Emily Bender
Natural Language Processing

1. Applications of NLP
   - grammar and spell checking
   - computer assistend language learning
   - assistive and augmentative communication
   - machine translation
   - information retrieval
   - information extraction
   - HCI

2. Approaches
   - knowledge engineering (rules)
   - machine learning
   - hybrid

3. Subtasks
   - identify which language
   - tag parts of speech
   - disambiguate word sense
   - recognize named entities
   - detect phrases
   - segment documents to sentence and then words
   - parsing
   - generation
   - reference resolution

Emily Bender (con’d)

4. Evaluation requires
   - test set with gold standard answers
   - metrics of comparison
   - baseline to compare against
   - there are many ways to parse a single sentence
   - just counting n-grams may be enough

5. Head-Driven Phrase Structure Grammar (HPSG)
   - declarative, order-independent, constraint-based formalism
   - collection of feature-structure descriptions
   - organized into a type hierarchy
   - rules contain both syntax and semantics
   - can be used by a parser or a generator

Katrin Kirchhof
Statistical Speech and Language Processing

1. Speech Applications
   - dictation
   - transcription of voicemail, phone conversations, TV shows
   - automated dialog systems
   - call centers
   - hands-free control
   - household appliances
   - assistive devices
   - search of audio archives

2. NLP Applications
   - document sorting
   - question answering
   - machine translation
   - document summarization

3. Subsystems
   - language modeling
   - parsing
   - tagging
   - word sense disambiguation
   - co-reference resolution
   - machine translation

Katrin Kirchhof (con’d)

4. Methodology
   - Early systems used rules.
   - Current systems use statistical pattern recognition.
   - Noisy channel model uses Bayes' rule.
   - Acoustic model uses a hidden Markov Model, a stochastic FSA.
   - The probabilities for the acoustic model are learned via EM.
   - The language model also needs to learn probabilities P(x|x1 x2 .... xn)
   - Machine translation systems use phrase-based models, mappings between phrases.

Fei Xia
Machine Translation

1. Applications
   - rough translation of web dta
   - computer-aided human translation
   - limited domain translation
   - cross lingual information retrieval

2. Evaluation requires automatic measures, since no gold standard

3. Challenges
   - choosing the correct root form
   - getting the correct inflected form
   - inserting spontaneous words
   - putting words in the correct order
   - one language may have a concept the other does not
   - resolving ambiguity

4. Resources
   - bilingual dictionary
   - grammar books
   - parallel comparable data
   - thesaurs
   - NLP tools

Fei Xia (con’d)

5. Major Approaches
   - transfer-based: parse the source, transform parse tree, translate the words
   - interlingua: translate to a language independent representation
   - example-based: use the closest match in training data
   - statistical: given sentence pairs with 1-1 word mapping, learn parameters of a model
   - hypbrid