2004 Command and Control Research and Technology Symposium
The Power of Information Age Concepts and Technologies

TOPIC: C2 Human Factors Engineering

The Requirements and Applications of Speech Recognition Technology for Voice Activated Command and Control in the Tactical Military Environment

By

Mr. Lockwood Reed

US Army CERDEC

Fort Monmouth, NJ 07703

AMSRD-CER-C2-SS

732-427-2559/FAX 732-532-0134

Lockwood.Reed@us.army.mil

Initial research into speech recognition technology

- Interactive Speech Technical Advisory Committee
  - All Uniform Services
  - Numerous Govt. Agencies (Agriculture, Post Office, etc.)
- Evaluate the effectiveness of COTS speech recognition technology for Avionic Command and Control.
  - NAP-of-the-earth flight
  - Hands & Eyes Busy
  - High Noise (98 dBA – 123dBA)
- Unassisted COTS technology failed to perform reliably at sound levels as low as 80dBA.
  - Various Noise Filtering approaches investigated
  - Additive Noise approach most successful

RDECOM

The evolution of our speech recognition technology

- Fort Hood Database Collection
  - Apache (AH-1), Iroquois (UH-1), Kiowa (OH-58) & Mohawk (OV-1)
  - Various Maneuvers – Straight Flight, Banking Turns, Hover (In & Out of ground effect), etc.
  - Testing Communication Systems
  - Recording Ambient and Comm. Environments
- Noise Processing
  - Analysis of noise characteristic impact on recognition performance
  - Multi-Level approach to Noise
    - Noise Cancellation Microphone
    - Active Noise Processing
  - Current solutions address Quasi-stationary noise
  - New Solutions to Impulsive Noise in progress.

RDECOM

The evolution of our speech recognition technology (Cont.)

- Gain Management
  - Sensitivity to Signal-to-Noise Ratio
  - Noise Cancellation Microphone Impacts Gain Management
  - Deterministic Automatic Gain Control
- Recognizer Activation
  - Conventional Press-to-Talk (PTT)
    - Comm. System Activated Side-Tone
  - Non-Conventional Press-to-OFF (PTO)
    - Unintended Vocalization
    - Word Switch
    - VR System Activated Side-Tone
- Audio Feedback
    - Command Completion Response
    - Command Interrupt Response on Error

RDECOM

The evolution of our speech recognition technology (Cont.)

- Beyond Threshold Functionality
  - Whispered/Shouted Speech
    - Requires improved front-end dynamics
    - Currently achieved through re-enrollment
    - New technology will eliminate re-enrolment
  - Multiple Speaker Confusion Avoidance
    - Command Post Environments (Close Quarters)
    - Adjacent Voice Spillover (AVS)
      - Exacerbated by Far-Field microphones
    - New technology will reduce affects of AVS

SRS Comparative Testing

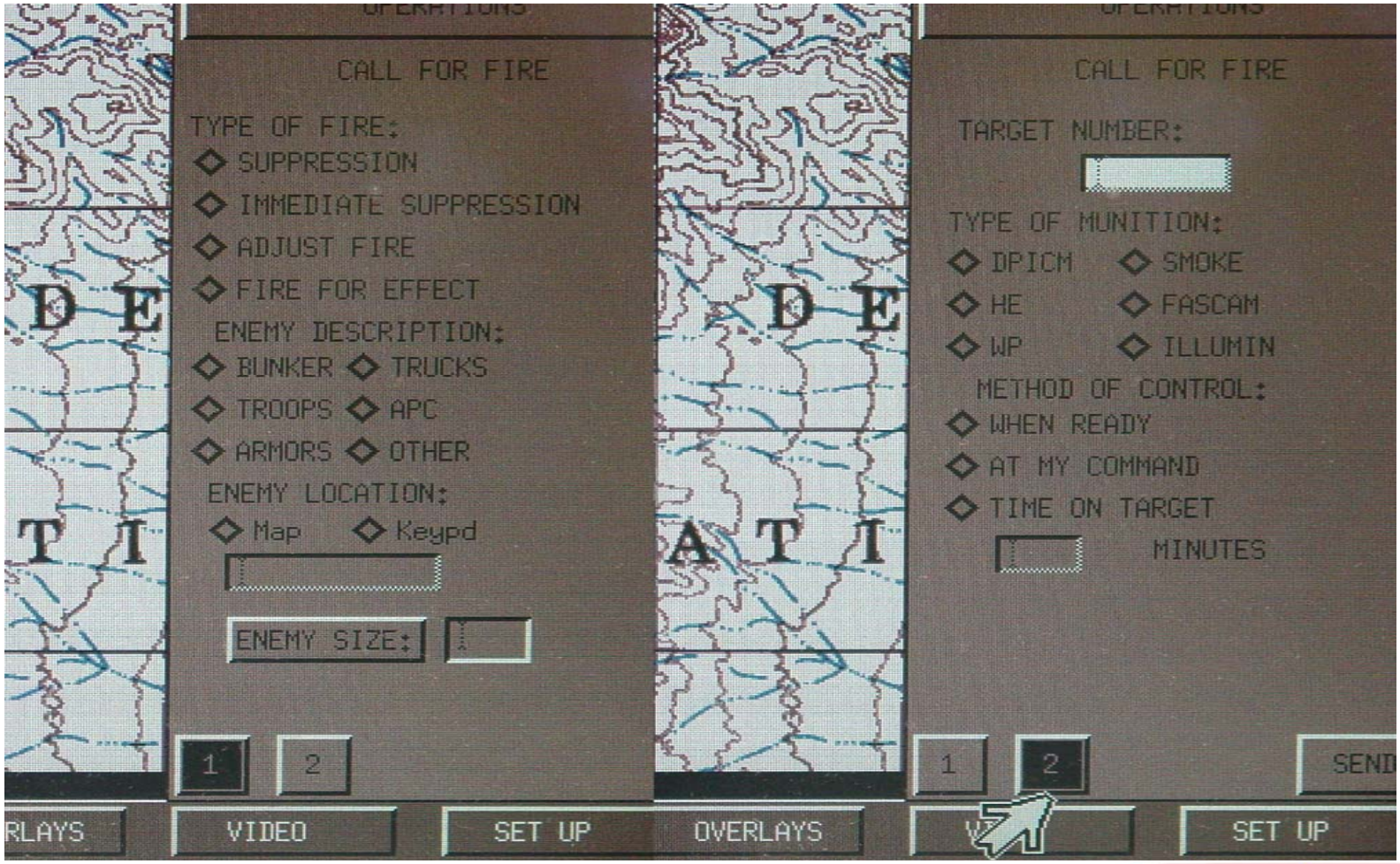
- Test Application: Brigade & Below, Command & Control (B2C2)
- Input Modality
  - Natural Speech Recognition
  - Isolated-Word Speech Recognition
  - Manual Entry
- Task – Call For Fire message
- Metrics
  - Task Completion Time
  - Accuracy
- 18 Test Subjects

RDECOM

Test Configuration

RDECOM

Call-For-Fire Task



Composite Image, CFF Pages 1 & 2

Comparative Test Results

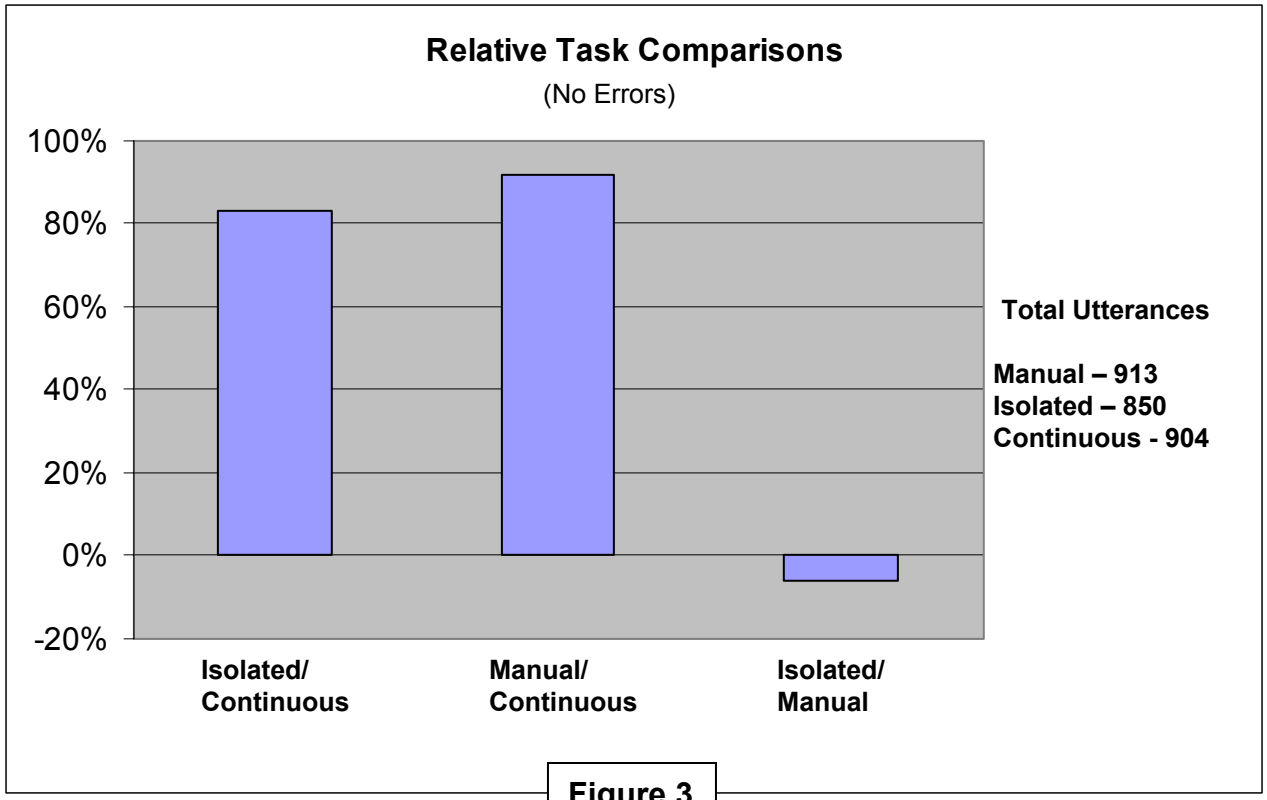


Figure 3

Comparative Test Results

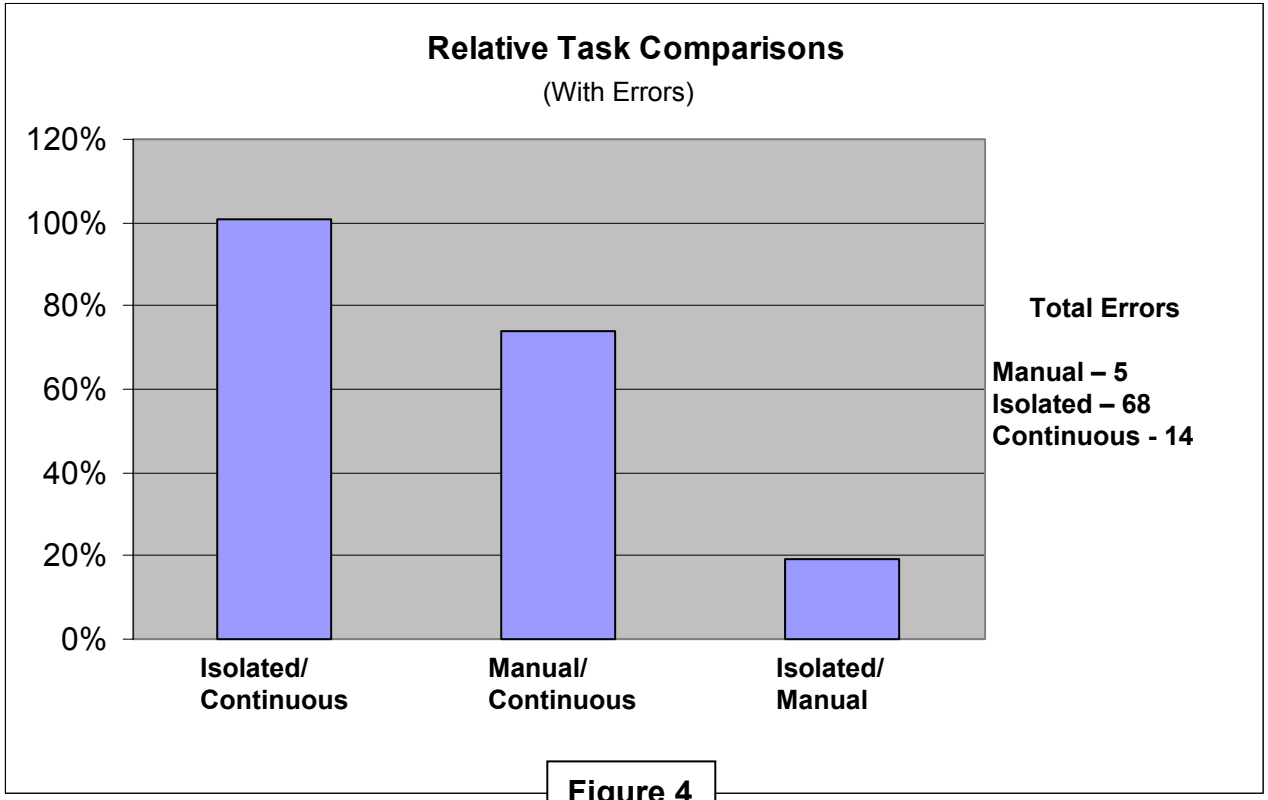


Figure 4

Subjective Questions

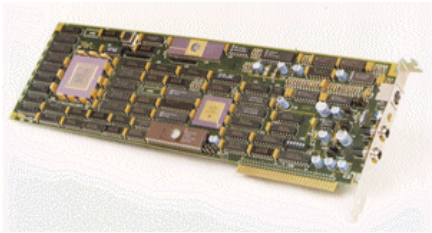
- How easy was it to use (continuous recognition, isolated recognition, and manual entry)?
  - Scale: 1 to 5, where 1 is easiest and 5 is hardest
  - Results: Continuous (1.3), Isolated (2.5), and Manual (2.3)
- How comfortable did you feel with (continuous recognition, isolated recognition, and manual entry)?
  - Scale: 1 to 5, where 5 is the most comfortable and 1 is the least
  - Results: Continuous (3.7), Isolated (2.5), and Manual (3.1)
- Speech Recognition was preferred over Manual Entry

Tactical Voice Activation System

Command Voice!™ All-Software Tactical Voice Recognizer

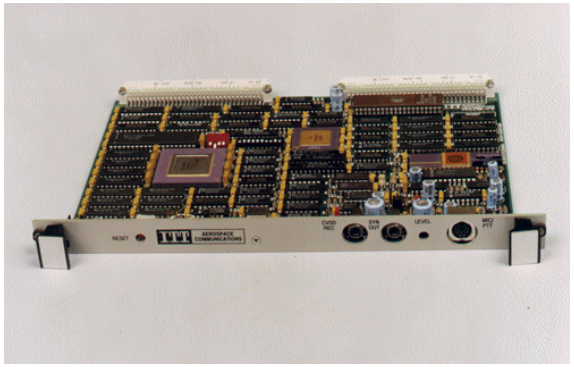
Command Voice!™ PCMCIA Tactical Voice Recognizer

Command Voice!™ Voice Application Toolkit



Command Voice!™ ISA Bus Tactical Voice Recognizer

Command Voice!™ SBus Tactical Voice Recognizer



Command Voice!™ VME Tactical Voice Recognizer

TVAS In ACTION

RDECOM

In to the *Future*

OBJECTIVE: Intelligent Soldier/Machine interaction, providing increased accuracy, reduced task time, yielding greater Survivability and higher Lethality.

Advanced Cognitive Interactive Speech Technology (ACIST)

RDECOM