A corpus search methodology for focus realization

Jonathan Howell
Department of Linguistics
Computing and Information Science

Mats Rooth
Department of Linguistics

Project Goals
- Study phonetic realization of focus in cases where formal semantic theories make predictions
- Develop and apply methodology for accessing controlled and quantifiable, but diverse and naturally occurring speech data

Focus in Comparative Clauses
- Coherent syntactic-semantic theory about focus placement
- Possibilities constrained: main clause usually antecedent for focus interpretation in comparative clause

Subject focus
(1) He stayed longer than [I] did.
-er [[he stayed x long]], then [I stayed x long] ~2

Non-subject focus
(2) I should have that like song more than [I did],
more λx[should λw[I like that song x well in w]]
then [I like that song x well in w]
(3) I understand even less than I did [before].
even less [[I pres understand x much]],
then [I understood x much before] ~4

Corpus search methodology
- commercial search engine
- corpus of podcasts in mp3 format
- transcription and indexing with BBN speech recognition
- Web harvest
  - Retrieval workflow

Web harvest
- 179 alleged tokens of “than I did”
- of which 91 unique correct tokens (56%)

Subphonemic annotation
- Experimenter annotation using landmarks

Syntactic-semantic classification
- Transcription into English prose
- Experimenter binary classification into subject focus condition “s” and non-subject focus condition “ns”

Phonetic feature extraction
- automated with Praat scripts
- 308 acoustic measures
  - Quantity: F0, amplitude, intensity, power...
  - Quality: formants, bandwidth, tilt, jitter...
  - Duration: for vowels, stops, syllables
  - Measures taken: mean, extrema, range
  - Locus of measure: regular intervals, time of other extrema (e.g. F0 max, intensity max)

Supervised machine learning
- many parameters highly correlated (collinearity condition number k=127)
- want a focus detector for classifying novel/controversial data
- SVM (support vector machine) increasingly common in machine learning
- SVM uses separating hyperplanes (margins) rather than comparison of mean

Comparing statistical classifiers
- Classifiers evaluated by one-held-out cross validation (OHOVC)
- Best measures for each category (using feature selection techniques)

Comparing human classifiers
- Informal forced-choice discrimination task:
  - “s” condition vs. “ns” condition
  - Stimuli: 91 “than I did” snippets
  - Non-naive listeners (n=5)

Phonetic models of accent: hyperarticulation vs. sonority expansion
- Hyperarticulation / featural enhancement
- vowel features or gestural targets maximized
- Sonority expansion / tongue or jaw lowering
e.g. Edwards & Beckman (1988), Beckman et al. (1992), Erickson (2002)
- lower jaw / more open vocal tract

Predictions
- Preliminary results consistent with featural enhancement
References