IEEE Workshop on Speech Coding for Telecommunications
Digital Voice for the Nineties
Whistler, British Columbia, Canada - September 1991

Generating Voice Intelligibility and Quality Test Material in Simulated Acoustic Noise Environments

John D. Tardelli — ARCON Corp.
Charles Teacher - KETRON, Inc.
Paul D. Gatewood - ARCON Corp.

Abstract

In 1989 an effort was initiated to provide the capability for the USAF's RL/EEV office to generate voice intelligibility and quality test material in acoustic backgrounds representing platforms of interest to DoD. Specific material was to include master audio tapes for the implementation of the Diagnostic Rhyme Test (DRT) and the Diagnostic Acceptability Measure (DAM). RL/EEV intended to record noise environments in the field, reproduce them in an acoustic isolation chamber, and record various talkers in that chamber speaking over a microphone resident to the simulated environment.

Consequently, two major objectives were required:

1) The establishment of techniques to record stereo ambient noise recordings in the field.
2) The establishment of techniques to reproduce these noise environments in an acoustic isolation chamber in order to generate materials to evaluate narrowband voice communication systems.

The environments used to develop these techniques were USAF road vehicles, specifically, a 1988 Plymouth Reliant Staff Car, a 1988 Chevrolet Suburban Truck, a 1987 Dodge Ram 50 Sport 1/2 Ton Pickup, and a 1989 Lincoln Towne Car. Conditions included: highway travel at 55mph with windows down and windows up with air conditioning on and off, residential roads at 25 mph with windows up and air conditioning on, the Suburban truck off road at 25 mph with the windows up and down.

For field recording a microphone mounting system was designed to allow calibrated, laboratory standard microphones to be placed on each shoulder of a subject at a position of interest within the environment. The system also allowed for a third microphone to be mounted for the collection of sound pressure level data. Windscreens were installed when necessary. The stereo pair was routed to a two—channel, portable digital audio processor/recorder. Using specific microphone calibrators, a reference tone of known acoustic sound pressure level was recorded after initial level adjustments and before actual data collection. SPL and level values with "A" and "C" weightings were logged periodically throughout the field recording.

In addition to the noise field recordings, a number of in situ DRTs were recorded in the field as controls for simulations performed in the sound chamber. Microphones used by the DRT talkers, both in the field and in sound room simulations, were a Motorola cellular handset microphone and a noise canceling Roanwell "Confidencer" handset.

During simulations, two channel recordings were generated, one channel with the cellular handset or the "Confidencer, the other using an Altec 659A, a high quality, dynamic microphone.

For sound chamber simulations, equipments to reproduce the acoustic noise platforms were procured with an emphasis on accurate SPL and frequency response within the bandwidth of secure narrowband speech (100Hz — 4kHz)

The digital field recordings were played through a preamplifier and into a stereo power amplifier, whose output was routed via a patch panel to stereo 15" speakers.

A precision measuring amplifier and calibrated microphone were used to calibrate room sound levels and to compare tape vs. room levels. This configuration produced reasonable results within the specified bandwidth up to an SPL of 120dB using a white noise source. A problem concerning accurate low frequency (less than 100 Hz.) reproduction of the simulated noise fields and its effect on test materials recorded within these fields was discovered. Spectral analysis and comparisons of field vs simulated noise environment DRTs were used to study this discrepancy. The solution of this problem led to a modification of the calibration procedures and the addition of an 18" subwoofer speaker to the sound reproduction system. This speaker received a filtered mono mix of the stereo field recording, processed by an additional power amplifier. Although this problem may be specific to the noise environments studied, it does indicate that extreme care must be taken when performing intelligibility and/or quality testing of digital voice communication systems in vehicular environments.

DRTs generated by the same talkers in the field and in the simulated noise environment were evaluated unprocessed, with LPC-1O processing at 2.4Kbps and APC processing at 9.6Kbps. Final results indicate a reasonable agreement between tests performed in the field and the sound chamber. Scores were within 2 DRT points of each other. These differences are within the range of talker and listener variation.

It has been the intention of RL/EEV to include DRT/DAM masters generated with the techniques discussed in this paper into the DOD Digital Voice Processor Consortium's audio library of voice intelligibility (the DRT) and quality (the DAM) master input tapes. The library includes talkers in a quiet environment and a host of simulated acoustic noise platforms using resident microphones. Originally housed in analog format, these tapes have been digitally remastered, presentation levels equalized, and Speech-to-Noise ratios measured and documented. As the custodian of this library, the USAF's RL/EEV office has generated a sub library of material that has been digitally dubbed onto R-DAT cassettes. This sub library is available to US Government agencies and their sponsored contractors. Procedures for receiving copies of this test library will be presented.