

Final Review

Linguistics 384, Summer 07 (Martin-Lozano)

Date of the Exam: Friday July 20th, 1:30-3:18, Derby Hall 29

1 Topics to be covered

1. Writer's Aids
2. Machine Translation

2 Format of the exam

1. Term Matching: (see below)
2. Calculations:
 - Minimum Edit Distance
 - Venn Diagrams
 - Bag of Words Method
 - Context-free grammars
 - Syntactic Trees

3 Terms to know

3.1 Writers' Aids

- | | | |
|---------------------------------|-------------------------------------|----------------------------|
| – productivity | – automatic spelling correctors | – dictionary construction |
| – inflection | – phonetic errors | – insertion |
| – distributional properties | – run-on errors | – deletion |
| – tokenization | – split errors | – substitution |
| – error detection | – nonword error detection | – transposition |
| – error correction | – isolated-word error correction | – single-error misspelling |
| – spoonerism | – context-dependent word correction | – multi-error misspelling |
| – word recognition | – dictionary lookup | – semantics |
| – interactive spelling checkers | | |

- positional bigram array
- nonpositional bigram array
- domain-specificity
- minimum edit distance
- graph
- acyclic graph
- dynamic programming
- transition probabilities
- confusion probabilities
- local syntactic errors
- long-distance syntactic errors
- syntax
- linear order
- constituents
- syntactic tree
- phrasal category
- lexical category
- morphology
- function word
- open class
- closed class
- phrase structure rule
- generative
- structurally ambiguous
- hierarchical
- recursive
- context-free grammar
- parsing
- garden path sentence
- local syntactic errors
- long-distance syntactic errors
- mother (in a syntactic tree)
- daughters (in a syntactic tree)
- terminals (in a syntactic tree)
- non-terminals (in a syntactic tree)
- top-down parsing
- bottom-up parsing

3.2 Machine Translation

- translation
- machine translation
- (lexical) ambiguity
- structural ambiguity
- source language
- target language
- hyponym
- hypernym
- synonym
- interlingua
- lexicon
- lexical gap
- light verb
- idiom
- reversibility
- transformer architecture
- direct transfer
- transfer component
- comparative grammar
- robustness
- text alignment
- sentence alignment
- word alignment
- machine learning
- training data
- bag of words
- intelligibility
- accuracy
- error analysis
- test suite
- taboo words
- interlingua
- semantic overlap
- parallel corpus
- pragmatic knowledge