Reductions in Speech Modulation

The Speech Transmission Index is a measure of the reduction of speech modulation by noise and reverberation in the environment. The essence of this method of clarity evaluation can be made obvious by simply observing the time vs. amplitude response of recorded speech.

Figure 1 shows the time vs. amplitude response of anechoic (no echoes) speech. Note the lack of energy between the speech syllables. The dynamic range of the speech has not been reduced by the environment. A room will tend to “fill in the gaps” of the time vs. amplitude record, therefore reducing speech modulation.

A recent site survey provided an opportunity to document the effect of a reverberant environment on speech modulation. The survey involved a comparison between devices of differing directivities in order to determine which one would provide an adequate direct-to-reverberant energy ration at the required listener distance of 41 ft.

The playback system was adjusted to produce the same L_P at the listener distance from each device, which in turn allowed the L_R generated by each to be compared. The room has a 6 sec RT_{60} at 2 kHz, and is very diffuse. I normalized the recorded waveforms to the same peak amplitude. This allowed the crest factors (peak-to-rms ratio) of the waveforms to be compared. Higher directivities should result in higher crest factors, since a lower sound power level L_W is required to produce the desired L_P. The results are documented on the opposite page.

The experiment reinforced our understanding of the relationship between directivity and intelligibility in reverberant spaces. Since the listener distance D_2 was defined (41 ft.), directivity was the only means of modifying the direct field level L_D. If D_2 can be varied, then other possibilities exist (lower Q at a closer distance with an increase in N factor). This is the language of the Hopkins-Stryker equation, which sets forth a set of relationships that all system designers should understand.

Our seminar attendees have enjoyed hearing these wave files to reinforce their understanding of Q, distance, and direct-to-reverberant ratios. *pb*

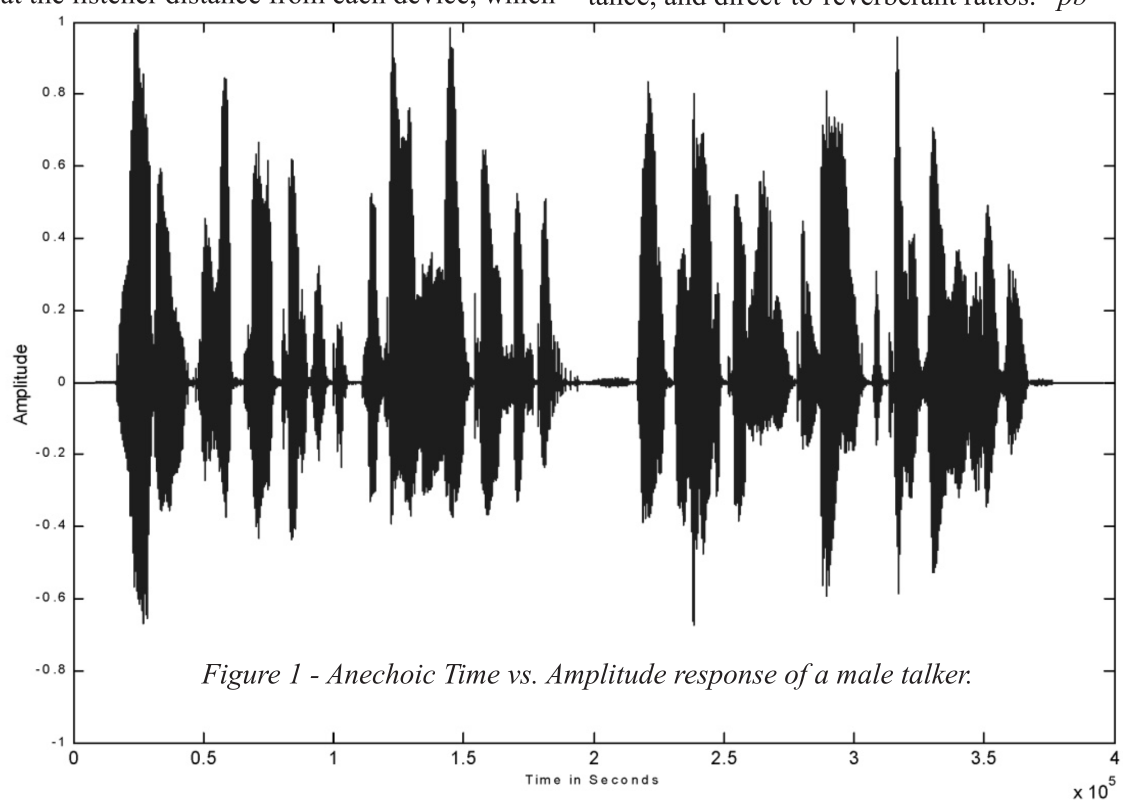
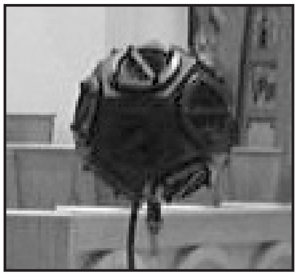
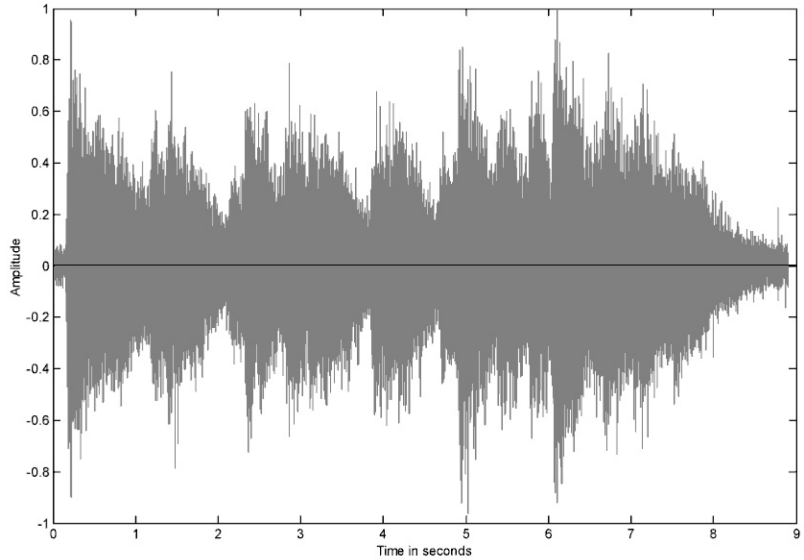


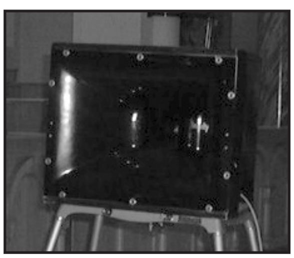
Figure 1 - Anechoic Time vs. Amplitude response of a male talker.

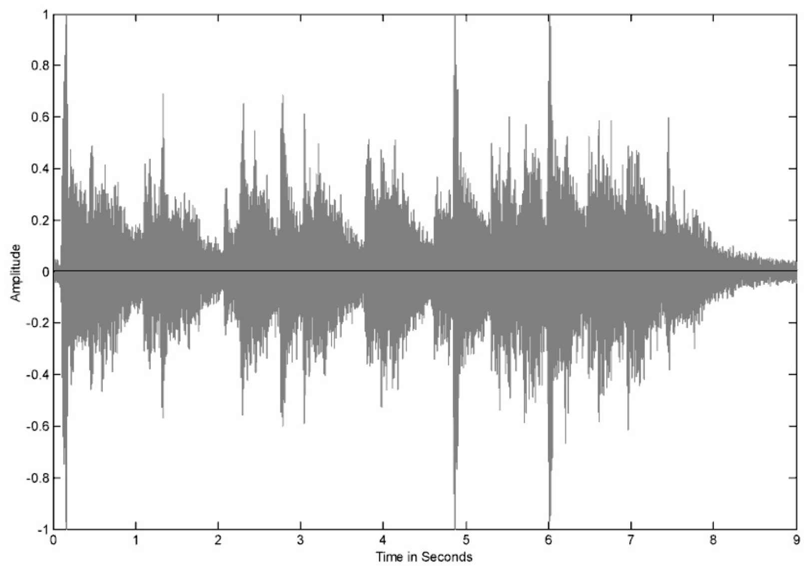
Q @ 2 kHz: 1
Crest Factor: 16.68 dB





Q @ 2 kHz: 13
Crest Factor: 19.84 dB





Q @ 2 kHz: 33
Crest Factor: 20.07 dB