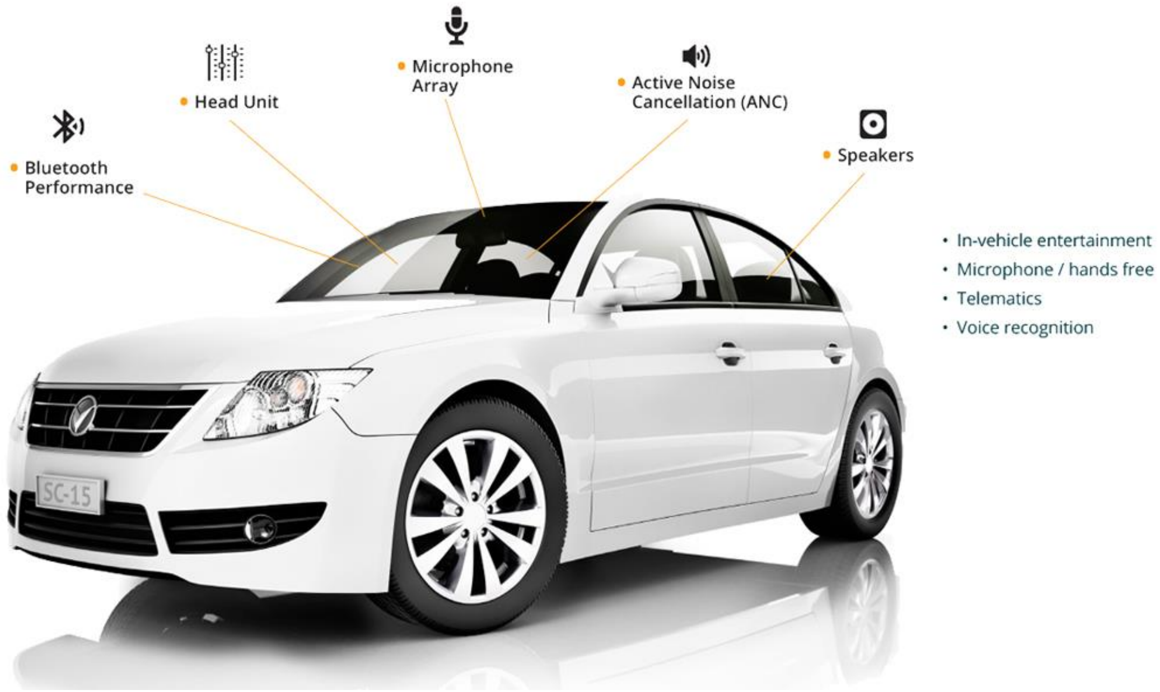
Smart Automobiles

No longer just 'a radio and speakers'…the automobile has become a mobile smart device!



Agenda

- What is a Smart Automobile?
- What makes testing Smart Autos difficult?
- Basic strategy for testing infotainment systems in Smart Autos
- Basic audio measurements;
  - Frequency/Impulse Response, L/R Front & Rear tracking, Distortion, Target Response, and Active Noise Cancellation (ANC)
- Overview of Open Loop testing
- Overview of test signals types and analysis options
- Advanced audio measurements;
  - Voice activation, speech recognition, voice quality, mic(s) directionality and non-coherent distortion
- Summary

What is a Smart Automobile?

- Smart Autos are designed not only to playback music but to;
  - Interface to a Smartphone
  - Responds to voice commands
  - Make hands-free phone calls
  - Play videos (for kids in back seat and sometimes front seats?)
  - Engine/exhaust sound enhancement
  - Warn pedestrians from getting run over
  - Minimize background noise (road, wind & cabin noises)
- This presentation focuses on how to measure their audio performance
  - Very similar to testing a Smart Speaker but with many speakers

What makes testing Smart Autos difficult?

- They have many possible interfaces;
  - hard-wired or auxiliary input, radio, CD, memory card, hard drive, USB, Bluetooth, A2B, smartphone (including Apple CarPlay and Android Auto)…
- They responds to Voice Commands e.g.
  - "Hey Siri, play me a song"
  - "Alexa, what is the weather?"
  - "OK Google, give me directions"
- Wireless interfaces e.g. Bluetooth & Wi-Fi
- Many microphones & speakers
- Complex Signal Processing e.g.
  - ANC, voice activity detectors, mic array beamforming, loudness control, equalization, compressors, etc.
  - Some of the signal processing is done in the smartphone and some is done in the head unit e.g. echo cancellers. This can lead to problems.
- Smart Autos are notoriously complex to test!

Some Possible Standards

- IEC 60268 Sound system equipment
- IEEE 1329-2010, “Standard Method for Measuring Transmission Performance of Speakerphones”, October 2010
- TIA 920.130: Telecommunications Communications Products Transmission
- ITU-T P.56 05/93 Method B: Active Speech Level
- ETSI ES 202 396-1: Speech and multimedia Transmission Quality (STQ); Speech quality performance in presence of background noise; Part 1: Background noise simulation technique and background noise database
- ***None of these standards address how to connect to them and how to measure them under real world conditions with real world signals!***

Test Strategies

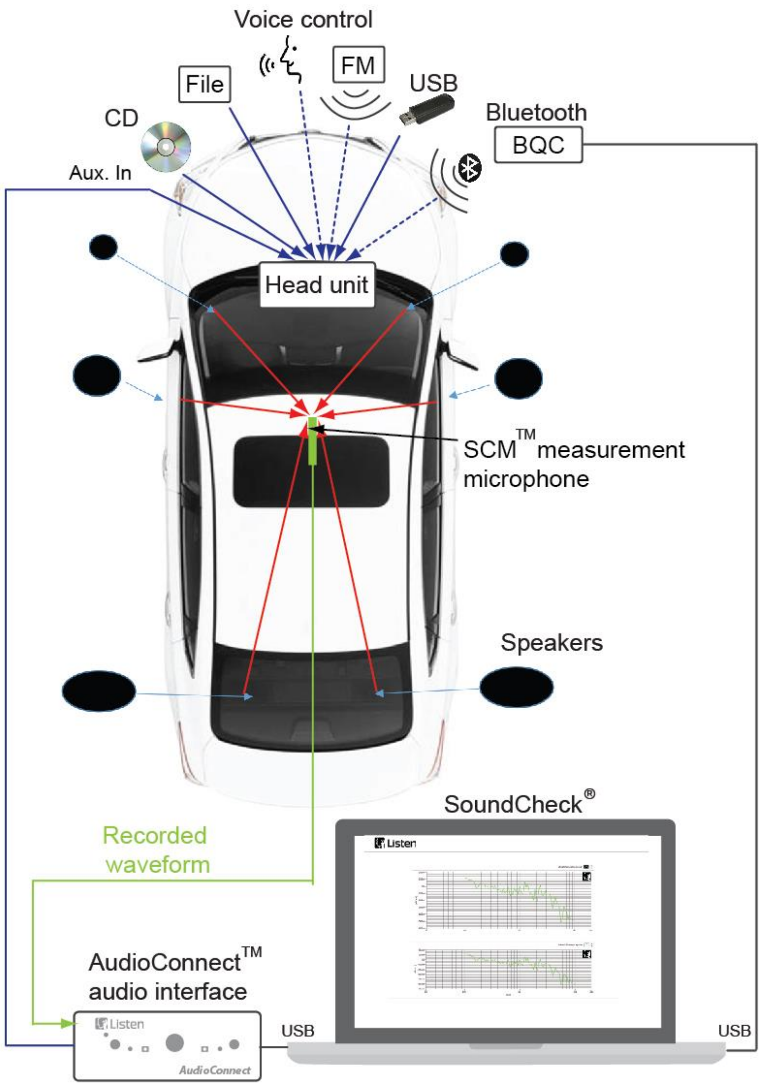
1. Analog
   – Wired (if available) typically found in older model cars with auxiliary input

2. Bluetooth
   – Most cars have a Bluetooth interface

3. Voice Interaction
   – Control the head unit via built-in voice recognition or cloud services e.g. Siri, Alexa, and Google

4. Smartphone
   – Interface to just about everything
   – This is where the industry is moving, e.g. GM just announced they are building Google into their 2021 infotainment systems

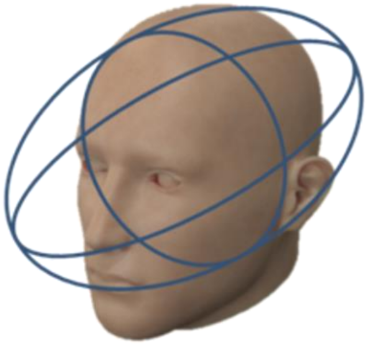
Many Possible Audio Interfaces



Auxiliary-input jack is almost **GONE** so use:

- CD – high quality but too clumsy and probably gone
- USB thumb drive – high quality but inconvenient
- Radio – convenient but low quality and possible RF interference
- Bluetooth – convenient but medium quality
- A2B - Automotive Audio Bus for digital devices
- File stored on Head Unit – high quality and convenient

Car Audio Tuning (Typical)



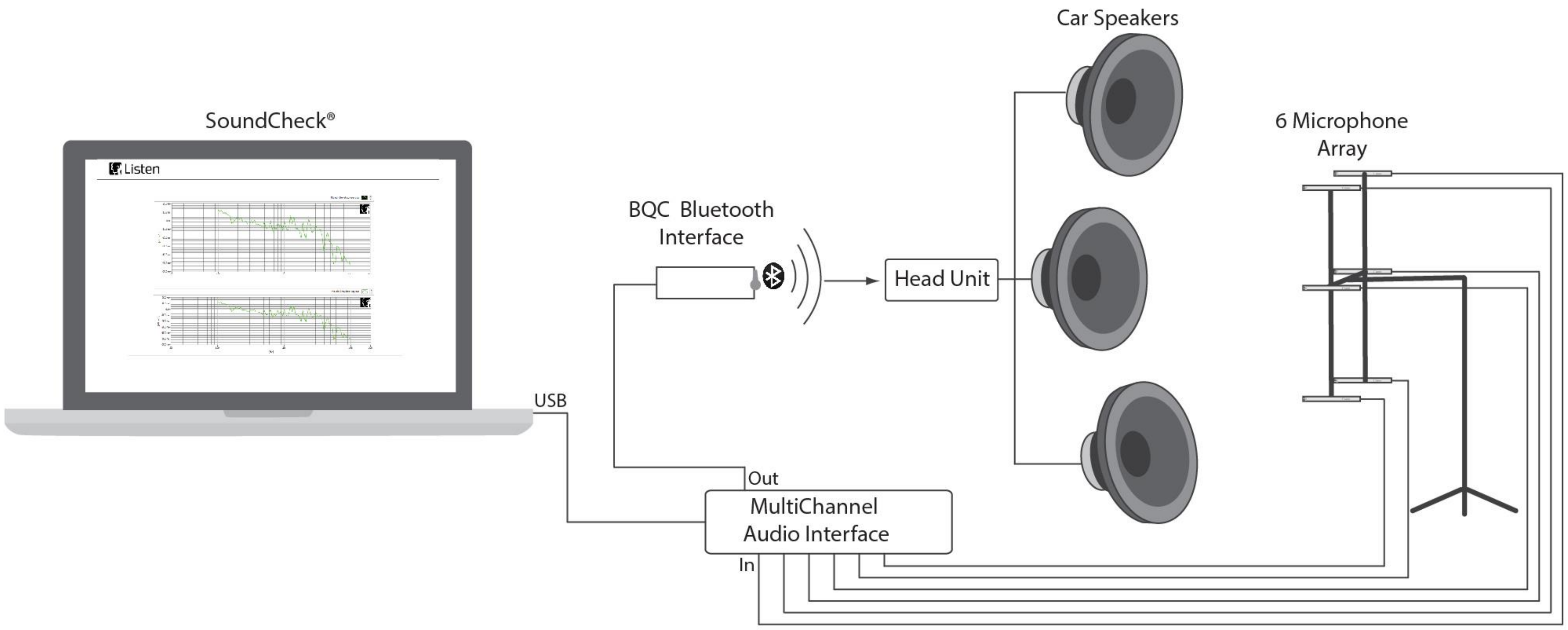
- Purpose: Create a ‘flat frequency’ response at listener position (typically driver)

- Play broadband stimulus out of all speakers simultaneously from head unit

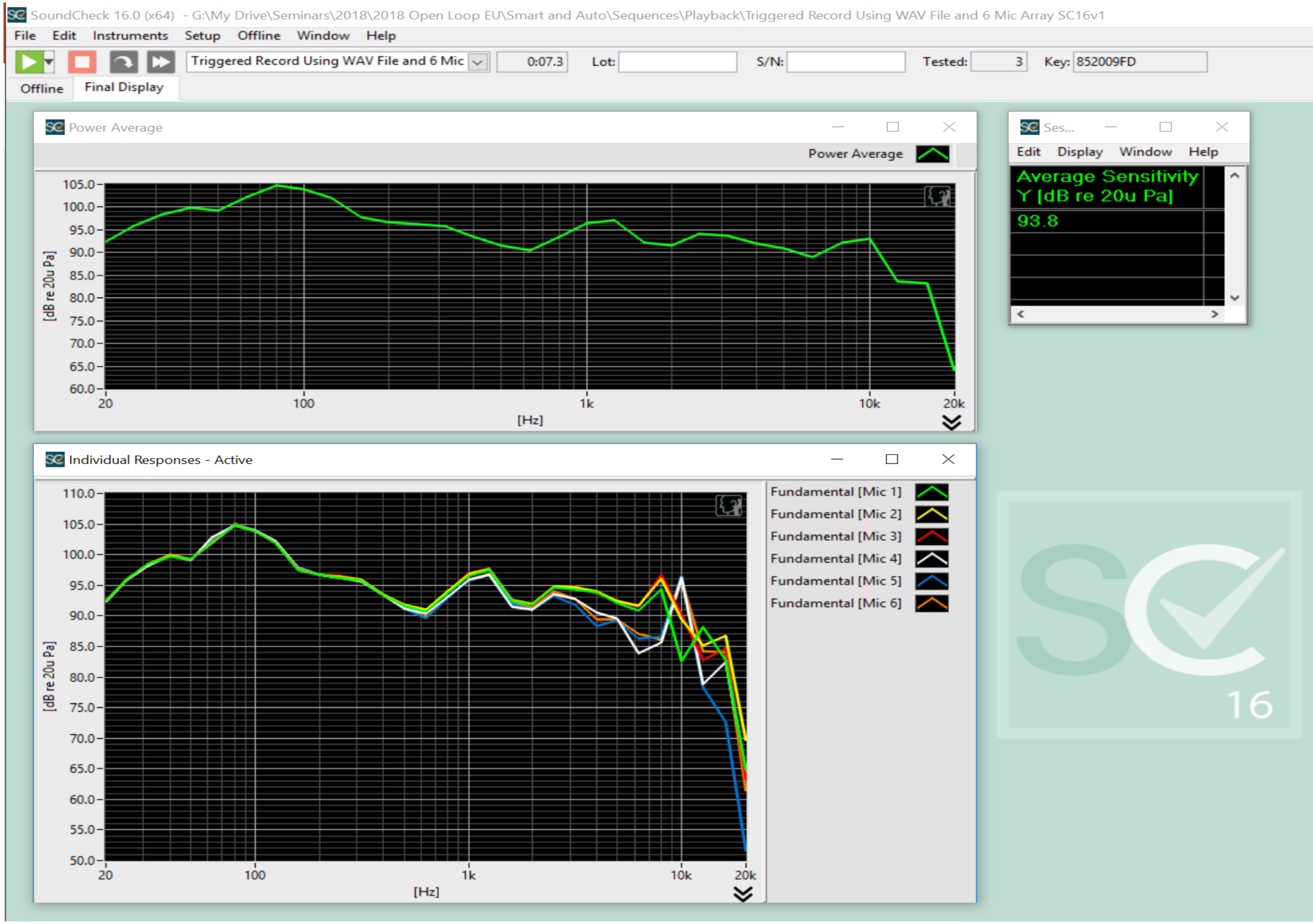
- Use a 6 microphone array where microphones are positioned in 3 seated positions (95th, 50th, and 5th percentile)
  - No standard positions!

- Spatially average 6 microphones and use the ‘inverse’ curve to tune the audio system

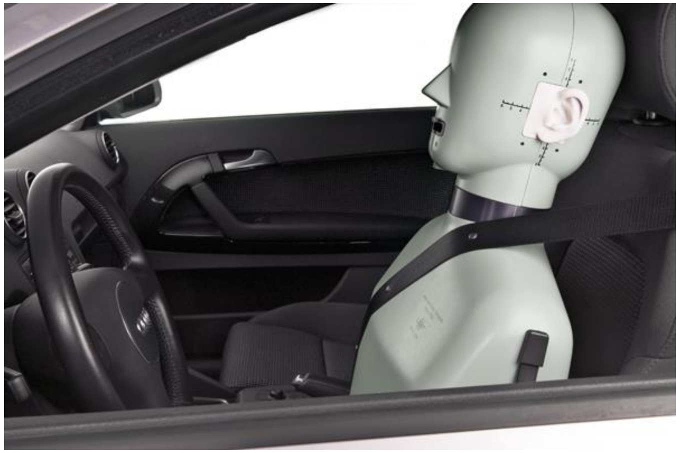
Test configuration for basic wireless (Bluetooth) in-vehicle loudspeaker measurements



Car Audio Tuning (Typical)

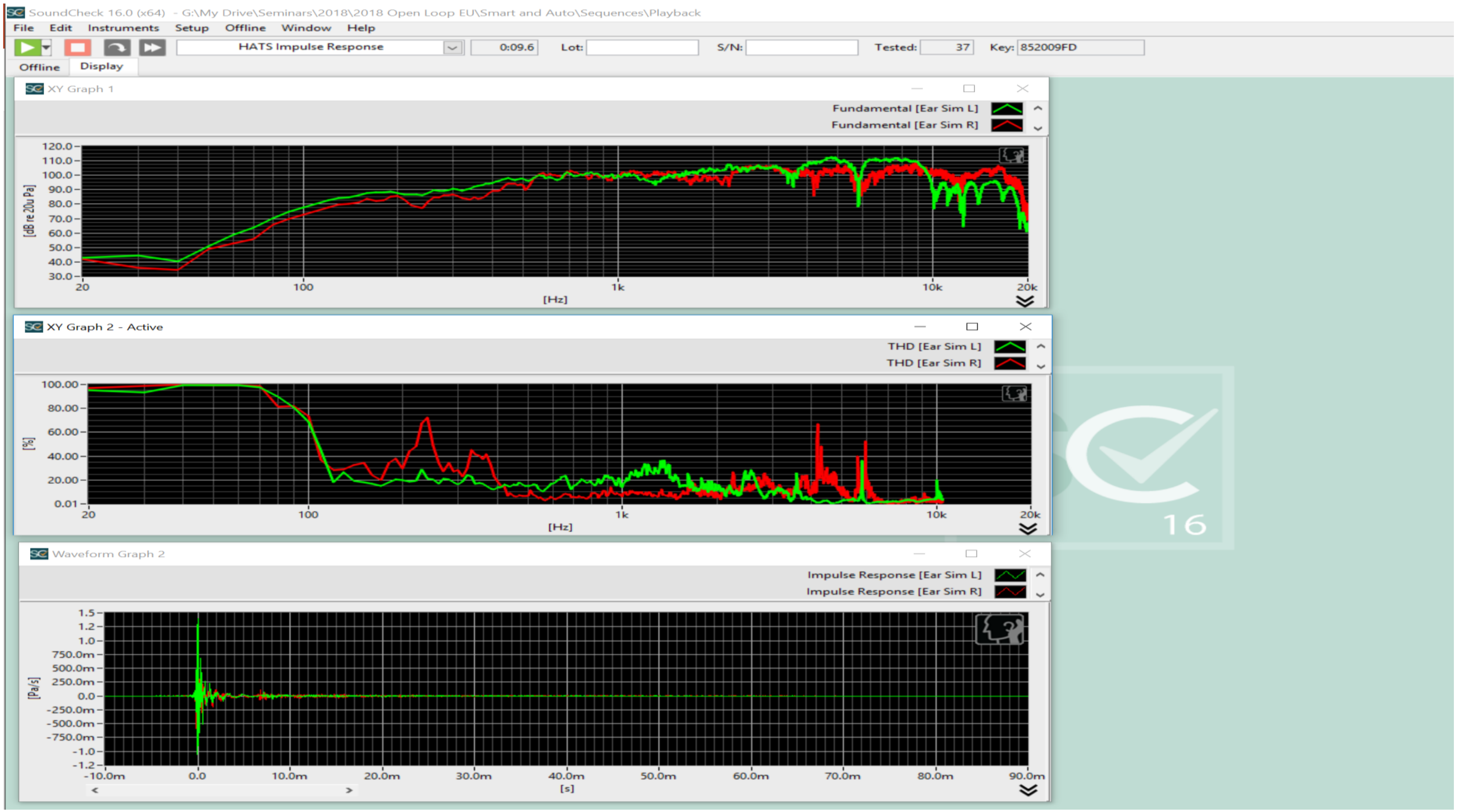


Car Audio Tuning (Advanced)

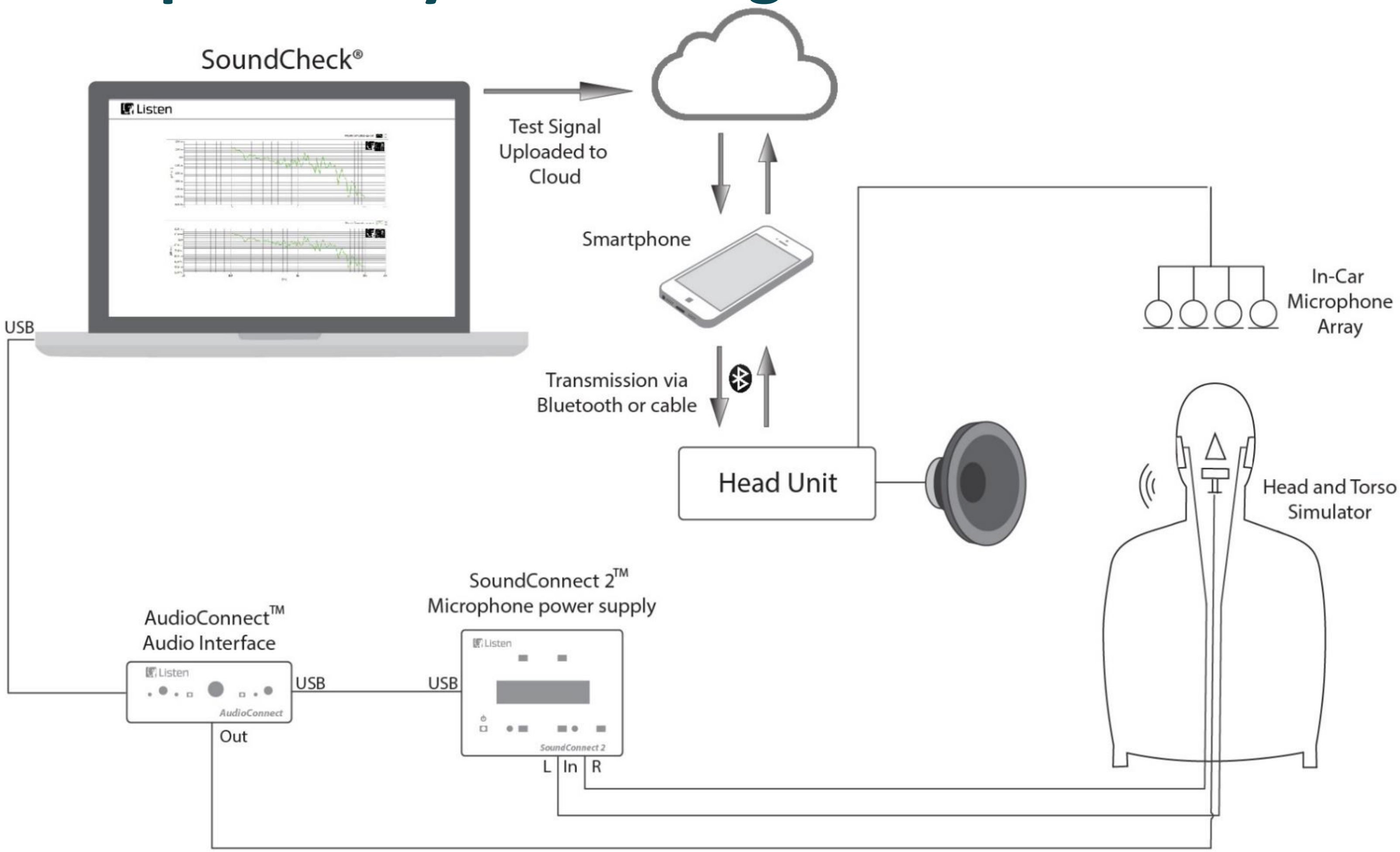


- Purpose: Capture vehicle cabin acoustic response at listener positioner. Typically used for offline simulations.
- Play broadband stimulus out of all speakers simultaneously
- Measure the response at the left and right ears of HATS
- Calculate Impulse Response for each listener position
- Distortion measurements??

Car Audio Tuning (Advanced)



Test configuration for testing infotainment speaker system using voice activation



Test via Voice Interaction

Control the DUT via voice commands

- Automated testing possible by recording and playing back voice commands.

Load test stimulus as music tracks

- Each voice service has it's own music playback system with different capabilities e.g. Apple & Amazon Music
- Uploaded audio is probably compressed, you are not necessarily playing back the signal you uploaded.

Retrieve recorded response via history function (if available e.g. Amazon Alexa app)

- By design, many smart speaker systems do not provide the ability to retrieve your voice command history.

Voice Interaction Speaker Measurement Steps

1. Create activation phrase

2. Upload test stimulus to cloud service
   – The example cloud music service would not accept a linear .wav file. Conversion to 320 kbps MP3 was used.
   – The transparency of the MP3 was tested by encoding and decoding the file and comparing them using Transfer function analysis.

3. Play activation word via mouth simulator and capture response with reference microphone

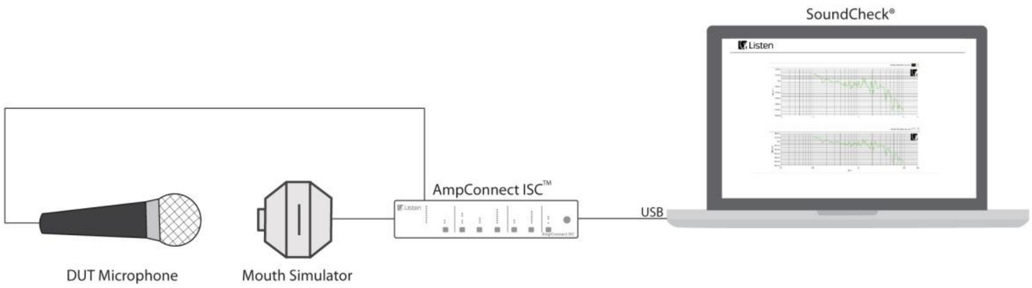
4. Analyze captured response signal

Advanced Test Details

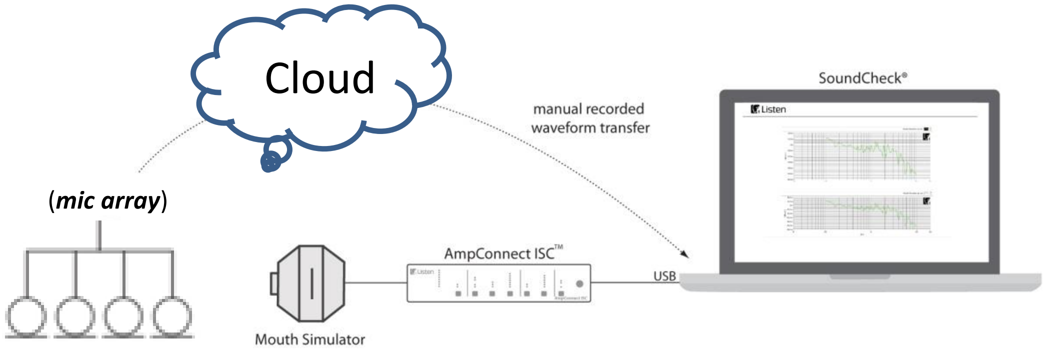
- What does “Open Loop” mean?
- Resampling and Frequency Shift
- Triggering and Windowing
- System Step

Closed Loop vs Open Loop Microphone(s) Measurement

What we are used to:



VS.



What's becoming more common.

Closed Loop vs Open Loop

Closed Loop Audio Test: Uninterrupted signal path from Output to Input

E.G. Audio Interface (Sound Card) -> Mouth Simulator -> Microphone -> Audio Interface

Properties:
- ✓ Input and Output are in the same domain (analog)
- ✓ Input and Output are synchronous (identical sample rate)

VS.

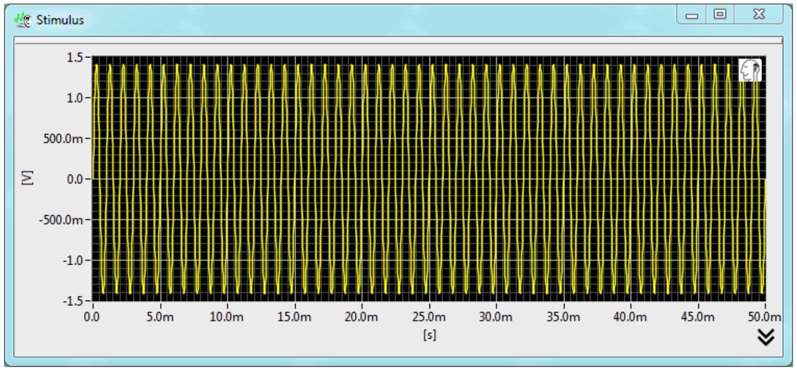
Open Loop Audio Test: Input and Output are not directly connected

E.G. Audio Interface (Sound Card) -> Mouth Simulator -> Microphone -> .Wav file

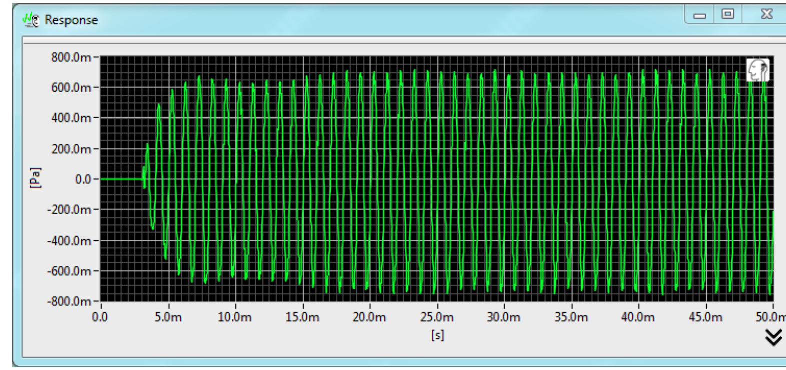
Properties:
- ✓ Input and Output may be in different domains (Output=Analog, Input=Digital)
- ✓ Input and Output are asynchronous (different sample rates)
- ✓ Variable delay between input and output

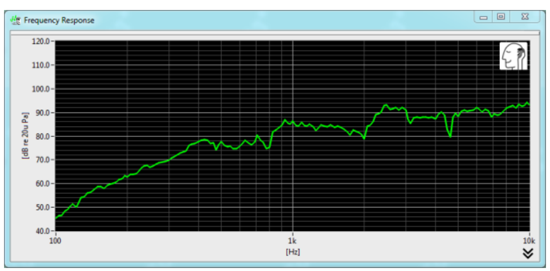
SoundCheck Audio Analysis does not require synchronous play & record

Stimulus

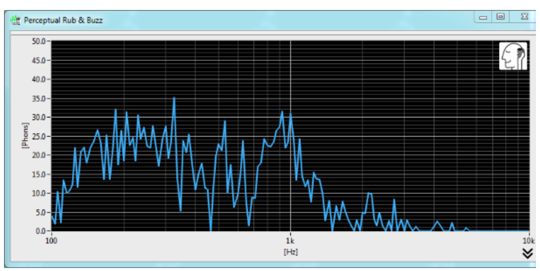


Response

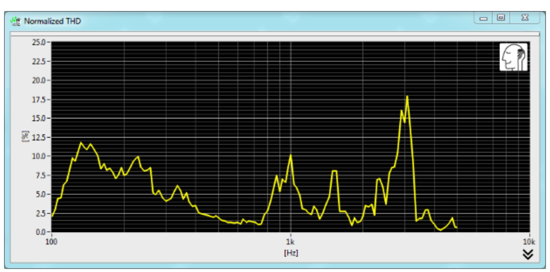




Frequency Response



Perceptual Rub & Buzz

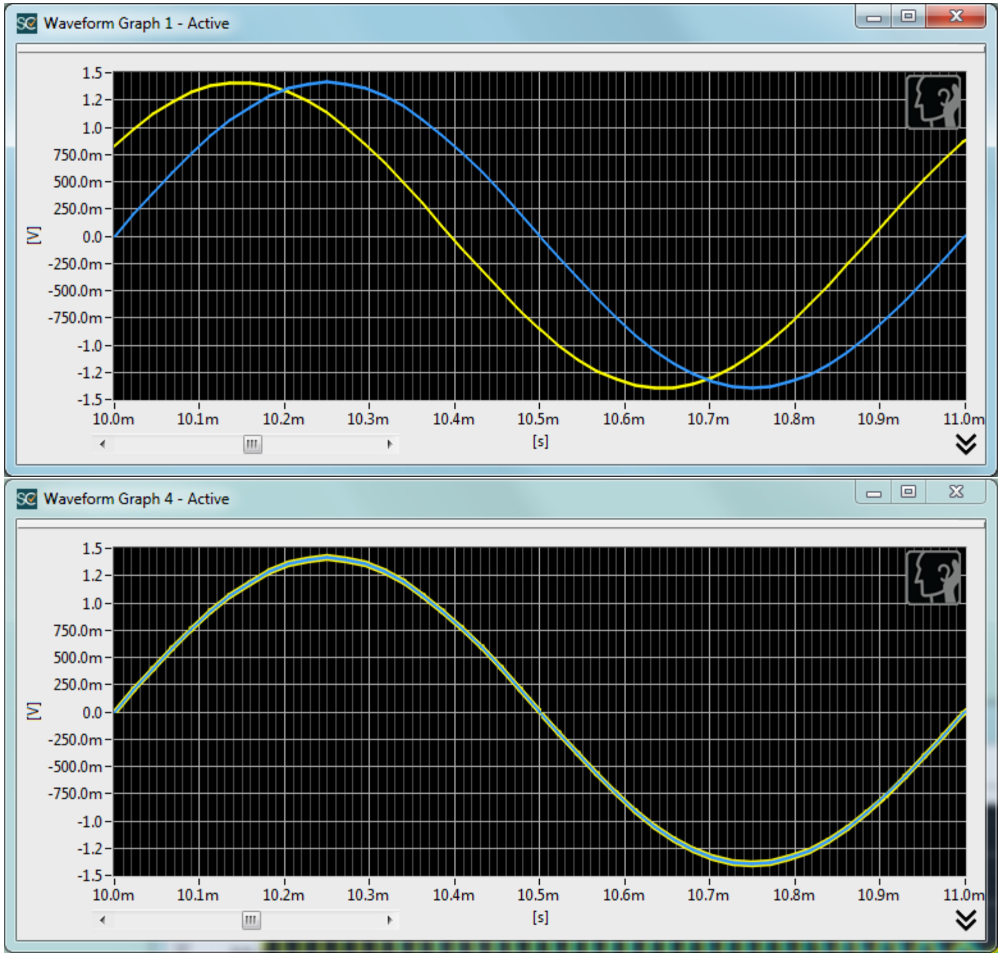


THD

Frequency Shift

With Open Loop tests there is almost always a sampling frequency error.

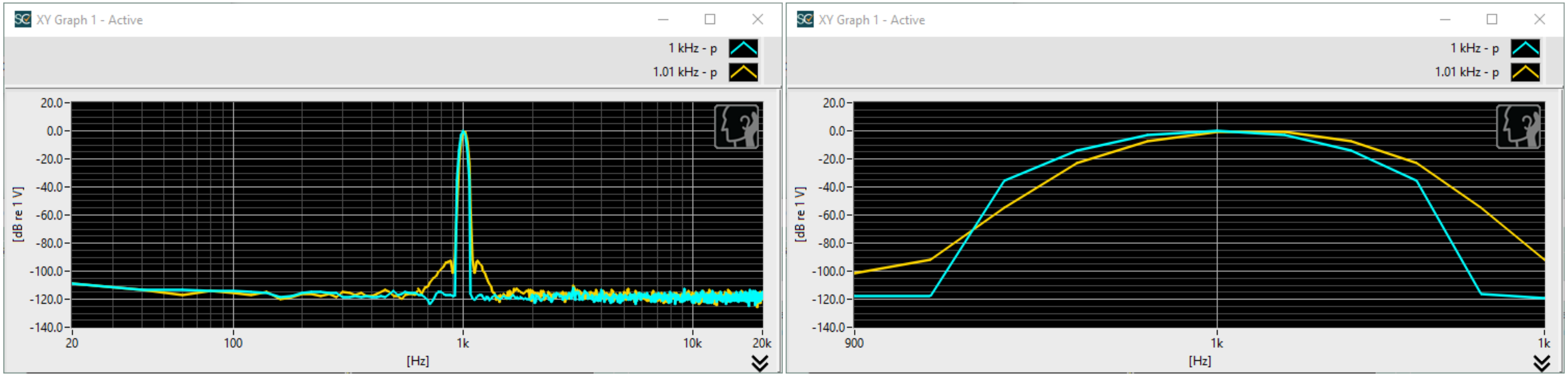
- ✓ Different devices have different clock crystals (no master clock)
- ✓ Will never produce ***EXACTLY*** the same frequency
- ✓ Uncorrected, this will lead to measurement errors (phase in particular)
- ✓ Post-processing, Frequency Shift re-aligns the waveforms in the time domain to correct for this error



Frequency Shift

How does it work?

Calculates the true center frequency using a curve fit:

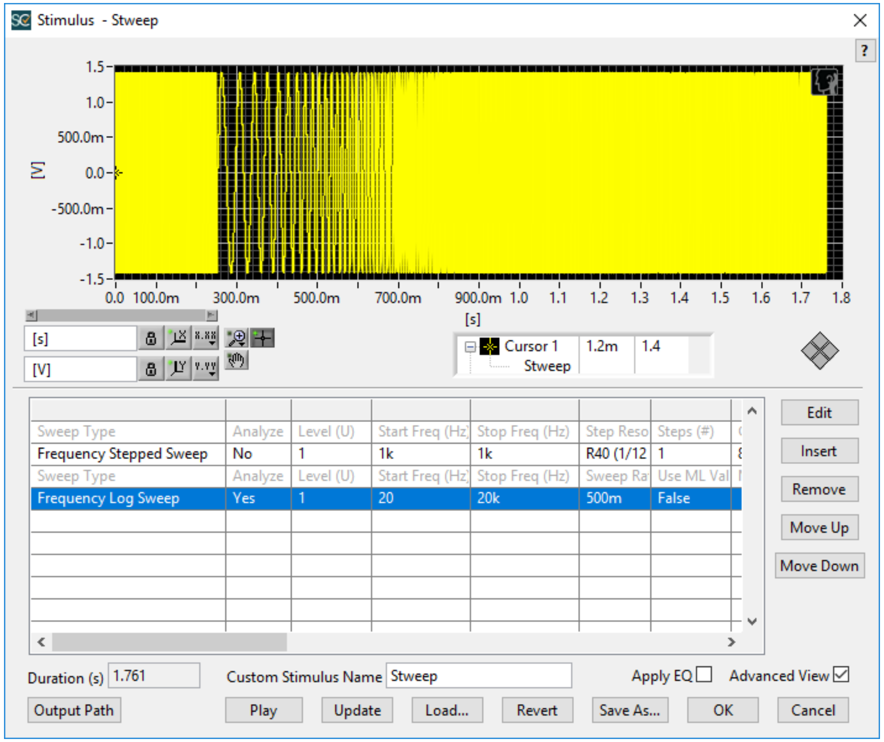


Can resolve frequencies below the FFT resolution.

Only works with sine waves!

Frequency Shift

Prepend a 1 kHz, 250ms tone to your stimulus when using frequency shift.



- ✓ Provides a solid signal for Frequency Shift to lock on to
- ✓ Provides a trigger tone for triggered record or Intersection
- ✓ By pre-pending a fixed tone, you can Frequency Shift logTSR, noise, speech and other non-sinusoidal waveforms.

Triggering and Windowing

- When testing open loop devices capturing the response waveform can be challenging.
- The response waveform needs to be isolated from other signals.
  - e.g. Voice feedback from the DUT.
- Triggered record only acquisition are a solution for playback/speaker tests.
- Windowing can be used to cutout the response waveform from a larger recording in microphone test.

Triggering

Setting the trigger frequency and threshold:

- Finding an appropriate trigger level can be difficult:
  - Too low and ambient noise will false trigger the acquisition
  - Too high and it will never trigger
  - Ideal is above the ambient noise and below DUT output level
- Triggering on a frequency tone and level makes it more robust to background noise and false triggers
- To find ideal trigger level, use multimeter and spectrum analyzer VI in Max mode. Remember, trigger is in Peak value (must multiply RMS result by 1.414 to find peak).

In-vehicle Microphone Testing

- **Handsfree**
  - Many vehicles come with a handsfree system and microphone selection/placement is critical for system to perform well
  - ANC is also becoming more common in vehicles
- **SNR**
  - Make sure that the Signal to Noise (SNR), frequency response, and distortion of the microphone are known
- **Required by:**
  - OEM
  - Tier 1
  - Tier 2

Typical Test Methods

1. Place HATS in seat (always driver, sometimes other positions)
2. Play stimulus out of HATS mouth and measure response at microphone
3. Calculate SNR, frequency response, distortion, and noise floor
4. Optionally…evaluate directionality of cardioid or beamformers
5. Optionally…test to ITU P.1100 / P.1110 / P.1120 narrow and wideband hands-free communication standards
6. Optionally…measure Impulse Response function from HATS mouth -> HF microphone

- **<u>SoundCheck Strengths</u>**
  - OPEN LOOP!
  - Many possible I/O combinations:
    - Use any commercial HATS (Bruel & Kjaer, HEAD acoustics, GRAS)
    - Bluetooth to radio
    - MEMS to microphone
    - Automotive Audio Bus (A2B) for digital devices
  - Can perform classic speaker tests as part of an integrated sequence

Demo Time!

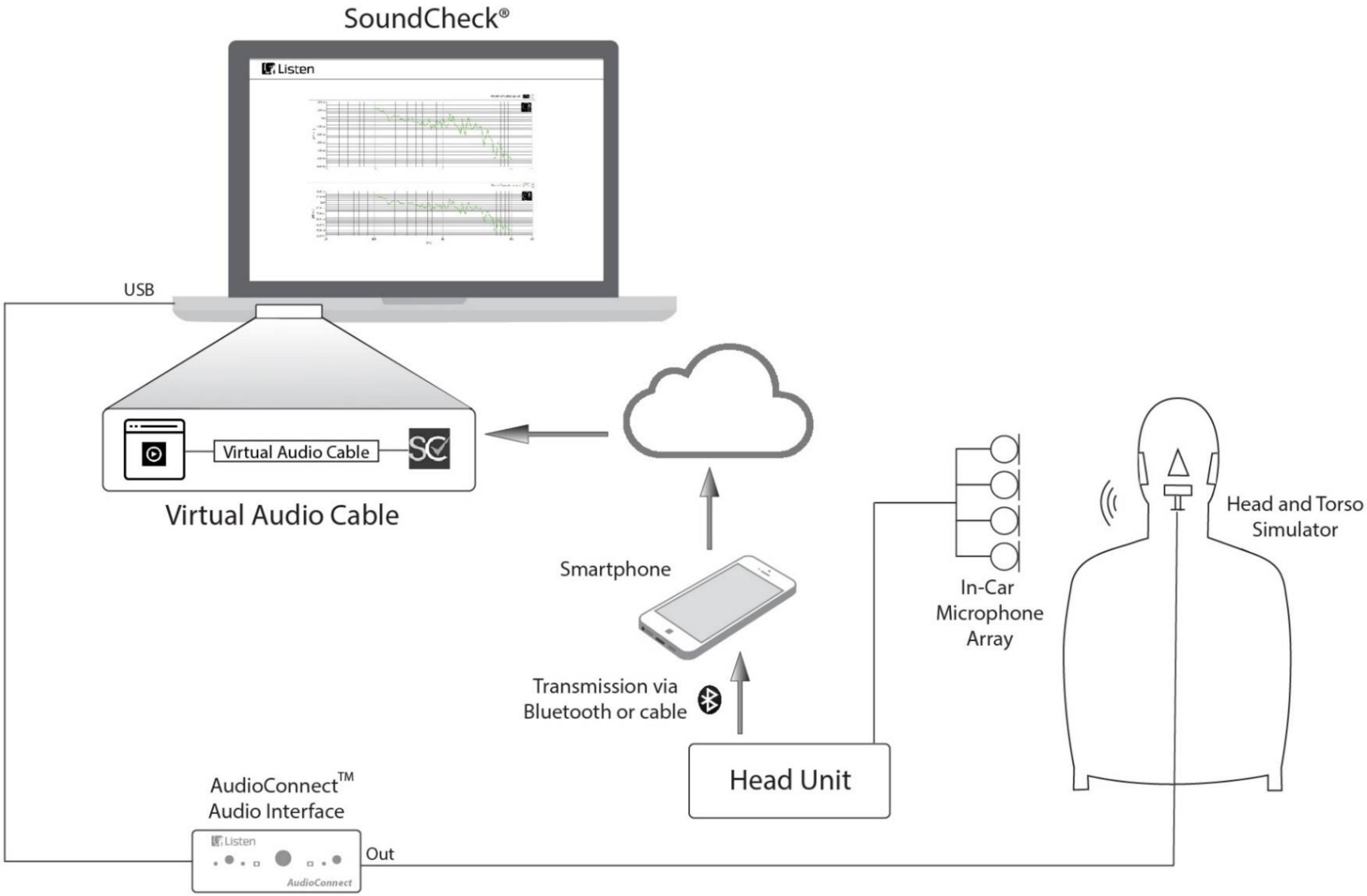
Voice Interaction Microphone Measurement Steps

1. Combine recorded activation word with frequency sweep equalized at the MRP of the mouth simulator
   – Sample stimulus

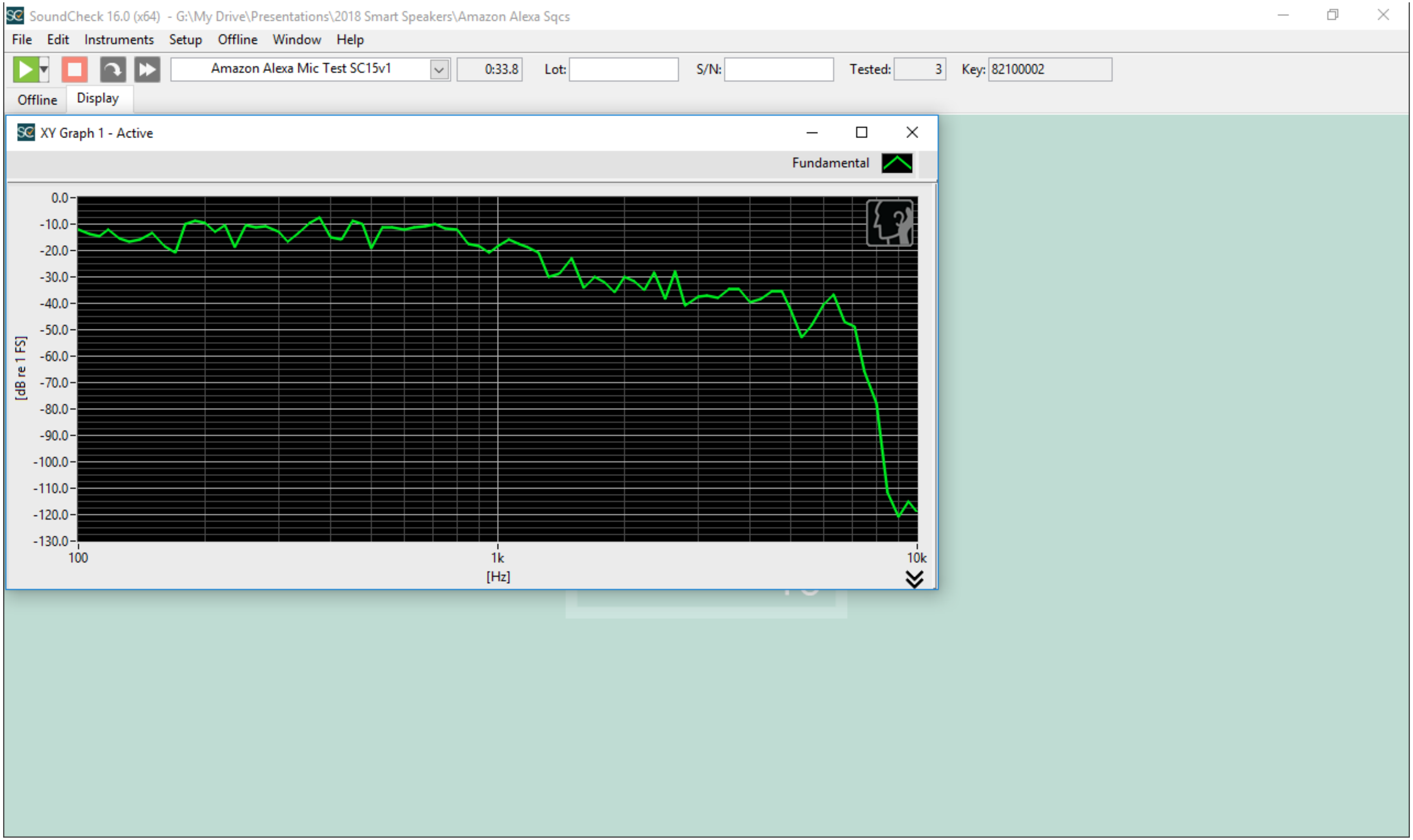
2. Playback stimulus, then download from the cloud service
   – Virtual Audio Cable can be used to capture the playback stream
   – Or Recall Step can be used to load captured recording

3. Analyze captured response recording

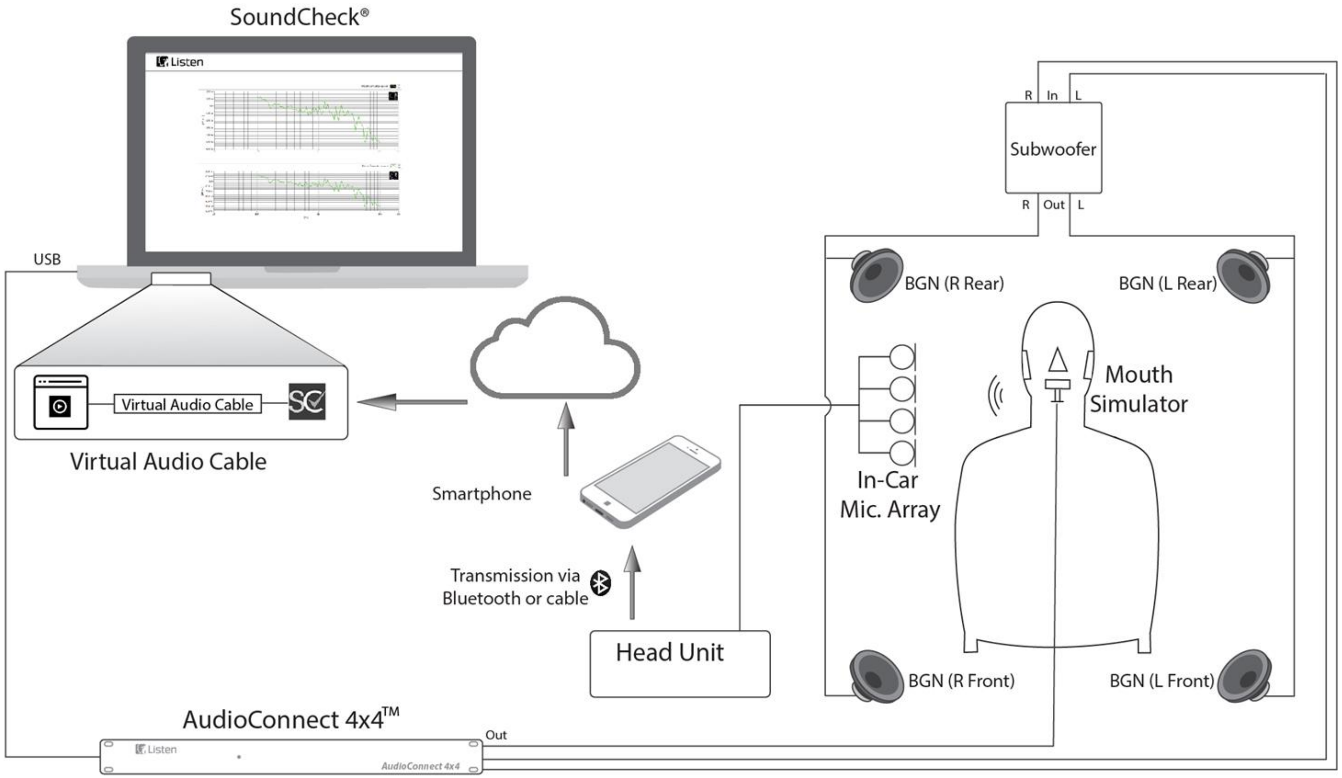
Microphone(s) measurement using voice activation and virtual audio cable



Voice Interaction Microphone Measurement demo

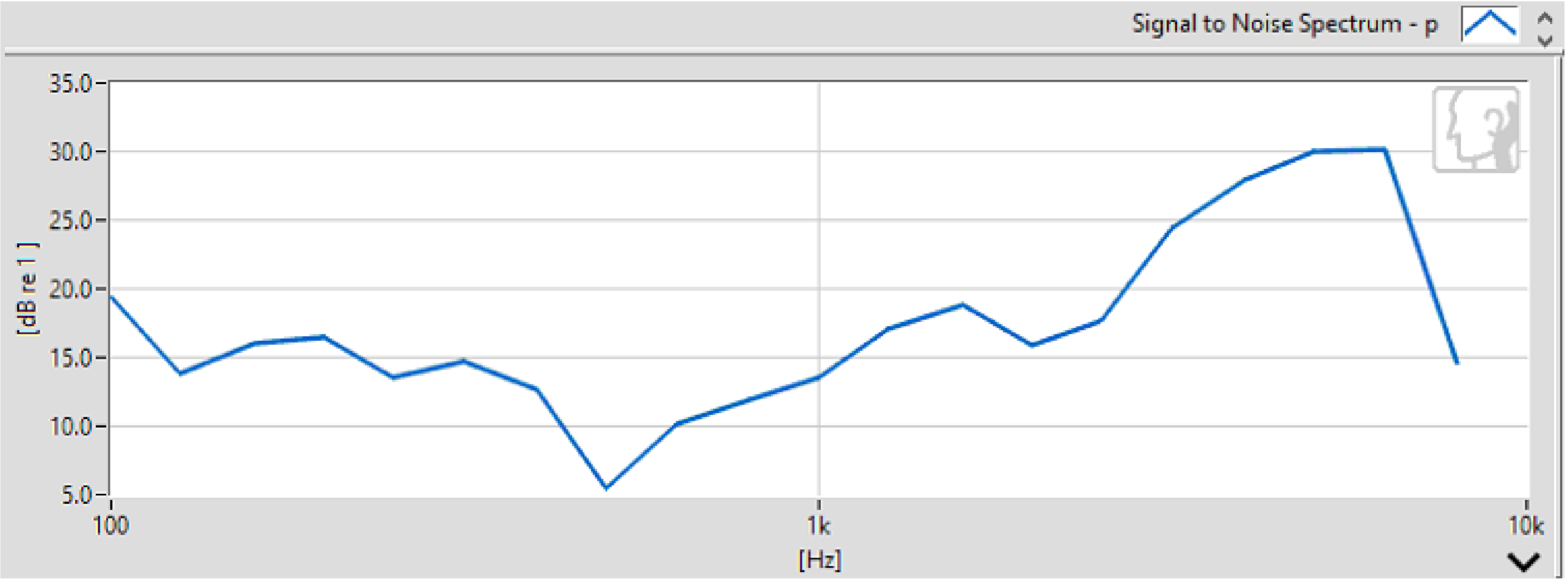


Background Noise System



- Automatically calibrates a standardized 4.1 speaker / subwoofer, providing an equalized, calibrated playback solution to stress your device in a standardized and repeatable way.
- The sequence includes a library of real world binaural recordings from the ETSI ES 202 396-1 Standard (cafeteria, pub, vehicle, etc.), and custom or user-defined binaural recordings may also be used.
- Applications of this sequence include evaluating;
  - ANC, noise suppression, voice recognition testing, SNR optimization of microphones, beamforming directionality studies of microphone arrays and more.

SNR for a microphone array measured with ETSI standard background noises

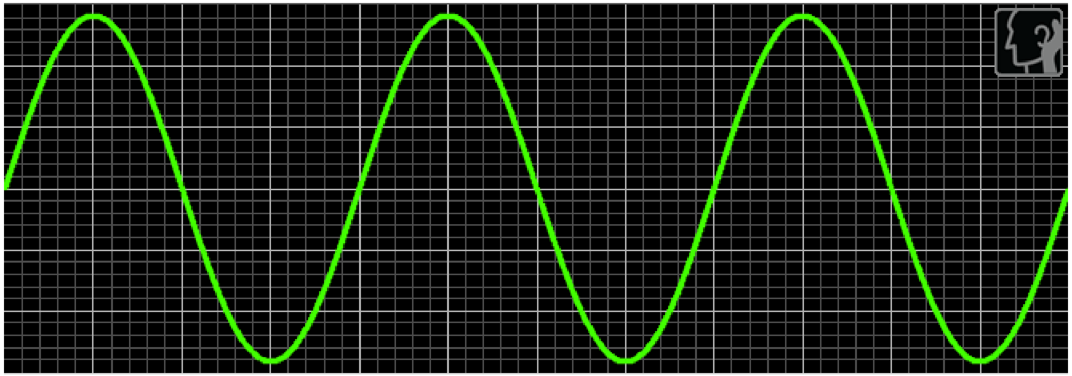


Test Signals

- In theory, the test signals applicable to normal speaker and microphone testing are equally applicable to smart headphone testing but…
  - Playing back the signal may require encoding via a lossy CODEC.
  - Microphone signal processing may be biased toward speech and/or use noise cancelling algorithms and/or speech activity detectors that preclude testing with sine waves.
  - Broadband test signals like pink noise allow viewing realtime changes when controls are changed

First, what is special about sine waves?

- The ***ONLY*** signal type that has all of it's energy at one frequency:



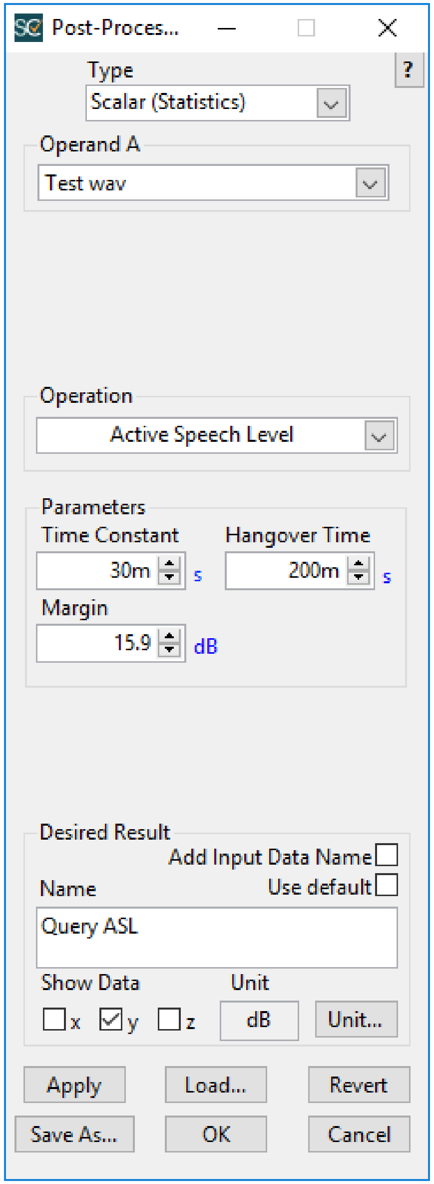
- Easy to analyze including harmonic distortion and noise
- Best signal to noise ratio of any test signal

Testing Using Speech & Music

- More realistic customer use case
- Signal processing inside Smart Device is designed to interact with voice and play music
  - e.g. voice activity detectors, mic array beamforming, sense acoustic environment, loudness control, equalization, compressors, etc.
- Use Active Speech Level for calibrating test level of voice activation and measuring speech playback
- Use Zwicker Loudness for measuring music level playback

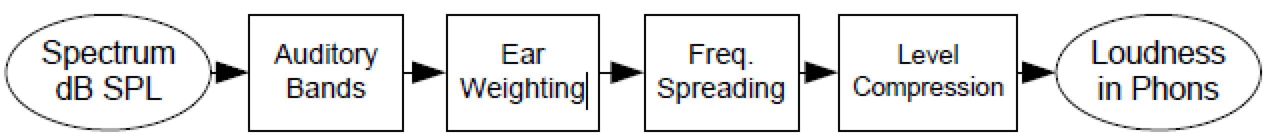
Active Speech Level

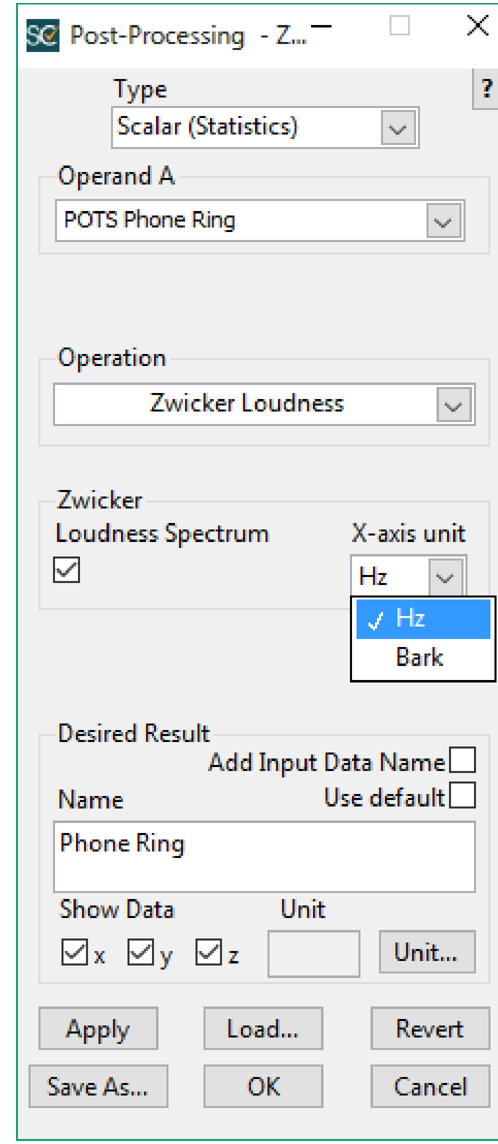
- Used with a speech signal, this operation evaluates levels for only the parts of the waveform where speech is actually present.
- Silent gaps are excluded but short interruptions that are part of continuous speech are included
- It is widely used in telephony applications, i.e.: testing to ITU-T P.56 05/93 Method B
- **Time Constant (Sec)** - Time constant of exponential averaging used to smooth the envelope of the speech signal
- **Hangover Time (Sec)** - Allowable time for silence during active speech. Longer silent gaps between active speech sections are ignored and left out of the calculation
- **Margin (dB)** - Difference, in dB, between threshold of activity and active speech level. When the level of the background noise is high, the margin can be reduced in order to exclude the noise
- The Active Speech Level of the WAV file can be set in the Stimulus Editor



Zwicker Loudness

- Zwicker Loudness calculates the overall perceived loudness of a sound
- Uses a psycho-acoustic model which takes into account the nonlinearity of the human ear to sound at different frequencies and levels
- Measure the perceived loudness of complex sounds, e.g. music





Speech Recognition

1. **White Box** – Direct comparison to automatic voice recognition algorithm output.

2. **Fully manual** – Use command phrases and evaluate response.

3. **Semi manual** – Use command phrases and compare to AVR output online.

4. **Automated** – Create specially titled music playback tracks, issue appropriate command phrases and test if correct track is played back.

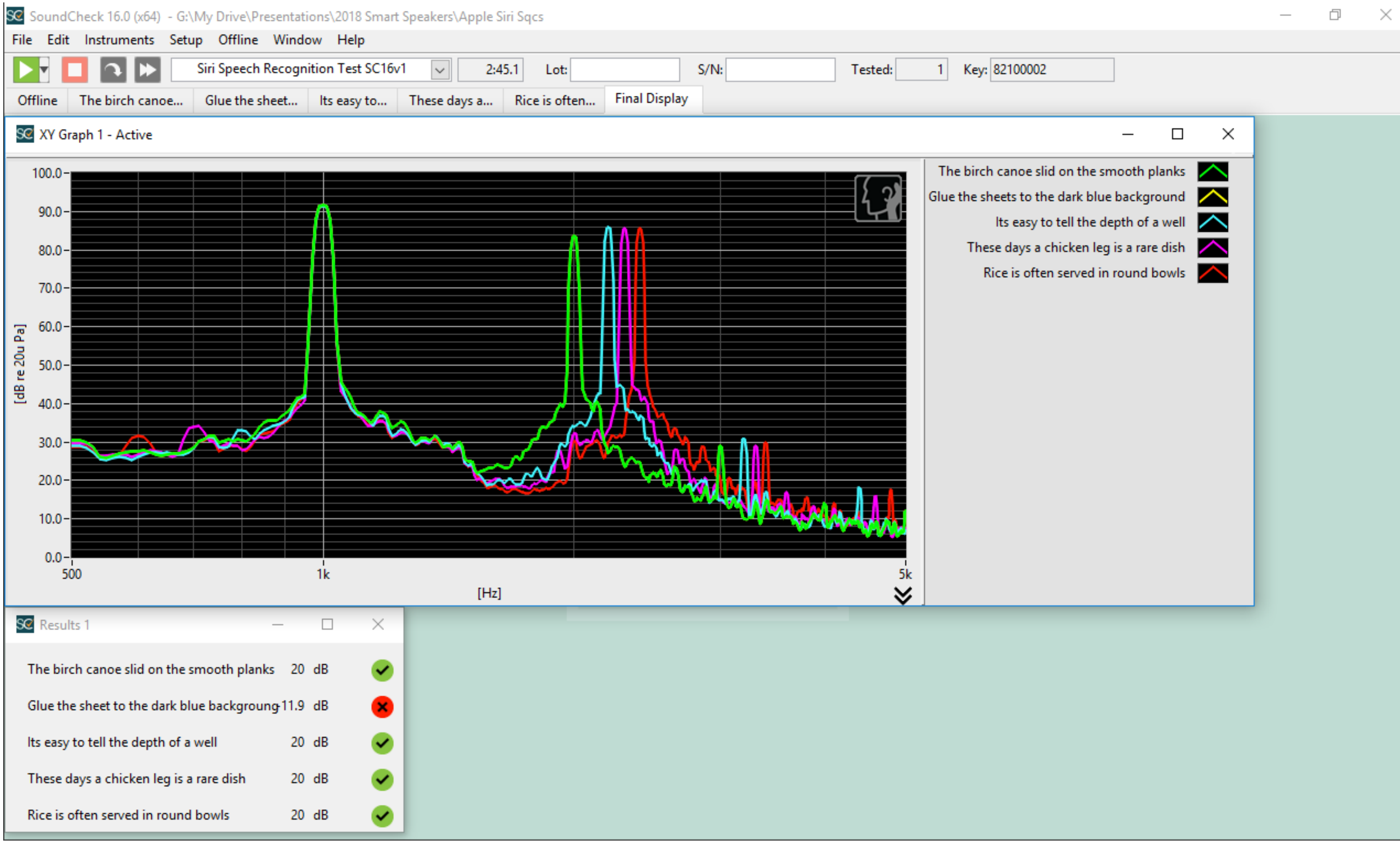
Automated Black Box Speech Recognition Test

1. **Upload a collection of special musical tracks**
   – The track titles are the words or phrases you want to test
   – The actual content of the tracks is a single, dual, or multitone that allows the track to be identified by it's audio content.

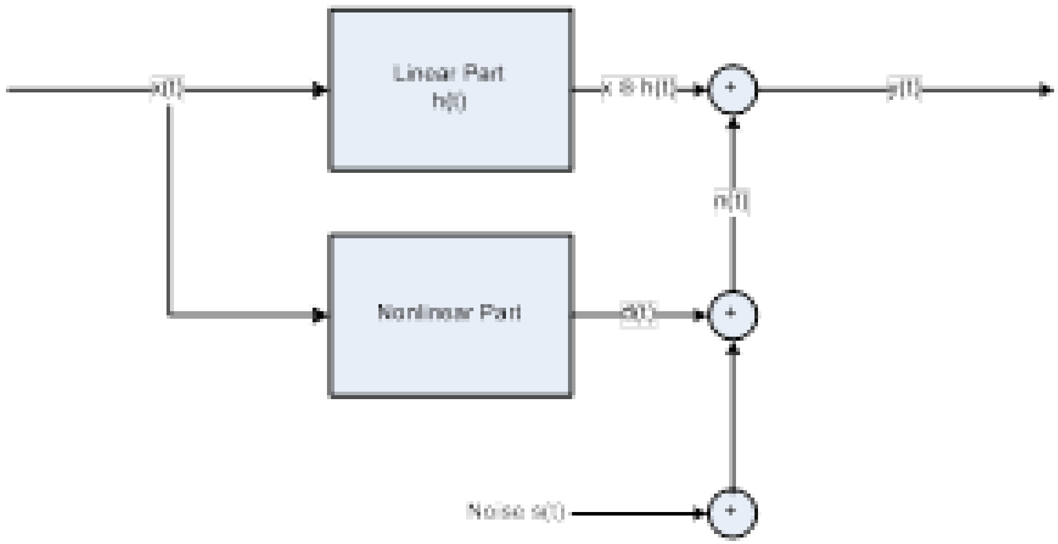
2. **Record people speaking the test phrases as requests to play back the specific tracks.**
   - Harvard Sentences - a collection of phonetically balanced sentences that use specific phonemes at the same frequency they appear in English – are used for the track titles.

3. **Capture the response of the device and use a limit to detect whether the correct, or any signal, was played back.**

Automated Black Box Speech Recognition Test



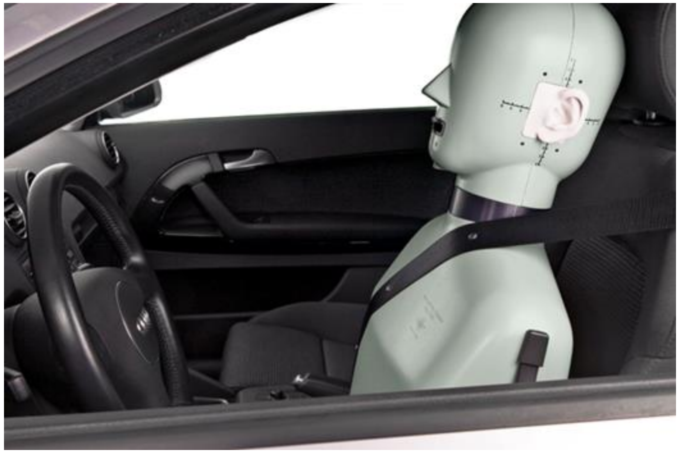
How to measure distortion with real signals?

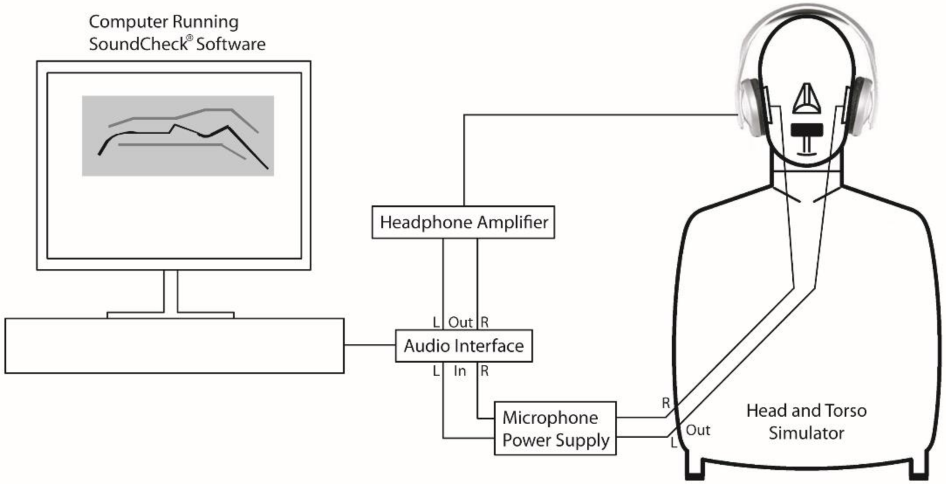


Non-Coherent Distortion is a normalized cross-correlation measurement that determines the degree to which the system output is linearly related to the system input (AES 121st preprint 6877)

- Works with any broadband signal e.g. music & speech

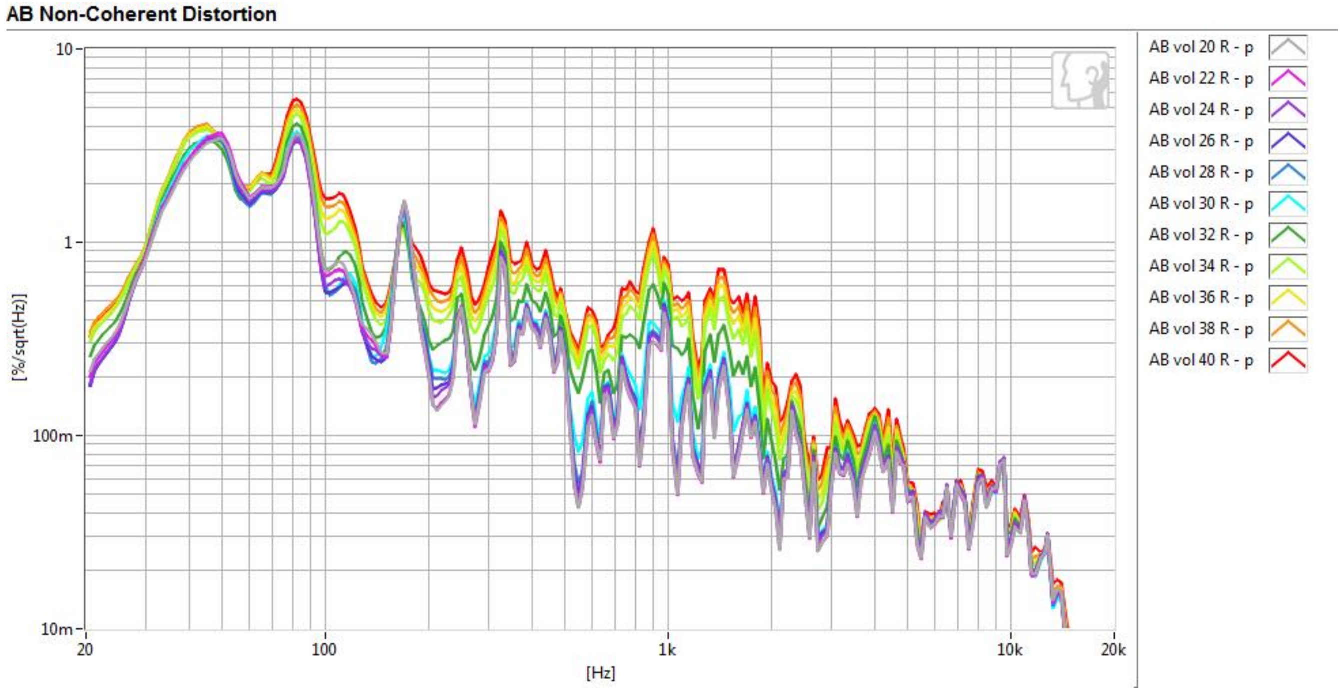
NCD Measurement Setup





- Binaural recordings were made from volume step 20 to 40 (maximum) in 2 step increments on head unit
- SoundCheck was then used to match each recording for overall loudness as well as spectrum in the bass region
  - Spectrum matching in bass region needed to offset the effect of dynamic loudness in the head unit
- Recordings at different levels were virtually played back over high quality & low distortion headphones at the same level e.g. 80dB SPL
- The adjustments helped make the distortion artifacts the main distinguishing difference between the recordings

NCD Distortion (American Boy) vs. Level



NCD with Estelle music as the test signal. The average SPL for volume level 30 was a loud 105dBC!

- Distortion curve shapes are similar with the previous songs but the distortion jumps at volume level 24 and above. This is probably due to the higher recording level (see next slide)

Questions?

Thank you

*[For more details, please read my AES Automotive Conference paper]