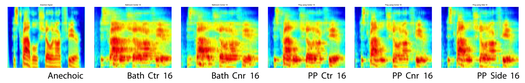
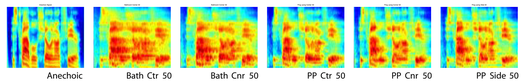
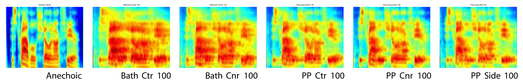
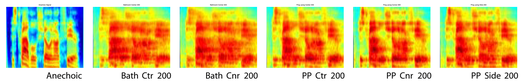
Effects of Reverberant Energy on Statistics of Speech

Madhu Shashanka[1], Barbara Shinn-Cunningham[1], and Martin Cooke[2]
[1]Boston University, [2]University of Sheffield
{shashanka,shinn}@cns.bu.edu, m.cooke@dcs.shef.ac.uk









Objective

Study how reverberation affects sparsity and statistics of speech.

Measurements: We measured impulse responses in five different room/position conditions and at four different distances from the source.

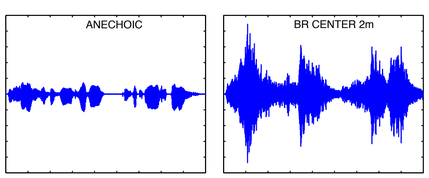
- Bathroom
  - Center
  - Corner
- PingPong Room
  - Center
  - Corner
  - Side

At every location/condition, responses were recorded at two mics (left and right) which were 0.4 m apart.

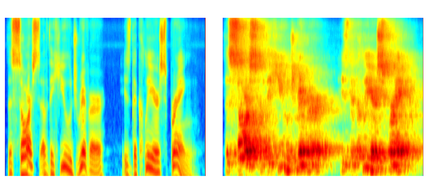
How does reverb distort a speech signal?

Reverberation smears energy across time. A reverb signal is less sparse in time and frequency than the anechoic signal. The waveforms and spectrograms of an anechoic signal and a reverb signal for a particular condition are shown below.

Waveforms



Spectrograms

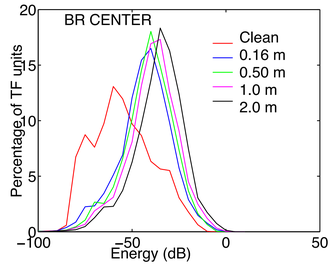


Approach

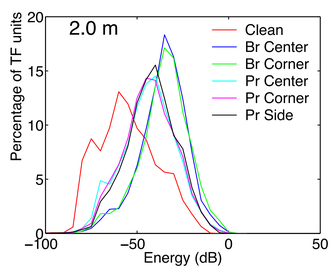
We compared energy differences between signals using a time-frequency representation (spectrogram). Reverberant signals were first time-aligned with the anechoic signal and levels were corrected so that RMS energy of the direct sound was equal to that of the anechoic signal.

Reverberation → More Energy

Below are histograms of the energy in time frequency (TF) units.



The relative amount of energy from reverberation grows with distance.
Histogram of TF energy for the bathroom center condition.



The hard-walled bathroom (BR) produces more reverberant energy than a classroom (the ping-pong room, PR). *Histogram of TF energy for different room conditions, for a 2-m distance.*

Speech and Silence

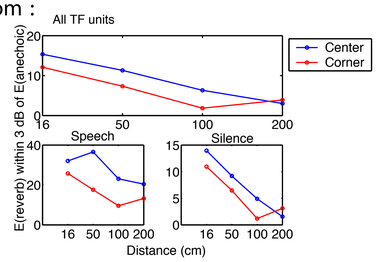
We consider another measure of the effects of reverberation – percentage of TF units for which anechoic energy was within ±3 dB of reverb energy.

We analyzed all TF units togethe, but also broke the TF units into two categories:

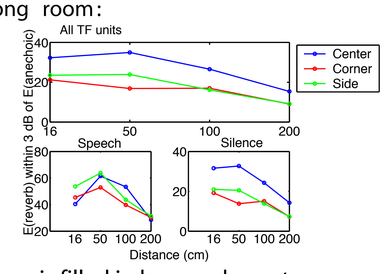
- Speech - if energy is within 20 dB of max energy.
- Silence - if energy is less than max energy by more than 20 dB.

Below, we plot these percentages separately for the bathroom and the pingpong room.

Bathroom :



Pingpong room:



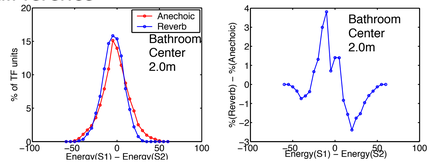
Silence is filled in by reverberant energy, especially in the hard-walled bathroom. As distance increases, the effects of reverberation increase, both for silence and for speech. However, filling in the silence destroys the normal modulations that convey speech meaning.
Plots show the percentage of TF units for which the energy in the reverberant condition is within 3 dB of what it was in the clean speech.

Simultaneous Speech

We now consider the effect of reverberation on simultaneous speech signals.

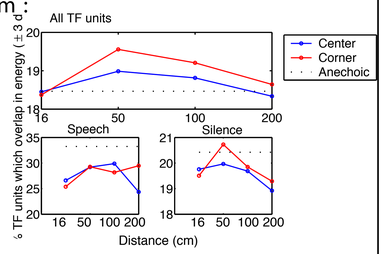
Below, we plot the histogram of the energy difference between S1 and S2 in the anechoic and reverberant cases.To visualize the difference more clearly, we also plot the difference between the two curves (right).

Energy(S1-S2) Histograms and their difference

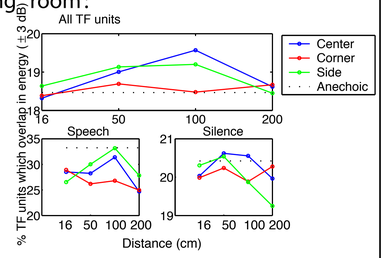


We can also see to what extent energies in the two signals overlapped. Below, we plot the percentage of TF units for which energy in speech sample 1 was within 3 dB of energy in the other sample.

Bathroom :

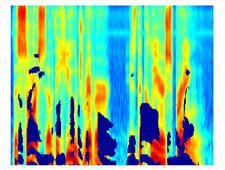


Pingpong room:



Energy of S2 in the *gaps* of S1

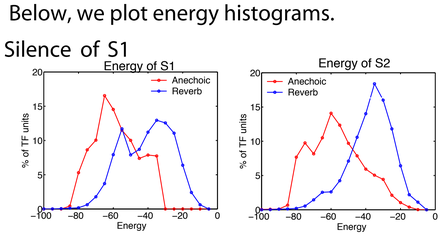
We consider how much energy of speech sample 2 is present in the silence regions of speech sample 1. The plot below shows the energy of speech sample 2 and dark blue regions represent TF units where sample 1 has high energy.



Sample 2 in the "silence" of Sample 1

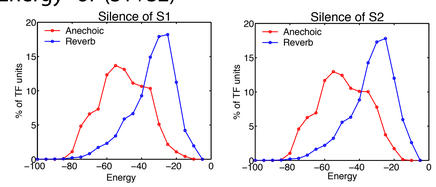
Below, we plot energy histograms.

Silence of S1



Histograms of energy of signal 1 (left panel) and signal 2 (right panel) in the silence regions of signal 1.

Energy of (S1+S2)



Histograms of energy of both signals in silence regions of signal 1 (left panel) and signal 2 (right panel).

Acknowledgements

1. Tim Streeter for helping with measurements of impulse responses.
2. Portions of this work were supported by grants from
(a) Air Force Office of Scientific Research
(b) National Science Foundation.