

Incremental Adaptation of Speech-to-Speech Translation

Nguyen Bach, Roger Hsiao, Matthias Eck, Paisarn Charoenpornasawat, Stephan Vogel, Tanja Schultz, Ian Lane, Alex Waibel and Alan W. Black
 {nbach, wrhsiao, matteck, paisarn, stephan.vogel, tanja, ianlane, ahw, awb}@cs.cmu.edu

Introduction

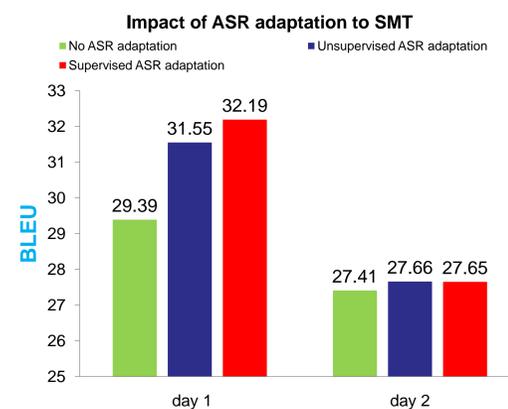
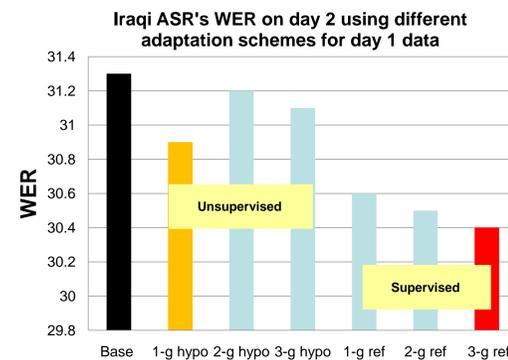
As speech-to-speech translation systems move from the laboratory into **field deployment**, we quickly see that mismatch in training data with field use can degrade the performance of the system. Retraining based on field usage is a common technique used in speech systems to improve performance. In the 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.

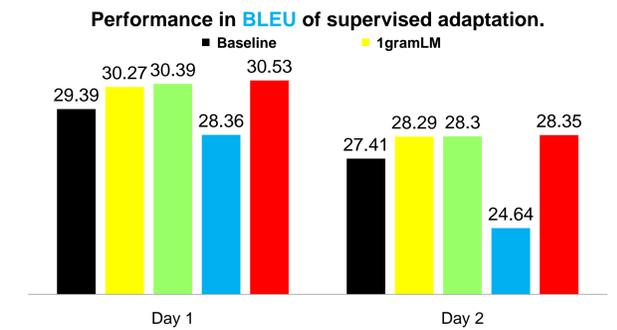
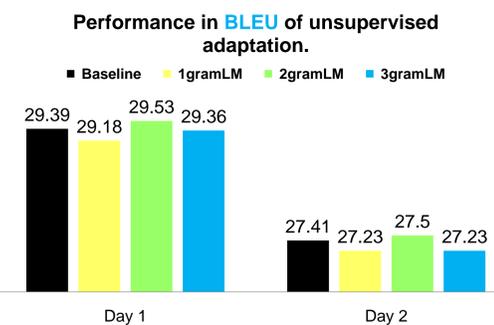
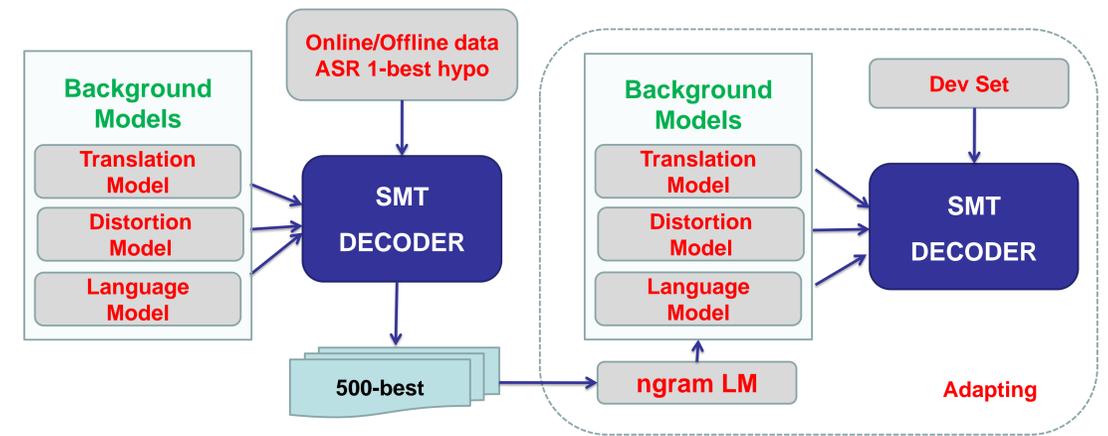
ASR LM Adaptation



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



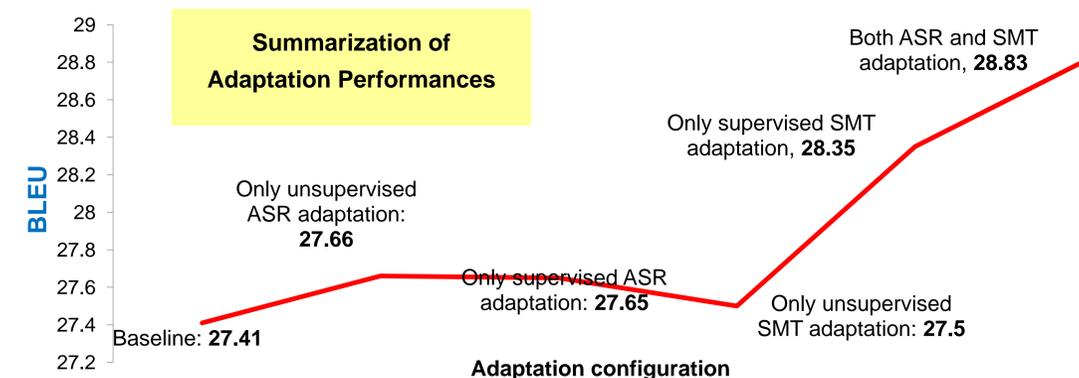
Joint Adaptation

ASR	SMT	Day 1	Day 2
No adaptation	No adaptation	29.39	27.41
Unsupervised ASR adaptation with 1gramLM ASR hypo	1gramLM 500-Best MT hypo	32.07	28.65
	1gramLM MT Ref	31.76	28.83
Supervised ASR adaptation with 1gramLM transcription	1gramLM 500-Best MT Hypo	32.48	28.59
	1gramLM MT Ref	32.68	28.6

Selective Adaptation

- 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.
- The subset of day 1 sentences is used to perform supervised adaptation.

No. Sents	Day 1	Day 2
Baseline	29.39	27.41
≥ 0	864	28.29
≥ 10	797	28.27
≥ 20	747	28.24
≥ 30	585	27.71
≥ 40	416	29.72
≥ 50	296	30.06
Correct	98	29.18



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