R&D Stories
EarPrint
Precision Listening for Earphones and Headphones
By Mike Klasco

You Can DIY!
Cherry Bomb
A VFET SET Amplifier
By Thomas Perazella

Sound Control
Assessing Audio Quality Using the MATT Test
By Richard Honeycutt

Sound Engineering
Head Space—A Binaural Recording Session
By Neville Roberts

Audio Electronics
A Sources 101 Update
By Walt Jung

Audio Electronics
Current Feedback and Voltage Feedback Fallacies
By Michael Kiwanuka

It’s About the Sound Loudness Revisited
By Ron Tipton

Questions & Answers An Interview with Dave Thomson, Raven Audio
By Shannon Becker
We humans have been paying attention to optimizing sound in rooms for a long time. The ancient Greeks and Romans took steps to optimize speech intelligibility when building theaters. But probably the earliest era in which specific attention was paid to good room design for music performance was in Baroque times (J. S. Bach and contemporaries).

Modern acoustical engineering started with the work of Wallace C. Sabine on Harvard University’s Fogg Lecture Hall. Sabine quickly recognized that the wretched lack of speech intelligibility resulted from excessive reverberation, so he developed methods for measuring and controlling reverberation. Thus, reverberation measurement and control became the first technological tool in the modern acoustician’s toolbox. Reverberation affects speech intelligibility and musical articulation, meaning that concert halls designed for sustained Gregorian Chants muddled the sound of highly articulated Baroque or Classical music. Sabine applied his “reverberation tool” later when he was the acoustic designer for the Boston Symphony Hall, and he recognized the need for sound to be evenly distributed throughout the seating area.

**Early:Late Energy Ratio**

In large rooms such as concert halls and theaters, the only significant frequency-dependent acoustical effect is the variation of acoustical absorption with frequency. There is no reason for anything like a “frequency response” measurement in these rooms. But in the latter part of the 20th century, other acoustical parameters were identified that have as much or more effect on music and speech articulation. These fell into the category of “early:late” energy ratios.

**Figure 1** shows the impulse response (IR) of a somewhat unusual listening room. This IR shows the decay of acoustical energy over time. An example of an early:late ratio is “Clarity index” or “C80.” C80 is the 10× the log of the ratio of total energy contained in the first 80 ms after the onset of an impulsive sound to the total energy after 80 ms. Clarity is specified in decibels and the acceptable range is about +1 to -4 dB, although the best concert halls have C80 between about -1 and -4 dB.

A related early:late energy ratio is “Definition” or “D50,” which is the ratio of energy in the first 50 ms to total energy. Definition is specified in percentage and correlates more closely with speech intelligibility than with music articulation—C80 is more applicable in evaluating a room for music.

**Speech Intelligibility**

After the introduction of early:late energy ratios to the evaluation of auditorium acoustics, quantitative evaluation of speech intelligibility became an area of great interest. The first speech intelligibility metric in common use was called “percentage Articulation Loss of Consonants” or “%Alcons” and was introduced by V. M. A. Peutz.

Learn how to use Musical Articulation Test Tones (MATT) to identify audio quality issues and obtain the highest quality sound possible from the components in your listening environment.

By
Richard Honeycutt
(United States)
in 1971.[1] The basis of determining intelligibility was the ability of listeners to correctly distinguish consonants pronounced by a talker in the room. A standardized list of test syllables was used. Peutz developed a method of predicting %Alcons from the reverberation time (RT) and background noise level in the room.

Tammo Houtgast, Herman Steeneken, et al., introduced a method of specifying speech intelligibility using the Modulation Transfer Function (MTF), which quantified the effect upon speech articulation produced by any defect in the electroacoustical and/or acoustical path between the talker and the listener.[2]

The metric is called the Speech Transmission Index (STI). The underlying theory is that speech can be considered to originate as a continuous wave created by the vocal folds, and then modulated by the throat, lips, and tongue. The “intelligence” or information is carried by the modulation and the depth of the modulation correlates with speech intelligibility. Reverberation, background noise, and other acoustical or electro-acoustical effects can decrease the modulation depth of the acoustic wave reaching the listener’s ears, hurting speech intelligibility.

Note that neither of these specifications of speech intelligibility addresses the naturalness of speech: Both exclusively target the articulation of speech. The sound of a letter “t,” when stretched by reverberation, masked by noise, or electronically distorted, is difficult to distinguish from the sound of a “p” or a “d.” Figure 2 and Figure 3 show an anechoically recorded speech signal and that same signal played in a room having a 1.5-second RT, respectively. Notice how the reverberation closes up the spaces between sound bursts (i.e., decreases the depth of modulation). Speech in the reverberant room would have a lower STI, indicating poorer speech intelligibility.

Small Room Acoustics

In the early part of the 20th century, radio programming frequently originated as live performances in the radio station’s studios, so the sound of the studio and the control room became important. With the advent of home studios and home theaters in recent years, the subject of small-room acoustics has become more prominent.

Contrary to the situation in auditoriums and theaters, small rooms seldom have issues with reverberation. In fact, the smaller dimensions and corresponding shorter delay times in control rooms, studios, and so forth make the use of Sabine’s statistical methods for specifying RT questionable for small rooms.

Small rooms do have two problems of their own: “boxy” sound quality and resonant modes. Unfortunately, these issues are often confused with one another. Boombiness can result from resonant modes, especially when the acoustical absorption in the room is inadequate at low frequencies. However, boxiness can exist along with boombiness or by itself. It is often caused by sound reflecting from surfaces near the speakers, then reinforcing or cancelling the direct sound at specific frequencies. This is called “comb filtering” because the resulting frequency response looks like a comb. Boxiness can be described as the impression of a small, acoustically reflective room we perceive from the aggregate of all sound (direct plus early reflections) within the 35-ms “fusion window” perceived by the ear-brain system. (Thanks to Art Noxon for this definition.) Decreasing the early reflections decreases boxiness.

Small Room Acoustics

In the early part of the 20th century, radio programming frequently originated as live performances in the radio station’s studios, so the sound of the studio and the control room became important. With the advent of home studios and home theaters in recent years, the subject of small-room acoustics has become more prominent.

Contrary to the situation in auditoriums and theaters, small rooms seldom have issues with reverberation. In fact, the smaller dimensions and corresponding shorter delay times in control rooms, studios, and so forth make the use of Sabine’s
Naturally, the first attempts at taming acoustically unruly small rooms used tools already in the toolbox: reverberation control by absorption, sound diffusion by means of barrel diffusers, and later, equalization of the sound system. These methods provided some improvement. In particular, boxiness can be reduced by placing acoustical absorption on offending surfaces. However, even though modes can cause boominess, they also cause articulation problems. Applying a fix in the frequency domain via equalization will not correct problems in the time domain.

At any modal frequency, not only is the level of the sound increased by the modal effect, but the effective RT at that frequency is increased. This is particularly true for pronounced modes having very little damping (acoustical absorption). Such a mode is characterized by a high quality factor or “Q.” Quality factor is the ratio of energy stored to energy dissipated per cycle. Thus, a high Q mode stores a lot of energy in each cycle at the modal frequency. This energy is then released into the room, lengthening the reverberant decay, which in turn muddles articulation. So while a notch filter can flatten the frequency response of the room at low frequencies, it does nothing to improve articulation.

**Measuring Room Effects**

One of the earliest tests tied in an attempt to measure the room’s effect on articulation was the low-frequency square-wave test. The results were not always enlightening, because no speaker can truly reproduce a square wave, although an excellent low-frequency horn comes closest. Next, engineers realized that a series of low-frequency tone bursts would more nearly simulate musical articulation, and would be a test signal that a good speaker system could inject into the room. Figure 4b shows a series of 40-Hz tone bursts recorded in the room whose impulse response was shown in Figure 1. Notice that the tone-burst test shows that the pulse train takes a few cycles to build to full amplitude, and then takes a longer time to decay after the pulse signal ends.

In 1980, Siegfried Linkwitz reported on investigations using shaped five-cycle tone bursts with a raised-cosine envelope.[3] Figure 5 shows shaped tone-burst tests with the frequencies at various proximities to resonant modes. The effect of the modes in extending the decay—filling in the spaces between bursts—can clearly be seen. Figure 5a has a frequency far from any modes. Figures 5b, 5f, and 5g show the frequencies of response peaks. Figures 5c, 5d, and 5e show the frequencies of response notches.

---

*a* Figure 4a: This tone-burst test signal is part of a series of half-second, 40-Hz sine-wave bursts separated by approximately half-second periods of silence. 4b: The series of tone bursts was played and recorded in the listening room.

*b* Figure 5: The shaped tone burst test tones appear at the top of each pair of waveforms; the playback recorded in a listening room is at the bottom. The burst frequencies are: (A) 112 Hz, (B) 73 Hz, (C) 132 Hz, (D) 47 Hz, (E) 81 Hz, (F) 57 Hz, and (G) 141 Hz. (Figures courtesy of Siegfried Linkwitz, Linkwitz Lab, Inc.)
**Figure 6** shows the subwoofer signal of an anechoic speech track, low-passed at 75 Hz, with the same track played and recorded in a room having a 1.5-second RT. Notice that, despite the widely held notion that all transient energy resides in high frequencies, there is substantial low-frequency energy in this signal. Reverberation smears it just as with a full-spectrum signal. For more information on shaped tone burst testing visit Linkwitz Lab (www.linkwitzlab.com/burst-cd.htm).

**MATT Testing**

Although tone-burst testing can reveal articulation degradation by room effects, evaluating the results takes some understanding and experience. Also, a single-frequency tone-burst test provides limited information. Around 1990, Art Noxon of Acoustic Sciences Corp. extended the MTF testing method, using a test tone composed of a swept sine wave gated at 8 Hz with a 50% duty cycle (1/16 second of tone burst followed by 1/16 second of silence). The frequency is swept from 28 Hz to 780 Hz and then back down to 28 Hz (see **Figure 7** and **Figure 8**). This is called the Musical Articulation Test Tone (MATT) signal. It is available for download at www.acousticsciences.com/matt.

Angelo Farina, et al., published a paper on Acoustical Quality Testing (AQT) in 2001.[4] The procedure included MATT testing. It seems to have been intended primarily as a procedure for determining audio quality in automobiles. In 2006, B. M. Fazenda published a paper on using the MTF to measure room acoustics performance.[5] The MATT signal can be analyzed using specific software, but it also produces specific sound signatures that can be identified by listeners. Quoting the developer:

“During room playback, a number of different effects will be audible.

Ta-Ta-Ta-Ta, the sound of an articulate group of tone bursts. Usually there will be some 8 to 10 clean bursts in such a group, lasting about one second. A typical room will have only a few articulate groups of signals in the 75 second test.

Tattle-Tattle-Tattle-Tattle, the tell-tale sound of the room’s double-tongue response. Large spans of the track will have this sound. Notice that the tonal pulse rate is really twice that of the real signal. Too much energy occupies the dwell period of the test signal.

Toodle-oodle-oodle-oodle, the sound of the garbled room. Notice that it is a softer, less impacted sound. It’s close to a slurred double-tongue response.

![Figure 6: The anechoic subwoofer signal for a speech track (a) can be compared with the same track played and recorded in a room having a 1.5-second RT (b).](image_url)

![Figure 7: The complete MATT test signal is shown for both left and right channels.](image_url)

![Figure 8: By zooming in on the time axis, we can clearly see the tone bursts and silent periods.](image_url)

![Figure 9: The MATT signal was played back and recorded in the listening room.](image_url)
Sound Control

Tathump-Tathump-Tathump, is a more accurate presentation of the TA-TA. The “thump” is the turn-on and turn-off transient effects. This subtle transient coloration becomes totally inaudible with anything but articulate room playback. The thump is a damped 45 Hz ringing with only two oscillations of presence following each burst transition.

**Figure 9** and **Figure 10** show the MATT signal played and recorded in the same room used for the tone-burst tests shown in **Figure 4**. The envelope of the recorded signal (see **Figure 9**) shows the low-frequency response of this unequalized room. In the expanded view (see **Figure 10**), it is difficult to distinguish the tone bursts until about 1 second into the test signal, corresponding to about 33 Hz in the sine sweep. **Figure 11** shows the recorded-in-room MATT signal after post-processing by Acoustic Sciences Corp. to clarify the frequencies and amplitudes within the wave. More quantitative test results can be had by submitting a recording of the MATT playback in the user’s listening room and paying a small analysis fee to Acoustic Sciences Corp.

Just for comparison, **Figure 12** shows the clarity, definition, %Alcons, STI, and other acoustical parameters for the listening room. The aural impression from the MATT playback was “Ta-ta-ta,” indicating good articulation in this room, as supported by the very high C80 and STI scores. It is important to note, though, that C80 and STI are full-bandwidth calculations; whereas the MATT test targets tonal frequency effects, especially including the low-frequency (subwoofer) range in the room.

Among the test CDs on my shelf, there are several signals for testing acoustical and electro-acoustical environments. These include:

- Sine sweeps
- Pink noise: wideband and narrowband (to 1/18 octave)
- Anechoic music recordings
- Square waves

Although each of these has valid uses, none is as effective for identifying low-frequency defects in small listening rooms as is the MATT test. All of the others also require recording and/or post-analysis of the test signal reproduction to yield meaningful information. MATT stands alone in its ability to yield repeatable results through subjective evaluation.

**Author’s Note:** I would like to thank Art Noxon of Acoustic Sciences Corp. for his assistance in researching and preparing this article.

---

**References**


**Resources**

Acoustic Sciences Corp., www.acousticsciences.com