Local Coherences in the Wild: An Eye-tracking Corpus Study

Klinton Bicknell¹, Vera Demberg², and Roger Levy¹

¹University of California, San Diego  ²University of Edinburgh

Abstract

The phenomenon of local coherences has recently generated significant interest in the sentence processing literature. Tabor, Galantucci, & Richardson (2004) showed that participants are slower to read object-modifying reduced relative clauses (RCs) when the RC verb is part-of-speech ambiguous (tossed) than when it is unambiguous (thrown). This result is problematic for any fully incremental framework because the main-verb interpretation is incompatible with the global syntactic context and should thus be ignored by the processor. An unresolved issue with this result, however, is the extent to which such effects occur in natural reading. We test the extent to which such effects occur in natural reading, using eye-tracking data from free reading of newspaper text. Our results support both of the major theories of local coherences, and add ecological validity to the study of local coherences.

1. Background and Motivation

background

Tabor, Galantucci, & Richardson (2004) varied verb ambiguity and relative clause resolution in object-modifying relative clauses, using sentences such as in (1):

(1) a. The coach smiled at the player tossed a frisbee . . . [+amb, +red]
b. The coach smiled at the player who was tossed a frisbee . . . [+amb, -red]
c. The coach smiled at the player who tossed a frisbee . . . [+amb, -red]
d. The coach smiled at the player who was thrown a frisbee . . . [-amb, -red]

- found an interaction such that [+amb, +red] was super-additively hard
- problematic result for any fully incremental framework because the main-verb interpretation is incompatible with the global syntactic context

theories

Tabor et al. (2004):
- difficulty arises because the ‘locally coherent’ parse of ‘the player tossed a frisbee’ competes with top-down syntactic expectations

Gibson (2006):
- difficulty arises from word part-of-speech (POS) disambiguation
- difficulty inversely proportional to context-independent probability of tag given word multiplied by the smoothed syntactic expectations

question

- both theories make broad claims about the parser from one example involving a very rare construction, a reduced relative of a passivized ditransitive
- how often do such effects arise in the reading of natural text?

plan

- build a model of the reading times in a corpus of naturalistic eye-tracking data
- test factors representing each of the two major theories

2. Corpus, Model, and Control Factors

 corpus
- used the Dundee corpus (Kennedy & Pynte, 2005) of eye-movement data from 10 participants reading 51,000 words of The Independent (a British newspaper)
- to get syntactic category information, we parsed the corpus with the Charniak parser (Charniak, 2000)

 model
- each of our factors was tested in a linear mixed-effect model (Bauyen, Davidson, & Bates, in Press) of the first pass times on each word, with a fixed effect control factors and participant as a random effect, as in Dembger & Keller (To Appear)

 control factors

- Demberg & Keller’s control factors included:

  linguistic properties:
  - word length and frequency
  - bigram probability
  - position in the sentence
  - lexicalized and unlexicalized syntactic surprisal (i.e., $-\log P(w_i|$context))

  eye movement properties:
  - landing position
  - launch distance
  - whether the previous word was fixated

3. Factor One: $P(t|w,s)$

 motivation

Gibson (2006) predicts a word $w$ to be read slower with a particular POS tag as the context-independent probability $P(t|w,s)$ decreases

 estimation
- for each word-tag pair in the Dundee corpus, we estimated $P(t|w,s)$ from a Charniak-parsed version of the British National Corpus (BNC)
- two versions of this factor:
  - $P_c$: maximum likelihood estimate (MLE)
  - $P_s$: MLE smoothed with a Bayesian prior, which is the type-averaged $P(t|w)$ (see box)

 results

- tested factor on all words in Dundee corpus that were found in the BNC

  - coefficient estimates
    - MLE (P_c): $r = 0.2$ ($t(199994) = -2.28; p < 0.05$) small effect (MLE too noisy?)
    - Smoothed (P_s): $r = 0.28$ ($t(199994) = -2.22; p < 0.05$) small effect

 further analysis
- performed a natural spline regression on $P(t|w)$ as before
- spline regression shows evidence of an effect for all values of the factor

4. Factor Two: $P(|w_i,w)$

 motivation

- Tabor et al. (2004) suggest that difficulty can arise when a locally coherent parse conflicts with a top-downparse
- we investigate a subset of these cases by using $P(t|w_i,w)$ as a factor

 estimation
- for each word-word-tag triplet in the Dundee corpus, we estimated $P(t|w_i,w)$ from the Charniak-parsed BNC
- two versions: $P_n = \text{MLE}$ $P_s = \text{MLE smoothed as before}$

 results

- tested factor on all words for which the bigram $w_i,w$ occurred in the BNC

  - coefficient estimates
    - MLE (P_n): $r = 0.2$ ($t(197556) = -2.28; p < 0.05$) small effect (MLE too noisy?)
    - Smoothed (P_s): $P(t|w_i,w) = -16.7; p = 0.0011$ large effect

 further analysis
- performed a natural spline regression on $P(t|w_i,w)$ as before
- spline regression shows evidence of an effect for all values of the factor

5. Caveat

- our factors were somewhat correlated with control factors:
- because these control factors were explicitly included in the model, we have some confidence that our effect is ‘above and beyond’ them
- nevertheless, we plan to address this in the future by:
  - residualizing the covariates
  - constructing more estimates of these and other measures of conflict between top-down and bottom-up probabilities

6. Conclusion

 summary
- evidence for both Tabor et al. and Gibson’s theories of local coherence effects
- adds ecological validity to the study of local coherences

 next step
- test a larger subset of cases predicted by Tabor et al. to be difficult by using a modified parser to predict conflict between top-down and bottom-up parses

References