U.S. Army
Machine Foreign Language
Translation System
(MFLTS)

October 2012

Michael J. Beaulieu
Product Director, 703-704-0304
Outline

- United States Army Machine Foreign Language Translation System (MFLTS) – Who We Are
- A view of our Capability
- Types of Development
Mission: Develop, field, and sustain a basic automated foreign speech and text translation capability into Army tactical systems in order to augment and complement limited human linguistic resources across all echelons in all environments.
**Description**

MFLTS is a software-only program that will develop, acquire, field, and sustain a basic automated foreign speech and text translation capability into U.S. Army systems. The language development design allows for technical insertion of software language modules performing translation as required by the U.S. Army.

**Capabilities**

- **Initial Capability:**

<table>
<thead>
<tr>
<th>Speech</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>2WS2S = Two-Way Speech-to-Speech</td>
<td>2WT2T = Two-Way Text-to-Text</td>
</tr>
<tr>
<td>3 languages (Pashto, Iraqi Arabic, &amp; MSA)</td>
<td>2 languages (Dari &amp; MSA)</td>
</tr>
<tr>
<td>portable &amp; mobile configurations</td>
<td>mobile &amp; web-enabled configurations</td>
</tr>
<tr>
<td>2 domains (checkpoint operations &amp; base security)</td>
<td>2 domains (checkpoint operations &amp; base security)</td>
</tr>
</tbody>
</table>

- **Follow-on Capability:** modules for additional languages, dialects (for already developed languages), OCR for handwriting, and technological upgrades.

**Characteristics**

- Portable (wearable or handheld), mobile (laptop), and web-enabled configurations
- Interoperable with commercial-off-the-shelf (COTS), or government-off-the-shelf (GOTS) automation equipment
- Plug-n-play language modules - a unit that translates between specific languages
- Rapid development and deployment of new language modules to meet warfighter operational needs
- Capable of deploying on multiple operating systems (e.g. Windows, Linux, Solaris)

**Vision:** To be a high performance and learning organization that delivers world class products to the Warfighter and takes care of people.

**Benefits**

- Greatly increases points of presence for foreign language translation capabilities on the battlefield
- Enables communications with foreign speakers when human linguists are not available, enhancing mission success with:
  - Host Nations
  - Coalition partners
  - Indigenous peoples
- Enables rapid translations and triage of documents and media focusing analysts on the most important documents/media for exploitation

**Not a “device” ... but rather an innovative Software Intensive Program employing modular, open system architecture**

**President Obama “... in the 21st century, military strength will be measured not only by weapons our troops carry, but by the languages they speak and the cultures that they understand”**

**2WS2S** = Two-Way Speech-to-Speech

**2WT2T** = Two-Way Text-to-Text

**COTS** = Commercial Off-the-Shelf

**GOTS** = Government Off-the-Shelf

**MSA** = Modern Standard Arabic
**SPEECH-TO-SPEECH**

Speech Recognition (ASR) → Transcription Software → Translation Engine → Text File → Speech Synthesis

Audio File → Digitized Speech → Text File → Audio File

**Text-to-Text**

Text Recognition → Translation Engine → Text File

Image → Text File → Text File
Speech-to-Speech (S2S)

You: Hello how are you doing

 Iraqi

You: مرحبًا شلونك

 Iraqi

Sleep
In Jeddah began a meeting of foreign ministers of the Organization of Islamic cooperation, such as the launching of Mecca summit expected tomorrow of the leaders of the Islamic world, which is expected to focus on the situation in Syria, and on "problems that impede the process of Islamic work", according to Abdul Aziz bin Abdullah, Deputy Saudi foreign minister.

In his speech, he said the secretary-general of the organization Ekmeleddin Ihsanoglu, the organization of cooperation, "stresses the need to preserve the unity of Syria and stop violence against civilians immediately."

Cakmakoglu added that the seriousness of what is happening on the level of the Islamic world, "requires the consolidation of efforts."

The Secretary-General renewed condemnation of the Organization to Rohingya Muslim minority in Myanmar of robbery and the violation of rights and freedoms sacrilegious, referring in this regard to deprive them of their fundamental rights in nationality, where it considers the Myanmar authorities "refugees" from Bangladesh.

The number of Ihsanoglu a number of the steps taken by
Open Issues

Metrics

Automated Testing
- Determines State of the Art
  - MT: BLEU
  - ASR: WER
  - OCR: WER

SUE
- Limited Operational Data

SUT
- Soldier Task Completion Using Device
  - (small scale, man in the loop)

Test Activities
- DT

Manual Scoring Protocol (Developed)

1. ILR* $\approx$ ILR + (Definition of Applicability of Machine Translation)
   - Soldier knows how to do job
   - Soldier knows how to use machine
3. Automated Metric Which Correlates to ILR
   … Poor Correlation

ASR = Automatic Speech Recognition
BLEU = Bilingual Evaluation Understudy
DT = Developmental Testing
ILR = Intergovernmental Language Roundtable
MT = Machine Translation
OCR = Optical Character Recognition
OT = Operational Testing
SUE = System Under Evaluation
SUT = System Under Test
TRR = Test Readiness Review
WER = Word Error Rate
Open Issues (continue)

- **Data Collection**
  - Expensive
  - Time Consuming
  - Need Industry / Academia / S&T / Community Collaboration

- **Domain Development**
  - Expensive
  - Time Consuming
  - Need Industry / Academia / S&T / Community Collaboration
Conclusions

- **Initial Capability**
  - On schedule, within budget and performance parameters
  - Need Industry / Academia / S&T / Community Collaboration with Open Issue:
    - Metrics

- **Follow-on Capability**
  - Need Industry / Academia / S&T / Community Collaboration with Open Issues:
    - Metrics
    - Data Collection
    - Domain Development