

TransTac

Dr. Sherri Condon

703-983-5522 • scondon@mitre.org

DARPA IPTO/CECOM





Problem

- **Address the widespread operational need for translation support: the most critical impact is on soldiers**
- **Develop technologies and systems that enable spontaneous two-way speech communications: must be robust under real-world tactical use conditions**
- **Reduce time to solution for new languages and dialects: phase 2 includes a “surprise” language to be developed in 90 days**



Background

■ Technical Challenges

- Speech recognition and machine translation
- Fast development with few resources

■ Current Status

- Limited one-way handhelds in the field
- Two-way prototypes for 5 languages in FY 04–06
- 5 systems for Iraqi Arabic, 1 for Persian in FY 07

■ Previous MITRE work

- Domain definition and data collection procedures
- Evaluation methods and metrics for speech translation



Objective

- **Support development and evaluation of two-way speech-to-speech translation devices**
 - To address the needs of military users
 - To provide useful feedback for developers
- **Develop measures to characterize training data for effective data collection methods**
 - Measure variation in scenario-based data
 - Estimate domain coverage
 - Compare training data to test data



Activities

- **Support NIST evaluations of 5 Iraqi Arabic systems, January and July 07**
 - Guide planning, scenario preparation
 - Perform component testing
- **Evaluate Farsi system, February 07**
 - Component testing
 - Live scenarios with Farsi speakers
- **Data characteristics report, March 07**
- **Evaluate Iraqi Arabic headset system, April 07**



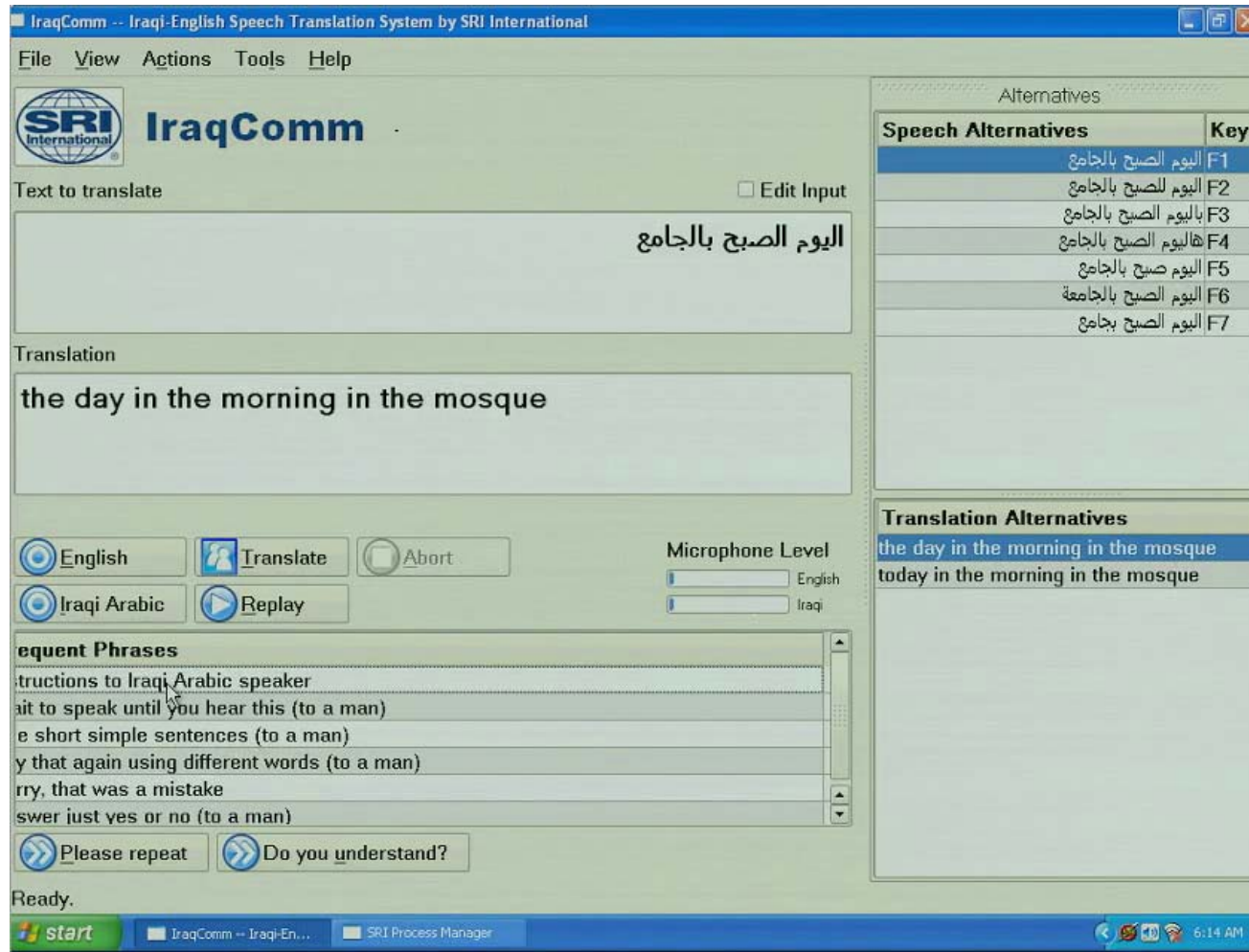
Highlight

- **Evaluation of English-Farsi (Persian) system**
 - Medical personnel interacted with Farsi speakers
 - Recorded 20 interactions in 10 domains
 - System also processed 20 interactions from held out training data for component measures
- **Benchmark measures for rapid development from minimal data**
 - 8 domains developed in 2 months

Paramedic's comment: "I believe in the future these will be required equipment"



Highlight



SRI Speech-to-Speech Laptop Translation System



Impacts

- **Ability to converse in multiple languages; enhance situational awareness and safety of the warfighter in different environments**
- **Translation accuracy, robustness, language and domain portability, collection logistics, and training**
- **Innovative methods and measures to evaluate speech translation devices**
- **Put the most effective technology into the warfighter's hands as quickly as possible**

Future Plans

- Develop task-based evaluation measures for a variety of system design features
- Identify effective data collection methods for rapid development of new languages



IWT-BBN Headset Translation System