Readers maintain and act on uncertainty about past linguistic input: Evidence from eye movements

Roger Levy    Klinton Bicknell    Tim Slattery    Keith Rayner
UC San Diego

CUNY Sentence Processing Conference
26 March 2009
Incrementality and Rationality

- Online sentence comprehension is hard
- But lots of information sources can be usefully brought to bear to help with the task
- Therefore, it would be *rational* for people to use *all the information available*, whenever possible
- This is what *incrementality* is
- We have lots of evidence that people do this often

“Put the apple on the towel in the box.”  (Tanenhaus et al., 1995)
Probabilistic models are good at capturing this use of evidential information (Jurafsky, 1996; Crocker & Brants, 2000; Hale, 2001; Levy, 2008)

(McRae et al., 1998)
Grammar ↔ Word modularity? (II)

- But implicit modularity assumption: decisions made during *word recognition* aren’t subject to revision

- Caveat: It’s clear that prior context can affect the recognition process for newly perceived words…

  - He swept the *flour*…
  - The baker needed more *flour*…

  *(Slattery, submitted)*

- …but in most models, the *outcome* of word recognition is “just a word”
Grammar ↔ Word modularity? (III)

• This partition between word-level and sentence level analysis could be real (Fodor, 1983)
• Or, it could be a practical modeling simplification
• *In this talk*, we use eye-tracking data to suggest that it is indeed a practical simplification
• We argue:
  • Readers *retain residual uncertainty* in their beliefs about the identities of words already read
  • Readers *change these beliefs* on the basis of these words’ consistency with subsequent input
  • Readers *rapidly act on these changes* via regressive eye movements
• We use local-coherence sentences (Tabor et al., 2004; Konieczny & Mueller, 2007) to show this
The puzzle of local coherences

• Try to understand this sentence:

(a) *The coach smiled at the player tossed the frisbee.*

…and contrast this with:

(b) *The coach smiled at the player thrown the frisbee.*

(c) *The coach smiled at the player who was thrown the frisbee.*

(d) *The coach smiled at the player who was tossed the frisbee.*

• Readers boggle at “tossed” in (a), but not in (b-d)
Why is \textit{tossed/thrown} interesting?

- As with classic garden-paths, part-of-speech ambiguity leads to misinterpretation
  - \textit{The horse raced past the barn…fell}
  - \textit{verb? participle?}

- But now context “should” rule out the garden path:
  - \textit{The coach smiled at the player tossed…}
  - \textit{verb? participle?}

- A challenge for fully incremental probabilistic models: \textit{failure to condition on relevant context}
Levy (2008) suggested that local coherences could arise from presence of “near-neighbor” sentences…

Hypothesis: the boggle at “tossed” involves what the comprehender wonders whether she might have seen

Levy (2008) quantified this boggle as the size of change in the probability distribution over earlier contexts (Error Identification Signal, or EIS)
The core of the intuition

- Grammar & input come together to determine two possible “paths” through the partial sentence: *(line thickness \(\approx\) probability)*

- *tossed* is more likely to happen along the bottom path
  - This creates a large shift in belief in the *tossed* condition
- *thrown* is very unlikely to happen along the bottom path
  - As a result, there is no corresponding shift in belief
Results for classic local-coherence sentence

The coach smiled at the player *tossed*

EIS greater for the variant humans *boggle more on*

The coach smiled at the player *thrown*
Near-neighbor manipulation

• Since the EIS depends on near-neighbor sentences, vary neighborhood of the context?

\[\text{The coach smiled at the player tossed the frisbee}\]

\[\text{The coach smiled toward the player tossed the frisbee}\]

• Substituting \textit{toward} for \textit{at} should reduce the EIS

• In free reading, we should see less tendency to regress from \textit{tossed} when the EIS is small
Indeed we do see this behavior predicted in the model:
Experimental design

• In a free-reading eye-tracking study, we crossed *at/*toward with *tossed/thrown*:

The coach smiled *at* the player *tossed* the frisbee
The coach smiled *at* the player *thrown* the frisbee
The coach smiled *toward* the player *tossed* the frisbee
The coach smiled *toward* the player *thrown* the frisbee

• Prediction: interaction between preposition & ambiguity in some subset of:
  • Early-measure RTs at critical region *tossed/thrown*
  • First-pass regressions out of critical region
  • Go-past time for critical region
  • Regressions into *at/toward*
Procedure

- 24 items + 36 fillers, 40 participants
- Isolated-sentence reading
- Each sentence followed by a comprehension question
- Eye movements monitored with Eyelink 2000 tower setup
- Regions of analysis:

    1        2          3           4        5 (critical)       6
The coach | smiled | {at, toward} | the player | {tossed, thrown} | the frisbee | ...
## Experimental results

<table>
<thead>
<tr>
<th>Condition</th>
<th>FirstPass RT</th>
<th>Gopast RT</th>
<th>Regressions out</th>
<th>Regressions in</th>
</tr>
</thead>
<tbody>
<tr>
<td>at/tossed</td>
<td>349</td>
<td>476</td>
<td>21%</td>
<td>36%</td>
</tr>
<tr>
<td>at/thrown</td>
<td>312</td>
<td>383</td>
<td>12%</td>
<td>31%</td>
</tr>
<tr>
<td>toward/tossed</td>
<td>352</td>
<td>390</td>
<td>10%</td>
<td>31%</td>
</tr>
<tr>
<td>toward/thrown</td>
<td>331</td>
<td>409</td>
<td>14%</td>
<td>31%</td>
</tr>
</tbody>
</table>

Main effect of category ambiguity (p<0.05)

Interaction with at/tossed slowest (p<0.05)

Interaction with most regressions in at/tossed (p<0.01)

n.s., but right trend

Critical region

at/toward

0.00  0.05  0.10  0.15  0.20

EIS

Noise level (low=noisy)

at...tossed

toward...tossed

at...thrown

toward...thrown
Where did readers regress to?

First regressive saccade from critical region rarely goes straight back to *at*:

The coach smiled at/toward the player {tossed, thrown}...
Where did readers regress to?

- Not much action going on from critical region:
Where did readers regress to? (II)

- **But** readers jump farther in first regressive saccade from the spillover region:

Interaction $t=1.934$ in random-intercepts multi-level model

Main effect of ambiguity; maybe an interaction too (needs further analysis before CUNY – random intercepts or random interactions?)

(no such pattern in regressions from critical region)
How far did readers continue to regress?

- More information contained in *regression sequences*

  2  3 (Prep)  4  5 (critical)
  
  | smiled | \{at, toward\} | the player | \{tossed, thrown\} | ...

- First-pass rate of ultimately regressing to Prep after fixating on Critical and before passing Critical:

  Interaction \(p<0.01\) in several types of multi-level analyses

  This bar is 1/3 of all first-pass regressions from the critical region!
16 of 24 questions were about the RRC, equally divided:

- *The coach smiled at the player tossed a frisbee by the opposing team*
  - Did the player toss/throw a frisbee? [NO]
  - Did someone toss/throw the player a frisbee? [YES]
  - Did the player toss the opposing team a frisbee? [NO]
  - Did the opposing team toss the player a frisbee? [YES]

**Significant main effect of question type ($p_z < 0.01$)**

**Significant interaction of question type w/ ambiguity ($p_z < 0.05$)**
What this result tells us

- Readers must have residual uncertainty about word identity
  - Word misidentification alone won’t get this result in a fully incremental model:

  The coach smiled toward the player. \textcolor{cyan}{thrown}

  The coach smiled at the player. \textcolor{cyan}{thrown}

  The coach smiled as the player. \textcolor{cyan}{thrown}

  The coach smiled toward the player. \textcolor{cyan}{tossed}

  The coach smiled at the player. \textcolor{cyan}{tossed}

  The coach smiled as the player. \textcolor{cyan}{tossed}

- Also, readers respond to changes in uncertainty in a sensible way

  Should be about equally hard

  Should be easier, if anything
What this result does *not* tell us

- Doesn’t show that Levy (2008)’s error-identification model is the only model that can get this results
  - But any model without uncertainty about prior-word identities of prior words has its work cut out for it
- Doesn’t show that local-coherence effects are exclusively a consequence of rational inference
  - But provides more supporting evidence that rational inference may be at least part of the story
Conclusion

• First experimental results indicating full bidirectional interactivity between word identification & sentence comprehension

• Readers…
  • …maintain uncertain beliefs about already-read word identities
  • …question these beliefs as a function of coherence with subsequent input
  • …act rapidly on changes in these beliefs through saccadic behavior

• Strengthens the case for sentence comprehension as rational probabilistic inference

• Methodological point: pay attention to the near-neighbors of your stimuli! Your participants will…
Acknowledgments

• Computational Psycholinguistics Lab
• UCSD Academic Senate
• NIH HD26765
• Jennifer Arnold for kindly lending me an AC adapter 😊
Thank you for listening!

http://grammar.ucsd.edu/labs/cpl

http://psy3.ucsd.edu/~raynerlab
Acknowledgements

- Thank you for all your help!
  - Collaborators:
    - Klinton Bicknell
    - Tom Griffiths
    - Keith Rayner
    - Florencia Reali
    - Tim Slattery
  - UCSD: Computational Psycholinguistics Lab & CRL
  - Liang Huang & David Chiang
  - Cyril Allauzen and the rest of the developers of the OpenFst (http://openfst.org) and GRM libraries (http://www.research.att.com/~fsmtools/grm/)
Behavioral correlates (Tabor et al., 2004)

- Also, Konieczny (2006, 2007) found compatible results in stops-making-sense and visual-world paradigms.

- **These results are problematic for theories requiring global contextual consistency** (Frazier, 1987; Gibson, 1991, 1998; Jurafsky, 1996; Hale, 2001, 2006)
Eye movements & question answering

- Hard to find much of anything
- But there was a significant relationship with first-pass regressions out of the critical region, only for RRC questions

<table>
<thead>
<tr>
<th>with</th>
<th>No Regression</th>
<th>Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>at ambig</td>
<td>0.50 (0.05)</td>
<td>0.55 (0.09)</td>
</tr>
<tr>
<td>at unambig</td>
<td>0.72 (0.04)</td>
<td>0.72 (0.11)</td>
</tr>
<tr>
<td>toward ambig</td>
<td>0.59 (0.04)</td>
<td>0.83 (0.11)</td>
</tr>
<tr>
<td>toward unambig</td>
<td>0.71 (0.04)</td>
<td>0.50 (0.11)</td>
</tr>
</tbody>
</table>

- Significant main effect of ambiguity ($p_z < 0.001$)
- Small but significant interaction of ambiguity w/ regression ($p_z < 0.05$)
- The predictions could have gone either way, though
How often did readers skip the prep?

- They definitely skipped “at” a lot more than skipping “toward”!

<table>
<thead>
<tr>
<th></th>
<th>tossed</th>
<th>thrown</th>
</tr>
</thead>
<tbody>
<tr>
<td>at</td>
<td>40.4%</td>
<td>33.8%</td>
</tr>
<tr>
<td>toward</td>
<td>2.9%</td>
<td>2.5%</td>
</tr>
</tbody>
</table>

- (main effect of *at/toward* only)
Bad comprehenders?

- Well, filler QA accuracy was 89.6%