Word Sense Disambiguation for Statistical Machine Translation

- Most SMT systems do not explicitly use WSD
  - static translation probabilities, not sensitive to context

- But using WSD for SMT first gave confusing results
  - WSD for SMT hurts BLEU score!? [Carpuat & Wu ACL-2005]
  - But WSD should help SMT... [Carpuat & Wu IJCNLP-05]

- Generalizing WSD to Phrase Sense Disambiguation for SMT [Carpuat & Wu, 2007]
  - PSD is fully phrasal just like conventional SMT lexicons
  - PSD predictions are fully integrated in SMT decoding
  - PSD models are trained on the same parallel data as SMT lexicons

  PSD improves translation quality consistently on 8 metrics and 4 tasks

Marine Carpuat  Columbia University Center for Computational Learning Systems