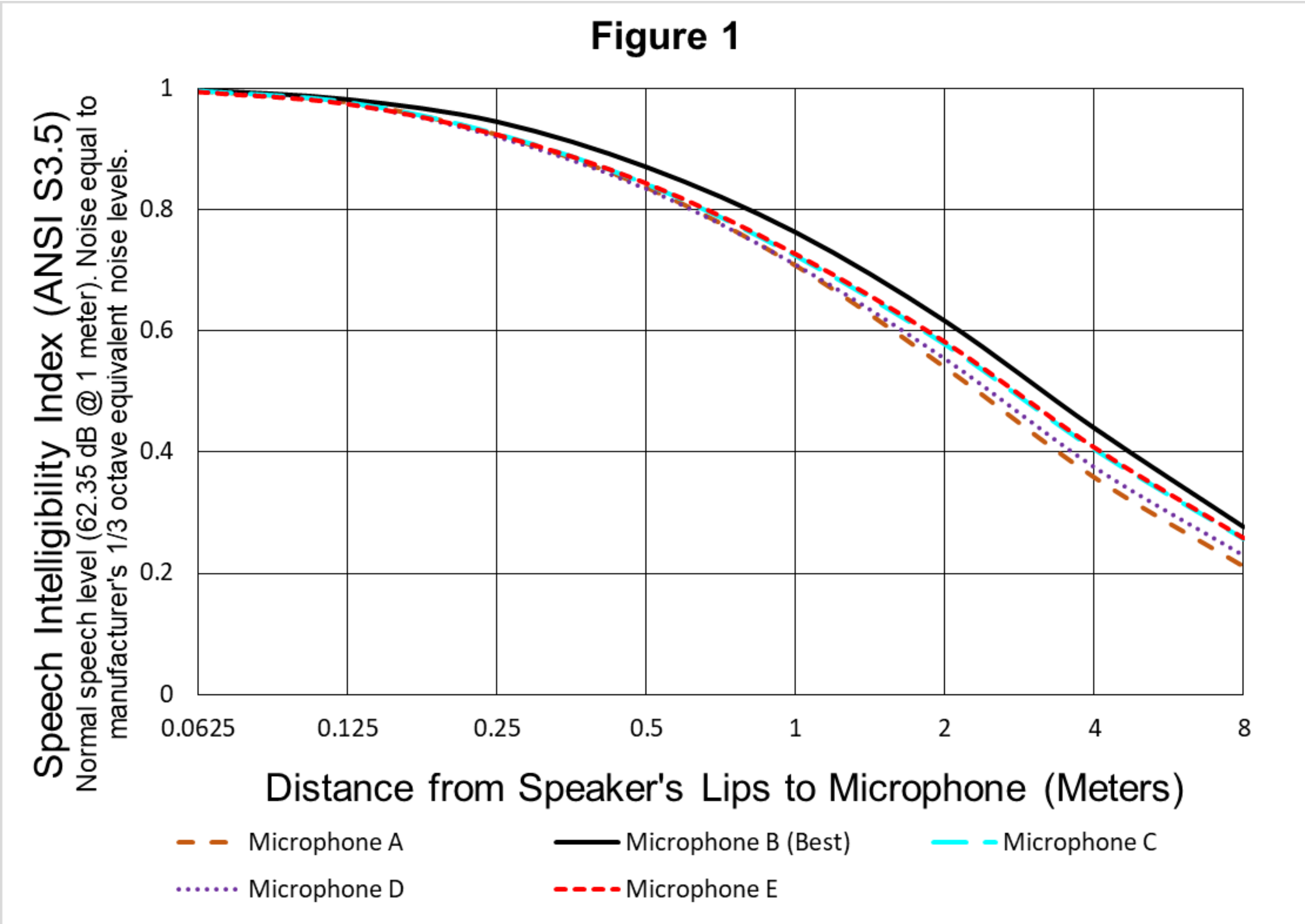


Figure 1 shows the maximum theoretical SII using this best-case model for five different, common, small microphones using the microphone manufacturer's published 1/3-octave equivalent noise levels as the only noise sources.

Proposed Solution

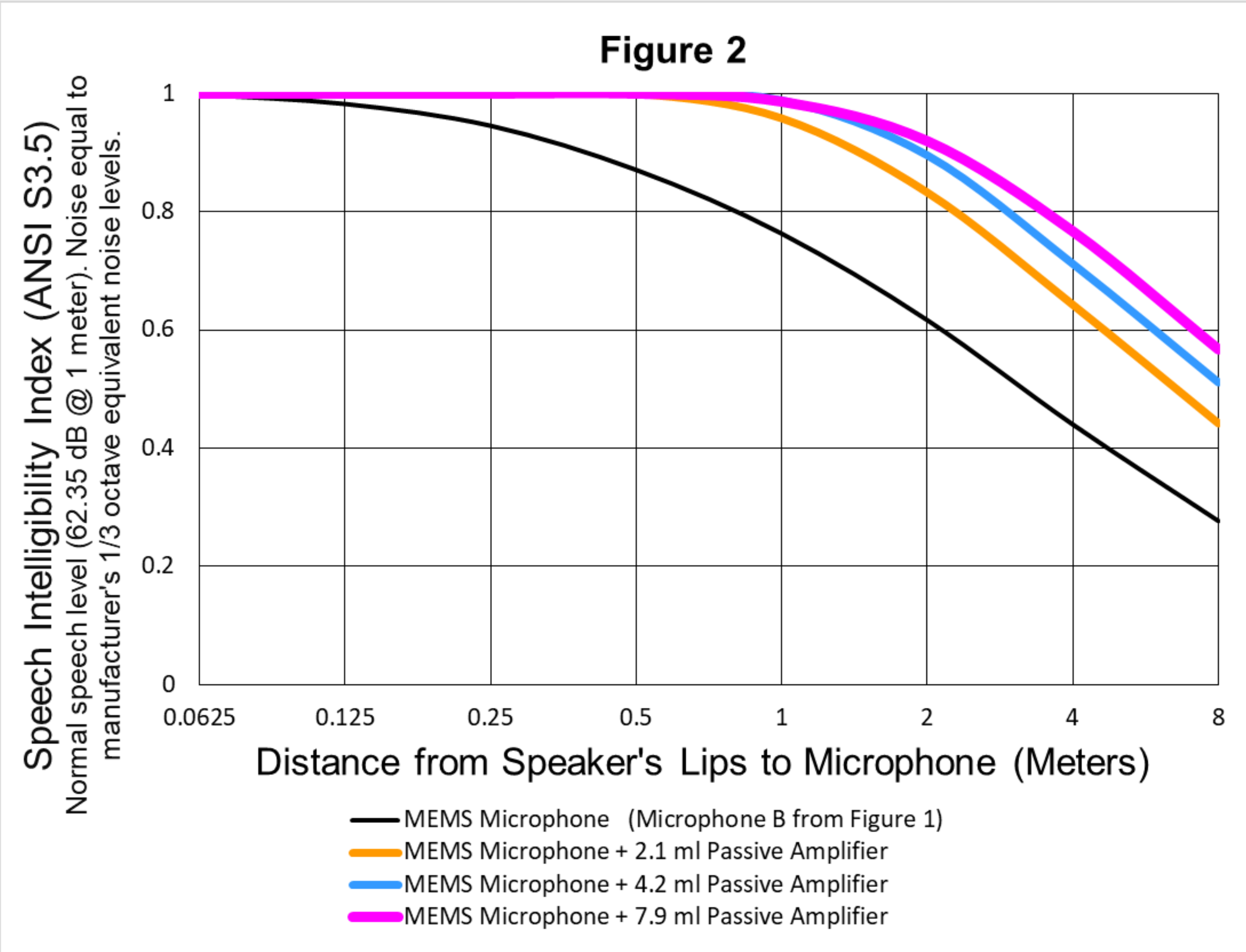
By passively amplifying a speech signal prior to the speech signal reaching the microphone (i.e. prior to the addition of microphone equivalent noise), the SII of a signal generated by a microphone can be greatly improved.

Pixation Corporation and its associates have developed various passive amplifier solutions for consumer devices such as smart phones, digital assistants, in-cabin audio systems and hearing aids. A description of these technological advancements in the field of passive amplifiers is beyond the scope of this paper. Instead, this paper will present some experimental results of these solutions as applied to small microphones.

A unique I^2S MEMS microphone was chosen to collect the differential data reported in this article. This I^2S MEMS microphone uses the industry standard 24-bit digital I^2S interface to report an audio signal sampled during every 32 microsecond period (a 31,250 Hz sampling rate). The samples were combined together to determine an audio measurement with 0.1 dB SPL precision during a 40.96 millisecond period. Two hundred fifty-six sequential measurements were used to determine an audio level during a 10.48576 second period.

The purpose of using differential data is to study only the effects of the passive amplifier rather than conflating microphone data with passive amplifier data. Thus, an insertion gain is computed by subtracting audio levels determined for the I^2S MEMS microphone without a passive amplifier from the audio levels determined using the same I^2S MEMS microphone with a passive amplifier.

Audio levels were determined in a typical, large, mixed-use, furnished room having various sound reflective surfaces (i.e., drywall, uncovered



window, hardwood floor) so as to be more representative of difficult, real-world, listening conditions. White noise recordings were used to suppress standing waves which might corrupt audio measurements. The white noise used had balanced power spectral density. The white noise was attenuated 48 dB per octave at band limits. Both the speaker and the microphone were free-standing on 16 mm poles, 1.4 meters above the floor and separated by 1 meter.

Insertion gains for three differently sized passive amplifiers were determined using the method described above. The passive amplifiers had interior volumes of 2.1, 4.2 and 7.9 ml respectively.

Figure 2 depicts the maximum theoretical SII for Microphone B (a MEMS microphone) from Figure 1 and includes depictions for this microphone with insertion gains benefitting SII for: (1) the 2.1 ml passive amplifier[6]; (2) the 4.2 ml passive amplifier[7]; and (3) the 7.9 ml passive amplifier[8].

It is noted that a large diaphragm microphone (e.g. 25 mm) could exhibit an SII similar to the SII of the MEMS microphones with passive amplifiers illustrated in Figure 2. However, large diaphragm microphones have greater power requirements than MEMS microphones with passive amplifiers. The passive amplifiers have no power requirement.

Why Passive Amplifiers Work

Understanding the contribution that distance has on the speech signal while microphone equivalent noise levels remain unchanged is essential to understanding both the limits of current technology and the potential benefits of including a passive amplifier. It is therefore helpful to take a more detailed look at the one-third octave procedure in ANSI S3.5 using the microphone equivalent noise levels from Microphone B in Figure 1.[9]

TABLE 1

Nominal Band Center Frequency (Hertz)	Band Impor-tance:	Standard Speech Spectrum for Normal Vocal Effort (dB) @ 1 meter	Standard Speech Spectrum for Normal Vocal Effort (dB) @ 2 meters	Best Micro-phone 1/3 Octave Equiv-alent Noise (dB)	2.1 ml Passive Amplifier Insertion Gain (dB)	4.2 ml Passive Amplifier Insertion Gain (dB)	7.9 ml Passive Amplifier Insertion Gain (dB)
1000	0.0818	25.01	18.99	10	1	1.6	3.7
1250	0.0844	23	16.98	9.8	1.5	2.3	5.7
1600	0.0882	20.15	14.13	9.7	2.3	4.2	12.1
2000	0.0898	17.32	11.3	9.7	3.3	7.2	19.8
2500	0.0868	13.18	7.16	9.9	5.9	17.3	20.4
3150	0.0844	11.55	5.53	10.2	12.6	21.1	13
4000	0.0771	9.33	3.31	10.6	20.9	21.5	14.4
5000	0.0527	5.31	-0.71	11.1	24.4	17.5	21.4
6300	0.0364	2.59	-3.43	11.8	24.9	21.9	20.5
8000	0.0185	1.13	-4.89	12.7	22.8	24.7	22.5

Hearing is segmented into critical bands.[10] Table 1 looks at the frequency bands starting with the 1000 Hz band because: (1) 70% of all speech cues are within these bands (the summation of band importance); (2) ambient room or car noise is mostly below the 1000 Hz band; and (3) passive amplification is greatest for the 1000 Hz band and above.

Note in Table 1 how distance affects standard speech spectrum levels which drop by 6.02 dB between 1 meter and 2 meters.

At 1 meter, standard speech spectrum levels for the bands from 4000 Hz to 8000 Hz are less than the corresponding microphone B equivalent noise levels.

At 2 meters, standard speech spectrum levels for the bands from 2500 Hz to 8000 Hz are less than the corresponding microphone B equivalent noise levels.

Adding insertion gains from any of the passive amplifiers studied results in raising the standard speech spectrum levels for all bands to levels above the corresponding microphone B equivalent noise levels for both 1 meter and 2 meters.

Conclusions

The effects of microphone equivalent noise on speech intelligibility are demonstrable and significant. Solving this problem for small microphones has the potential to meet consumer demands in multiple expanding industries.

Contact pixation@pixation.com for additional information.

[1] ASA Secretariat: Acoustical Society of America. ANSI/ASA S3.22-2009 Specification of Hearing Aid Characteristics. New York, NY.: Acoustical Society of America; American National Standards Institute, Inc. Approved November 10, 2009: pp.13.
[2] Ballou, Glen (Editor). Handbook for Sound Engineers (Audio Engineering Society Presents). 5th Edition. New York, NY: Routledge; 2015: pp. 601-700, 1387-1412.
[3] Ibid. pp.626.
[4] Ibid., pp.631.
[5] ASA Secretariat: Acoustical Society of America. ANSI S3.5-1997 Methods for Calculation of the Speech Intelligibility Index. New York, NY.: Acoustical Society of America; American National Standards Institute, Inc. Approved 6 June 1997.pp.5.
[6] The 2.1 ml passive amplifier exhibited the following insertion gains (Gi): 160 Hz/0.2 dB; 200 Hz/0.0 dB; 250 Hz/0.3 dB; 315 Hz/0.3 dB; 400 Hz/0.2 dB; 500 Hz/0.5 dB; 630 Hz/0.4 dB; 800 Hz/0.4 dB; 1000 Hz/1.0 dB; 1250 Hz/1.5 dB; 1600 Hz/2.3 dB; 2000 Hz/3.3 dB; 2500 Hz/5.9 dB; 3150 Hz/12.6 dB; 4000 Hz/20.9 dB; 5000 Hz/24.4 dB; 6300 Hz/24.9 dB; and, 8000 Hz/22.8 dB.
[7] The 4.2 ml passive amplifier exhibited the following insertion gains (Gi): 160 Hz/0.2 dB; 200 Hz/0.8 dB; 250 Hz/0.7 dB; 315 Hz/0.8 dB; 400 Hz/0.2 dB; 500 Hz/0.7 dB; 630 Hz/0.9 dB; 800 Hz/1.2 dB; 1000 Hz/1.6 dB; 1250 Hz/2.3 dB; 1600 Hz/4.2 dB; 2000 Hz/7.2 dB; 2500 Hz/17.3 dB; 3150 Hz/21.1 dB; 4000 Hz/21.5 dB; 5000 Hz/17.5 dB; 6300 Hz/21.9 dB; and, 8000 Hz/24.7 dB.
[8] The 7.9 ml passive amplifier exhibited the following insertion gains (Gi): 160 Hz/0.3 dB; 200 Hz/0.7 dB; 250 Hz/0.9 dB; 315 Hz/0.6 dB; 400 Hz/1.1 dB; 500 Hz/1.4 dB; 630 Hz/1.9 dB; 800 Hz/2.6 dB; 1000 Hz/3.7 dB; 1250 Hz/5.7 dB; 1600 Hz/12.1 dB; 2000 Hz/19.8 dB; 2500 Hz/20.4 dB; 3150 Hz/13 dB; 4000 Hz/14.4 dB; 5000 Hz/21.4 dB; 6300 Hz/20.5 dB; and, 8000 Hz/22.5 dB.
[9] ASA Secretariat: Acoustical Society of America. ANSI S3.5-1997 Methods for Calculation of the Speech Intelligibility Index. New York, NY.: Acoustical Society of America; American National Standards Institute, Inc. Approved 6 June 1997.pp.5.
[10] Fletcher, H. Speech and Hearing in Communication (1st ed.). Huntington, New York: Krieger; 1972:170-175.