**Introduction**

As speech-to-speech translation systems move from the laboratory into **field deployment**, we quickly see that mismatch in training data with field use can degrade the performance of the system. Retraining based on field usage is a common technique used in speech systems to improve performance. In the case of speech-to-speech translation we would particularly like to be able to adapt the system based on its usage automatically without having to ship data back to the laboratory for retraining. **This paper investigates the scenario of a two-day event. We wish to improve the system for the second day based on the data collected on the first day.**

**Data Scenario**

ASR and MT was trained on data from English-Iraqi force protection and civil affairs dialogs.

Most users change their language when using an automatic S2S translation system.

Users switch to a clearer pronunciation and use shorter and simpler sentences with less disfluency.

Collect data in two days, with around **2 hours of actual speech** per day. This data was transcribed and translated, resulting in **864 and 824 utterance pairs on day 1 and 2**, respectively.

**ASR LM Adaptation**

Unsupervised: Use the ASR hypotheses from day 1 to build a LM. Interpolated with the original 3gram LM to produce an adapted LM for day 2.

Supervised: Having transcribers provide accurate transcription references for day 1 data, and see how it may improve the performance on day 2.

**SMT Adaptation**

Online/Offline data

**Joint Adaptation**

<table>
<thead>
<tr>
<th>ASR</th>
<th>SMT</th>
<th>Day 1</th>
<th>Day 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>No adaptation</td>
<td>No adaptation</td>
<td>29.39</td>
<td>27.41</td>
</tr>
<tr>
<td>Unsupervised ASR adaptation with 1gramLM 500-Best MT hypo</td>
<td>1gramLM 500-Best MT hypo</td>
<td>32.07</td>
<td>28.65</td>
</tr>
<tr>
<td>Supervised ASR adaptation with 1gramLM MT transcriber</td>
<td>1gramLM MT Ref</td>
<td>31.76</td>
<td>28.83</td>
</tr>
<tr>
<td>Supervised ASR adaptation with 1gramLM transcriber</td>
<td>1gramLM MT Ref</td>
<td>32.48</td>
<td>28.59</td>
</tr>
</tbody>
</table>

**Selective Adaptation**

1. Take the translation hypotheses on day 1 of the baseline SMT and compare them with translation references, then select sentences which have BLEU scores higher than a threshold.
2. The subset of day 1 sentences is used to perform supervised adaptation.

**Conclusions**

Improvement is possible using collected data for adaptation.

The best results however still require producing translation references, notably ASR transcriptions do not seem to help, but may still be required in the process of generating translation references.