Mobile Speech Translation
By Farzad Ehsani, CEO
Technology Overview

- 4th generation
- 7 military deployments
- 2 successful hospital pilots
- Thin-client based
- Smartphone prototype
System
Capabilities & Limitations

• **Fast and easy to use**
  – Hands/eye Free

• **Accurate and consistent**
  – Works in high-noise environments

• **Interprets for specific jobs & settings**
  – Not the “Star Trek Universal Translator”
  – But provides significant language capabilities

• **Handles 80-90% of all common interactions for a given setting**
System Details

- **Audio to Speech Recognition (ASR)**
  - Nuance 8.5 Engine
  - Multiple Language models
  - Normally the small ASR is specific to the sub-domain/question and the large ASR covers the entire domain

- **Translation**
  - Statistical machine translation engine
  - Similar to Giza with lexical constrains and phrase alignment
  - Ability to handle multiple translation and language models

- **Template Translation**
  - Using templates to capture and translate common phrases
    - E.g., “How are you?”, “Where does it hurt?”
System Evaluation

• **Evaluation Performed at Kaiser San Francisco**
  - 3 nurses and 11 patients
  - 95 nurse-patient encounters
  - 500 conversational segments

• **User Training**
  - Each nurse was trained for 1-1.5 hours prior to system use
  - Patients received a 40 sec sound-byte on the system functionality before using it

• **System Evaluation**
  - Every utterance judged as Good, Fair, Poor, Mistranslated, or Not Translated by 2 people
    • Conflicts resolved by a 3rd person
  - Examples given in the next two slides
<table>
<thead>
<tr>
<th>What Nurse Said</th>
<th>S-MINDSTM Translation</th>
<th>Human Translation</th>
<th>S-MINDS Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>I will give you an I.V.</td>
<td>Voy a colocarle un cateter para liquidos intravenosos.</td>
<td>Voy a colocarle un cateter de liquidos intravenosos.</td>
<td>Good</td>
</tr>
<tr>
<td>Let me check if I can give you medication for that.</td>
<td>Dejeme chequear si puedo darle algun medicamento.</td>
<td>Permitame reviso si puedo darle algun medicamento para eso.</td>
<td>Fair</td>
</tr>
<tr>
<td>I will check your…</td>
<td>Yo voy a revisarle los vendajes</td>
<td>Voy a revisar su …</td>
<td>Poor</td>
</tr>
<tr>
<td>Did someone take your vitals?</td>
<td>¿Le tomare sus signos vitales?</td>
<td>¿Alguien tomo sus signos vitales?</td>
<td>Mis-translated</td>
</tr>
<tr>
<td>Your heart rate is normal.</td>
<td>---</td>
<td>Su frecuencia cardiaca es normal.</td>
<td>Not Translated</td>
</tr>
<tr>
<td>What Patient Said</td>
<td>S-MINDS Translation</td>
<td>Human Translation</td>
<td>S-MINDS Accuracy</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------------------------------------</td>
<td>------------------------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>No, no tengo dificultad en respirar.</td>
<td>I don't have difficulty breathing.</td>
<td>No, I don't have difficulty breathing.</td>
<td>Good</td>
</tr>
<tr>
<td>En la parte baja del estómago.</td>
<td>The lower part of my stomach.</td>
<td>In the lower part of my stomach.</td>
<td>Fair</td>
</tr>
<tr>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>Mistranslated</td>
</tr>
<tr>
<td>Los huesos.</td>
<td>---</td>
<td>My bones</td>
<td>Not Translated</td>
</tr>
</tbody>
</table>
Overall Translation Accuracy

92% Good

1% Fair
1% Poor
4% Mistranslated
2% Not Translated
Nurse Translation Accuracy

- Good: 375 (93%)
- Fair: 3 (1%)
- Poor: 3 (1%)
- Mistranslated: 21 (5%)
- Not Translated: 2 (0.5%)
Patient Translation Accuracy

- Good: 88% (84)
- Fair: 2%
- Poor: 0%
- Mistranslated: 0%
- Not Translated: 10% (10)

Fluential
Eliminating Language Barriers
Nurse/Patient Surveys (based on 5-point Likert scale)

- System was Easy to Use
- English/Spanish Voice Easy to Understand
- I understood the patient/nurse better
- I feel I am giving/receiving better care
- I would like to use system in the future

Nurse
Patient
Problems with Cloud-based Recognition

• **Startup time**
  – Up to 2 minute for bringing up a new processor
  – Need to have the processor before the user request

• **Real-time access**
  – Use of codecs decreases recognition accuracy
  – Bandwidth depends on network load
  – Even with good bandwidth the speed is variable
Future Effort

• More extensive technology evaluation
  – Low tolerance for errors in healthcare setting
  – Need to reduce user training time

• Improve system performance with additional data

• Better system response
  – To handle network load fluctuations