Andrea Audio Test Labs

Andrea
PureAudio BT-200 Noise Canceling Bluetooth Headset
Performance Comparative Testing

November 26th 2008

Rev B

Andrea Electronics Corporation
65 Orville Drive
Suite One
Bohemia NY 11716
(631)-719-1800
www.andreaelectronics.com

Andrea PureAudio BT-200 Noise Canceling Bluetooth Headset (Patent Pending) eliminates real world noise without any audio degradation of voice quality, better than any other Bluetooth headset tested. Though some Bluetooth headsets are capable of reducing ambient noise, the BT-200 is the only headset that does so with out voice distortion. This report provides a standard audio engineering analysis procedure to demonstrate the performance differences of the below products.

The Bluetooth devices subjected to performance comparative testing are:

Andrea BT-200

Jawbone

BlueAnt Z9i

Plantronics Calisto

Plantronics Voyager 855

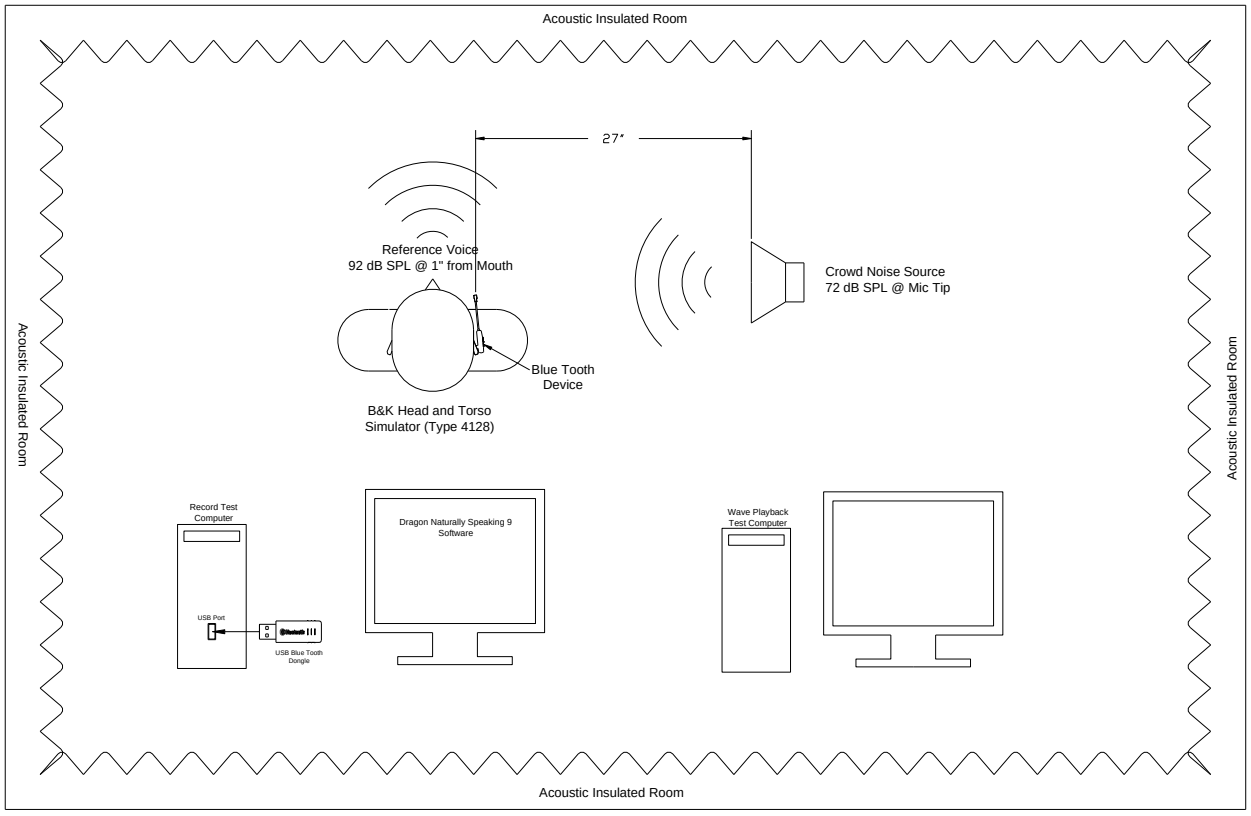
Logitech Cordless Headset

Motorola H500

Test Equipment Setup (Refer to Set up Figure illustration)

- All Bluetooth devices are placed in a fixed position on the right Right Ear of the B&K Head and Torso Simulator.
- A generated wave file of “Crowd Noise” (obtained from CBS Audio file sound effects library) is played through a speaker placed 27” away from the Bluetooth devices on the B&K Head and Torso Simulator. The “Crowd Noise” output is calibrated and adjusted at fixed setting of 72dB SPL, measured at the Bluetooth device under test.
- Two generated “Speech” wave file (15 Seconds and 3 Minutes) text readings are played through the B&K Head and Torso Simulator. The “Speech” output is calibrated and adjusted for 92 dB SPL max at 1” distance away from Mouth Simulator.
- Dragon Naturally Speaking Version 9 software application is used on the computer paired with the Bluetooth devices to record the transcribed text reading.

Set Up Figure



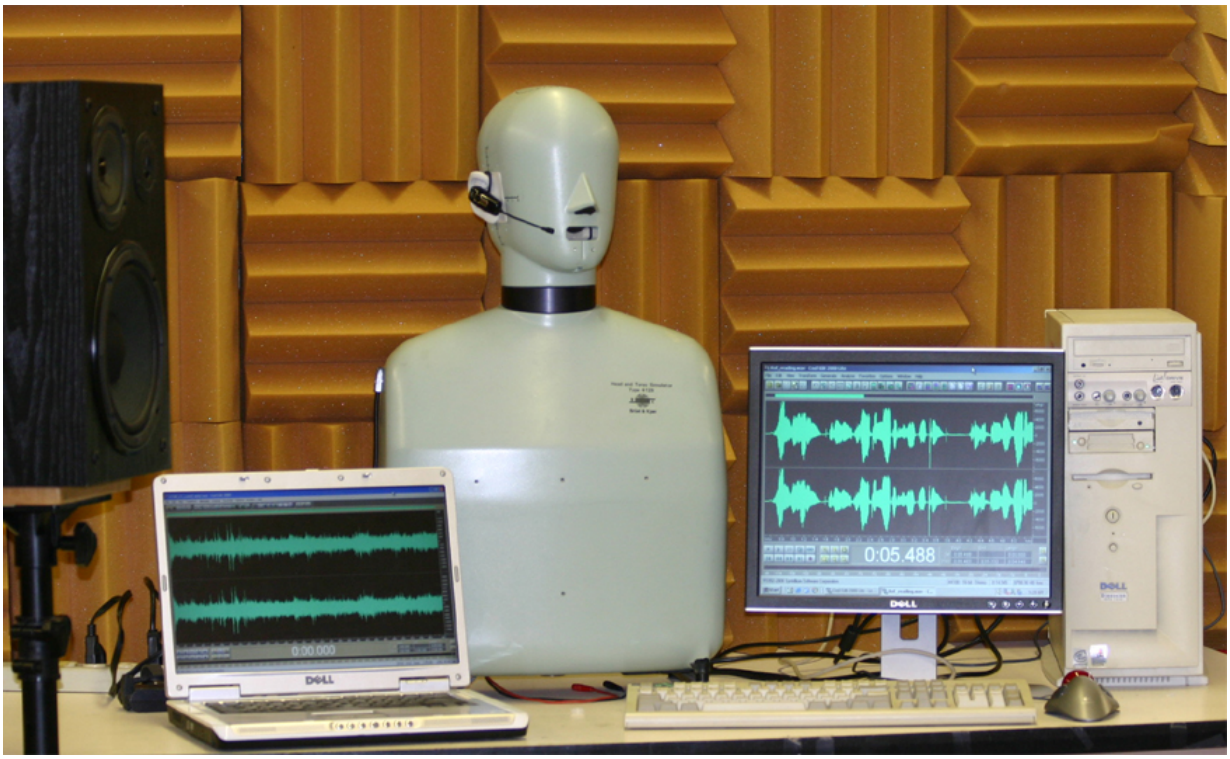
Test Setup

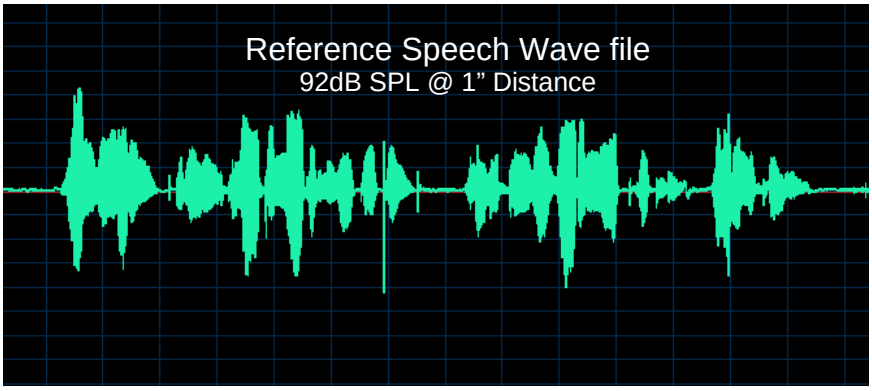
Testing was performed under the following controlled, but real life conditions:

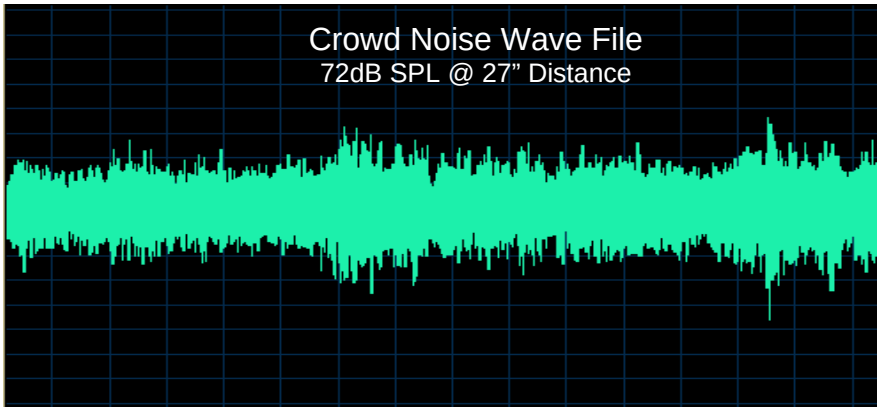
The setup used a Dell Notebook computer as the source of the 1411kbps high quality speech files that were played from the PC through a Brüel & Kjær artificial head and torso simulator. Noise files, when needed, are played from the PC through an amplifier then through an amplified off axis speaker. To ensure repeatable results the Speech and noise files are combined into a single stereo file that is played left channel to the head and torso simulator and right channel to the off axis speaker.
The Device Under Test (DUT) is connected to a second computer that is configured for speech recognition. The input for all tests is via an Andrea Bluetooth USB Audio Adapter with a full duplex audio input and output.

A source of variability is sound pressure level measurements. There are two main “weightings” for Sound Pressure Level (SPL): “A” weighting and “C” weighting. “A” weighting compensates for the non-linear response of the human ear, while “C” weighting is a flatter response. The response of the human ear to sound is worse at low and high frequencies. The two weightings will give different results depending on the frequency distribution of the sound. In an office environment, significant content can be missed by “A” weighting that we cannot hear, but the microphone will pick-up. The microphone and speech engine will be affected by this inaudible noise so we have considered it. Low frequency blower and air exchanger noise is of particular concern, as it is difficult to hear.
All Andrea SPL readings use “C” weightings. We believe “C” weighting reflects more accurately the range of sounds the microphone picks up.

Test Set Up

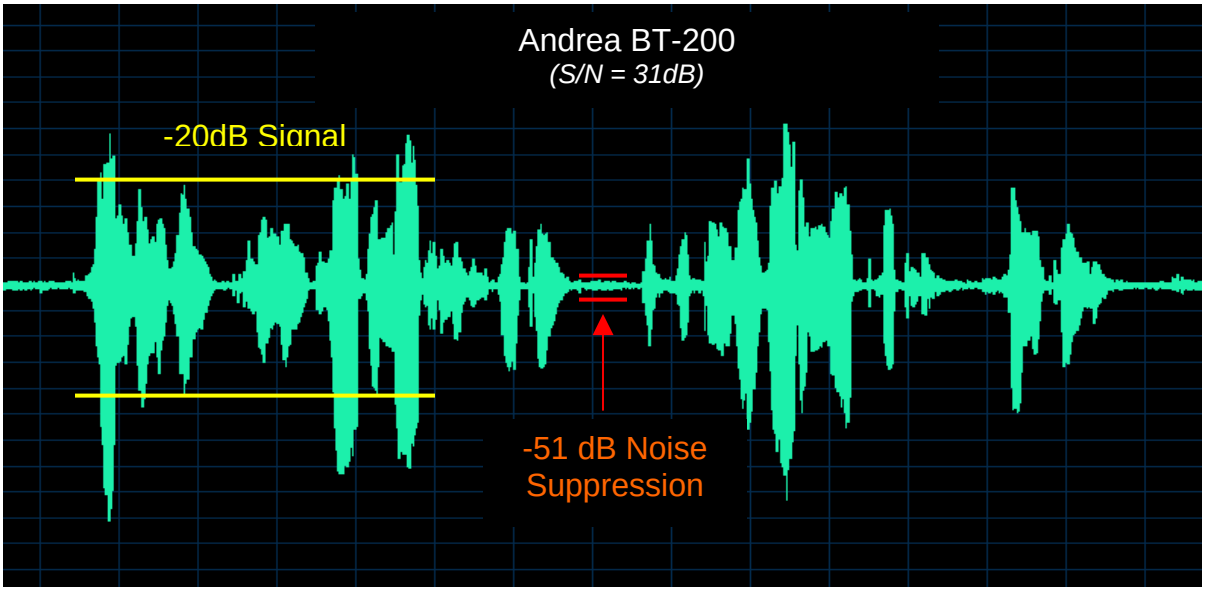

0:05.488

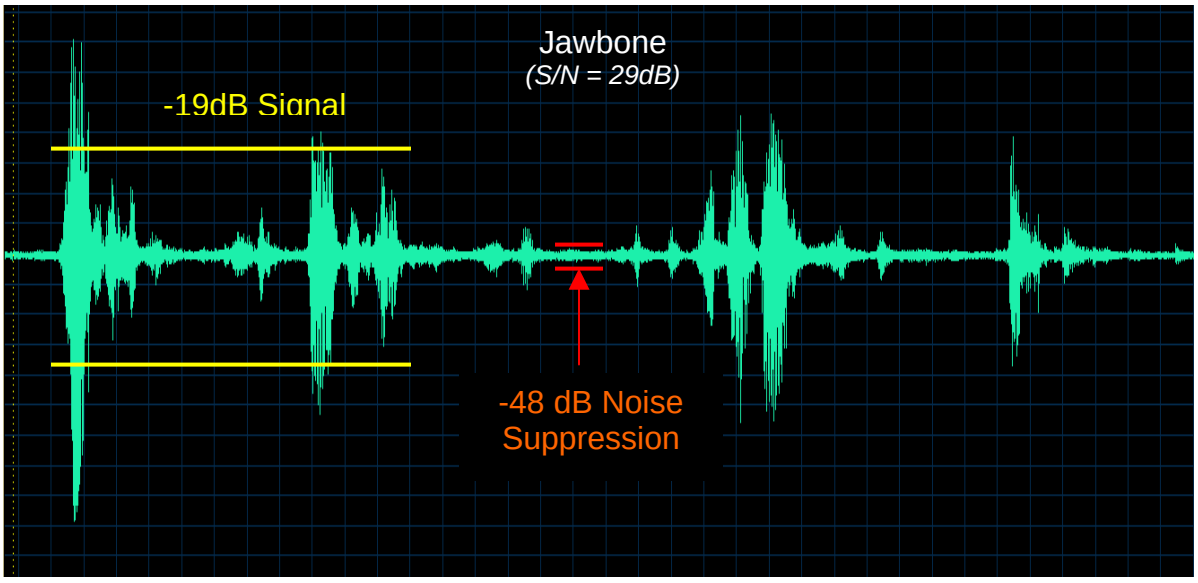

Reference Speech Wave file
92dB SPL @ 1” Distance

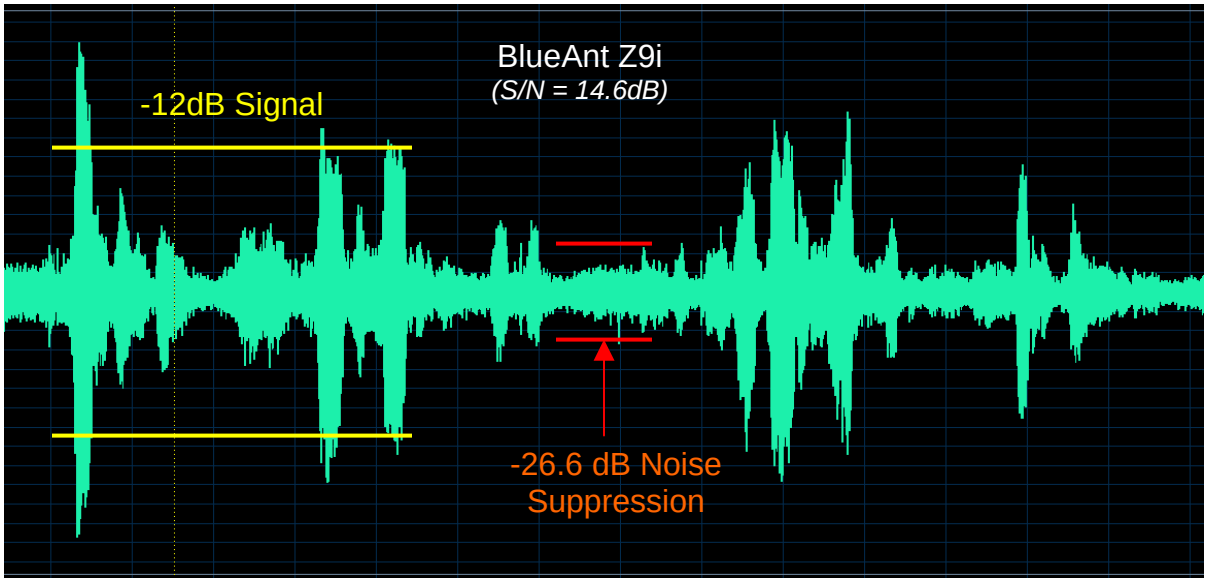

Crowd Noise Wave File
72dB SPL @ 27” Distance

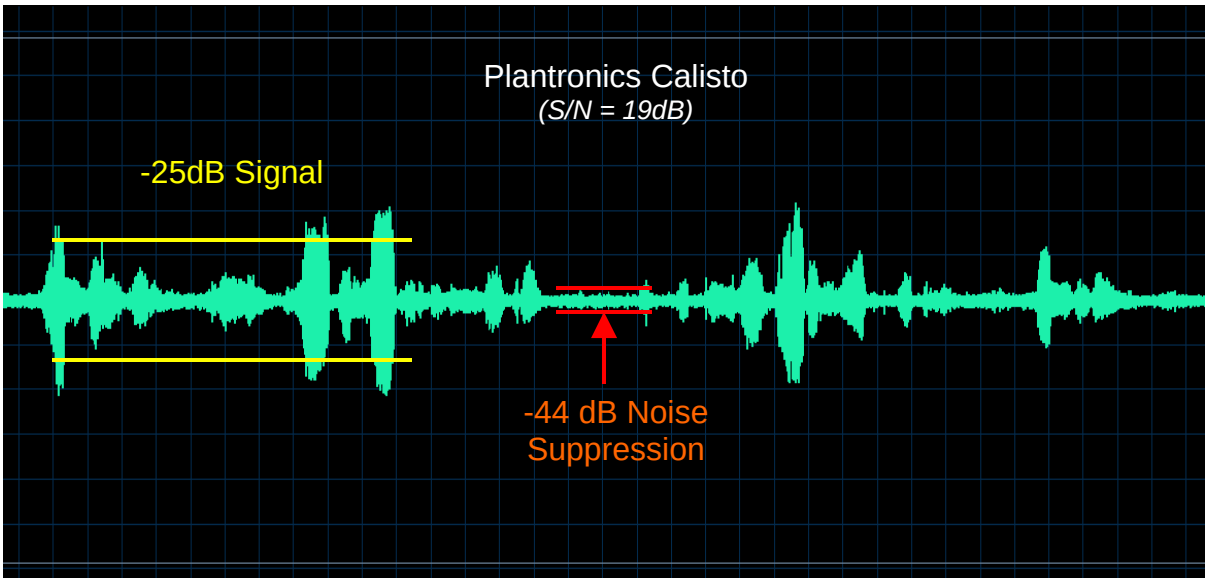
Noise Cancellation Results:

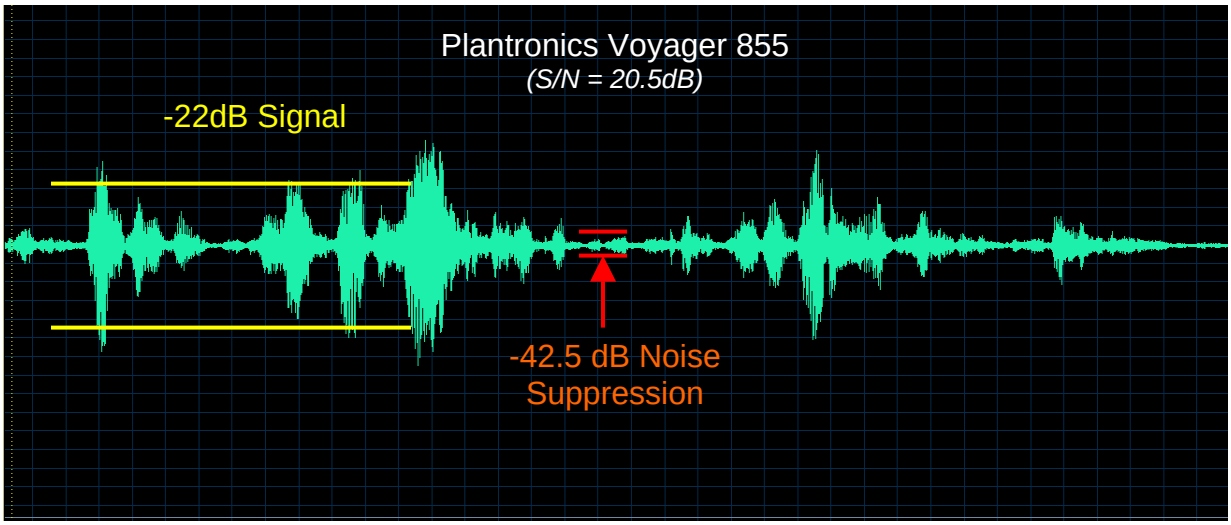
Andrea PureAudio BT-200 Noise Canceling Bluetooth Headset results indicate a 21 dB suppression of background “Crowd Noise” imposed into the microphone while speech remains clear and undistorted.

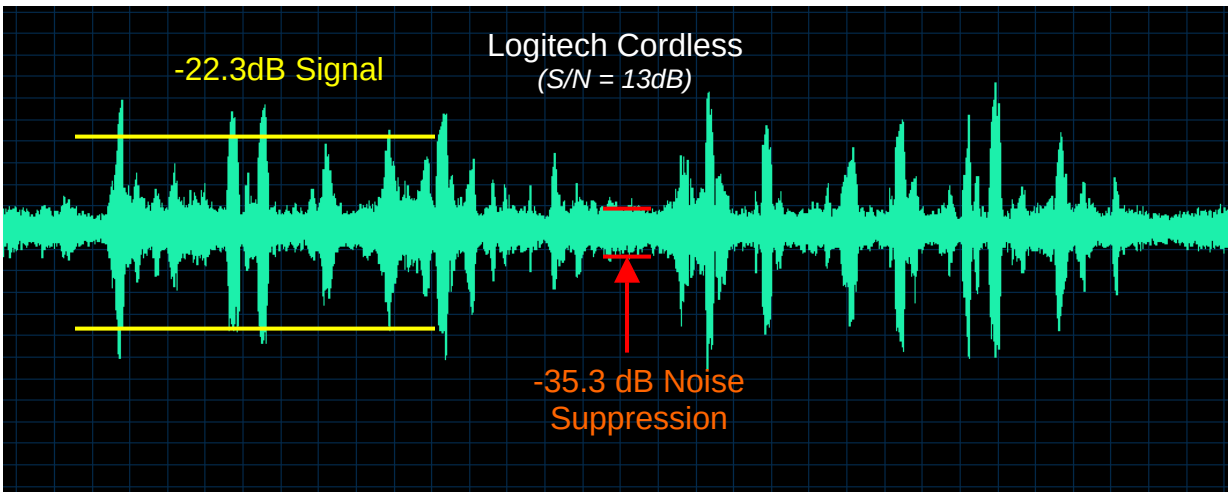


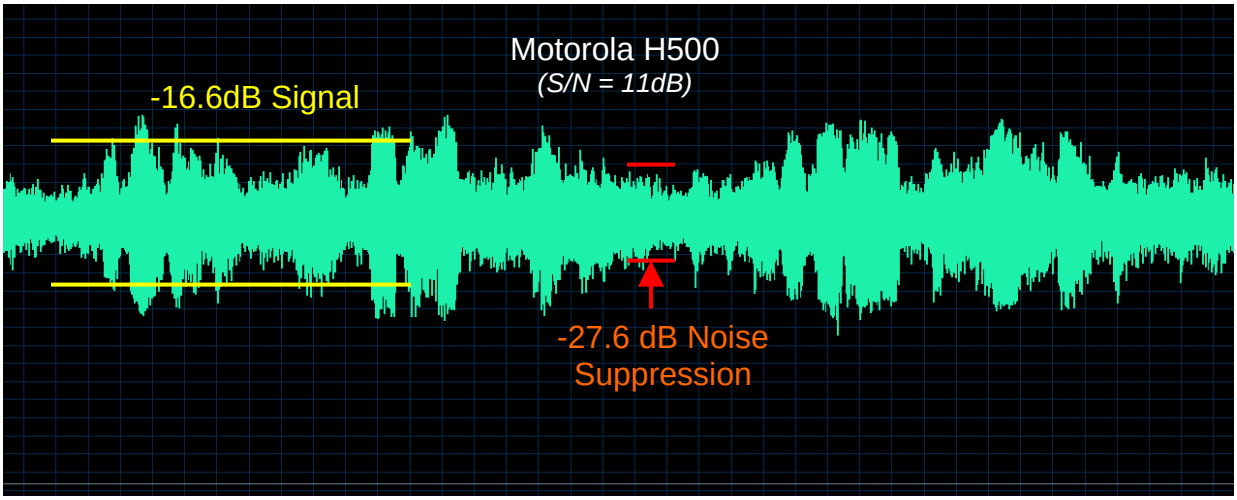



BlueAnt Z9i
(S/N = 14.6dB)
-12dB Signal
-26.6 dB Noise
Suppression


Plantronics Calisto
(S/N = 19dB)
-25dB Signal
-44 dB Noise
Suppression


Plantronics Voyager 855
(S/N = 20.5dB)
-22dB Signal
-42.5 dB Noise
Suppression


Logitech Cordless
(S/N = 13dB)
-22.3dB Signal
-35.3 dB Noise
Suppression


Motorola H500
(S/N = 11dB)
-16.6dB Signal
-27.6 dB Noise
Suppression

Speech Recognition (Distortion) results:

We compare the Andrea PureAudio BT-200 Noise Canceling Bluetooth Headset Microphone's Speech Recognition performance to other Bluetooth microphones. Tests were performed under the same conditions to determine, on a relative basis, how the BT-200 microphone compares to other noise-reducing Bluetooth headsets. The speech engine utilized in the tests was Nuance Dragon Naturally Speaking continuous speech recognition version 9.
The intention is not to determine absolute Speech Recognition accuracy. With this in mind, we are using only a single male voice for speech input. Minimum training is performed. Additional training will improve absolute accuracy.

Note: The accuracies herein are for comparison purposes only and do not reflect actual user experiences.

All text readings transcribed by Dragon Naturally speaking 9 Software via Bluetooth were rated with a scoring utility application that checks for word recognition efficiency. The scores can be used to determine the amount of distortion in each Bluetooth speech recording. Recordings are taken with and without imposed "Crowd Noise".

Results indicated best over-all speech recognition performance with the Andrea PureAudio BT-200 Noise Canceling Bluetooth Headset. There was less than 5% voice degradation with the background "Crowd Noise". Furthermore, the other Bluetooth Headsets distort the speech in the process of attempting to eliminate ambient background noise. The result of this distortion is a significant degradation in intelligibility and quality in order to achieve a perceptual reduction in background noise. Voice distortion of speech quality will adversely affect user satisfaction by impeding telephony systems accuracy of integrated speech recognition auto attendant command and control of voice systems.

	Scores (%)	
Blue Tooth Device	Normal	Crowd Noise
Andrea BT-200	94.1	89.8
Jawbone	84.2	34.1
Blue Ant Z9i	90.1	33.4
Plantronics Calisto	91.3	52.9
Plantronics Voyager 855	86.9	>10
Logitech Cordless	83.9	23.5
Motorola H500	69.1	>10

Test Methodology Notes

Enrollment is the process of training the speech recognition software to your voice so that it can improve its accuracy. Since each microphone has its own tonal qualities, signal to noise, and sensitivity, each microphone requires separate enrollment.
To minimize test time, and to accentuate any differences between microphones, the minimum enrollment is performed for each test microphone. Accuracy will most certainly improve if additional training is performed.

Our test script has a word count of 323. Punctuation is counted as words. A recording of a “trained voice” 30 yr old male is used. The script is direct dictation into Microsoft Word Vista using the DNS version 9 speech Engine.

Scoring is done using an Andrea Audio Test Labs File Comparison Utility developed for speech test scoring. The utility compares the reference script to the speech-recognized script of the microphone under test. The scoring method counts substitutions, which are words, or commands that the computer misinterprets, and deletions, which are words or commands that the computer misses. The Speech Recognition Accuracy score is calculated as follows:
Accuracy Score (%) = ((Total words – Substitutions – Deletions) / Total words) x 100
Each test is repeated at least three consecutive times, and the average score is calculated.

Because of the lack of an industry standard for “speech recognition accuracy” and the differing scoring approaches, all accuracy data should be made on a relative basis only.

The accuracy scoring is usually lower when insertions are considered as errors. For accuracies of 80% and above the scoring will be about 1-5% lower when considering insertions. For accuracies below 80% the scoring can be 5-20% lower when considering insertions.

Because of the many variables involved with speech recognition testing, recreation of our test environments may result in somewhat different results. The relative results however should be the same. We have made every attempt to minimize test variations, and after hundreds of tests our results have been very consistent