Incrementality (so far)

Evidences from Prediction


Soko!

The woman will spread the butter on the bread.
The woman will slide the butter to the man.
Incrementality (so far)

Reading_time_effect()

The readers/listeners have anticipations about what comes next;

if (continuation ~ anticipation)

   Faster reading time; //structures already made 😊

else

   Slower reading time; //reanalysis required

}
Incrementality (so far)

Questions about Ambiguity

Swels et al. 2008

Tabor et al. 2004

Global Ambiguity
Faster reading!!

Local Ambiguity
Slower reading!!
Overview

- Background
  - Top-down vs. Bottom-up constraints in parsing
  - Different parsing theories

- Papers
  - Effect of merely local syntactic coherence on sentence processing
    Tabor, Galantucci and Richardson. 2004
  - The interaction of top-down and bottom-up statistics in the resolution of syntactic category ambiguity
    Gibson. 2005

- Summary
Background

Top-down vs. Bottom-up constraints

The lawyer visited that cheep hotel to stay for the night.
Background
Top-down vs. Bottom-up constraints

The lawyer visited that cheep hotel to stay for the night.
Background
Top-down vs. Bottom-up constraints

The lawyer visited that cheep hotel ...
Background
Top-down vs. Bottom-up constraints

The lawyer visited ... that cheep hotel ...
Background
Top-down vs. Bottom-up constraints

S

NP

The lawyer

VP

visited ... that cheep hotel ...

Top-down

That (det)
That (pro)
That (com)...

Bottom-up
Background
Different parsing accounts

- Local coherence accounts (LCA):
  - Self-organized parsing
  - Fixed-width buffers
  - First-pass failure

- Self-consistent parsing accounts (SCPA):
  - Full grammatical parsing
Background
Different parsing accounts

- Local coherence accounts (LCA):
  - Self-organized parsing
  - Fixed-width buffers
  - First-pass failure

The lawyer visited ... that cheap hotel ...
Background
Different parsing accounts

- Local coherence accounts (LCA):
  - Self-organized parsing
  - Fixed-width buffers
  - First-pass failure

The lawyer visited that cheap hotel...
Background
Different parsing accounts

- Local coherence accounts (LCA):
  - Self-organized parsing
  - Fixed-width buffers
  - Firs-pass failure

The lawyer visited ... that cheep hotel ... ?

The lawyer visited that cheep hotel ...
DET ADJ N
Background
Different parsing accounts

- Self-consistent parsing accounts (SCPA):
  - Full grammatical parsing
Background
Different parsing accounts

- Self-consistent parsing accounts (SCPA):
  - Full grammatical parsing

Grammar

The lawyer visited ... that cheap hotel ...
The goal is to examine these contrasting parsing accounts:
- Partial parses are sometimes constructed (LCA)
- Only parses consistent with the whole input are constructed (SCPA)

Experiments:
1: Syntactic local coherence increases reading time (RT)
2: Semantic support for a local parse increases RT even more
3: People judge sentences with local coherence as ungrammatical
Papers
Local Syntactic Coherence. Tabor et al. 2004

- Experiment 1: local coherence (syntactic) slows reading?
  
a. The coach smiled at the player tossed a Frisbee by the opposing team.

b. The coach smiled at the player who was tossed a Frisbee by the opposing team.
Papers
Local Syntactic Coherence. Tabor et al. 2004

- Experiment 1: Items

<table>
<thead>
<tr>
<th>-6</th>
<th>-5</th>
<th>-4</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>The coach smiled at the player...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. tossed a frisbee by... (A/R)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. who was tossed a frisbee by... (A/U)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. thrown a frisbee by... (U/R)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d. who was thrown a frisbee by... (U/U)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ambiguous

Unambiguous
Papers
Local Syntactic Coherence. Tabor et al. 2004

- Experiment 1: Procedure
  - Self-paced reading

The coach smiled at the player tossed a Frisbee by the ...

----- ----- ----- ----- ----- tossed ----- ----- ----- ----- ...
Experiment 1: Results

Fig. 1. Mean residual reading times from Experiment 1. Error bars show one standard error around each data point.
Papers
Local Syntactic Coherence. Tabor et al. 2004

- Experiment 1: Verification
  - Argument: “throwing a Frisbee to a player” a typical co-occurrence?
  - Answer: No significant difference in rating

(7) (a) Someone tossed a frisbee to the player. (5.83, SE = 0.14)
    (b) Someone threw a frisbee to the player. (5.63, SE = 0.13)
Papers
Local Syntactic Coherence. Tabor et al. 2004

- Experiment 1: Verification
  - Argument: “thrown” recipient interpretation more acceptable?
  - Answer: No significant difference between verb types.

(8) (a) The player was tossed a frisbee by the opposing team. (Recipient/Ambiguous) $4.30, SE = 0.15$
(b) A frisbee was tossed to the player by the opposing team. (Theme/Ambiguous) $5.56, SE = 0.10$
(c) The player was thrown a frisbee by the opposing team. (Recipient/Unambiguous) $4.39, SE = 0.15$
(d) A frisbee was thrown to the player by the opposing team. (Theme/Unambiguous) $5.42, SE = 0.11$
Papers
Local Syntactic Coherence. Tabor et al. 2004

- Experiment 1: Verification
  - Argument: “thrown” is biased towards recipient-extracted passive sense while “tossed” is biased toward active interpretation.
  - Answer: Experiment 2 (let’s keep ambiguous verbs the same between cases).
Papers
Local Syntactic Coherence. Tabor et al. 2004

- Experiment 2: local coherence (syntactic + semantic) slows reading?

  a. The bandit worried about the prisoner transported by the capricious guards.

  b. The bandit worried about the gold transported by the capricious guards.
Experiment 2: Items

-2  -1     0       1
the  prisoner  transported  the
the  prisoner  who        transported  the
the  gold       was        transported  the
the  gold       that       was        transported  the
Papers
Local Syntactic Coherence. Tabor et al. 2004

- Experiment 2: Results

![Graph showing residual reading times from Experiment 2.](image)

Fig. 2. Mean residual reading times from Experiment 2. Error bars show one standard error around each data point.
Papers
Local Syntactic Coherence. Tabor et al. 2004

- **Experiment 2: Verification**
  - Argument: “transporting the gold” a typical **co-occurrence**?
  - Answer: No significant difference in (target) passive.

(11) The prisoner was transported. (Animate/Passive) (6.06, $SE = 0.12$)
The prisoner transported something. (Animate/Active) (6.03, $SE = 0.12$)
The gold was transported. (Inanimate/Passive) (6.11, $SE = 0.12$)
The gold transported something. (Inanimate/Active) (2.16, $SE = 0.14$)
Experiment 2: Verification

- Argument: Any semantic interpretation of the locally parsed text?
- Answer: “yes” answers to the second question increased with Reduction.

(12) The bandit worried about the prisoner (who was) transported the whole way.

(a) Was the prisoner transported? (prisoner = Theme of transporting)
(b) Did the prisoner transport something? (prisoner = Agent of transporting)
Papers
Local Syntactic Coherence. Tabor et al. 2004

- Experiment 3: Does grammaticality judgment interact with local coherence in a same way as reading time does?

The coach smiled at the player tossed a Frisbee
Experiment 3: Items

3.1: The coach smiled at the player (who was) tossed/thrown a Frisbee.

3.2: The bandit talked remorsefully of the prisoner/gold (who/that was) transferred the whole way by the guards.
Papers
Local Syntactic Coherence. Tabor et al. 2004

- Experiment 3: Procedure

- The coach smiled at the player tossed a Frisbee. 
  Grammatical? (yes/no)
### Experiment 3: Result

#### Table 4
Mean rates of positive grammaticality judgments for the two subexperiments and the fillers in Experiment 3

<table>
<thead>
<tr>
<th>Subexperiment 3.1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambiguous/Reduced</td>
<td>0.22</td>
<td>(0.040)</td>
</tr>
<tr>
<td>Ambiguous/Unreduced</td>
<td>0.72</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Unambiguous/Reduced</td>
<td>0.39</td>
<td>(0.042)</td>
</tr>
<tr>
<td>Unambiguous/Unreduced</td>
<td>0.72</td>
<td>(0.038)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subexperiment 3.2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Animate/Reduced</td>
<td>0.58</td>
<td>(0.045)</td>
</tr>
<tr>
<td>Animate/Unreduced</td>
<td>0.82</td>
<td>(0.030)</td>
</tr>
<tr>
<td>Inanimate/Reduced</td>
<td>0.75</td>
<td>(0.035)</td>
</tr>
<tr>
<td>Inanimate/Unreduced</td>
<td>0.84</td>
<td>(0.032)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fillers</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical</td>
<td>0.93</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Ungrammatical</td>
<td>0.08</td>
<td>(0.013)</td>
</tr>
</tbody>
</table>

The values shown in parentheses are standard errors.
Conclusions (with respect to incrementality):

- **Local parsing** happens where local coherence exists.
- Constructing a local parse **inconsistent with the global parse** suggests that preceding information is not maximally applied in parsing.
- Ignoring information from previous steps is **against incrementality**.
Papers
Top-down and Bottom-up Statistics. Gibson 2005

- Goals:
  - Top-down & bottom-up information contribution
  - Formulation suggestion (context-independent hypothesis)
  - Against context-dependent hypothesis (Tabor et al 1997)
  - Defend parallel processing

- Experiments:
  1 & 2 : Ambiguous words might increase RT independently from the context
  3: People keep track of alternative interpretations until it can be resolved
Papers
Top-down and Bottom-up Statistics. Gibson 2005

- Same word in different environments (from Tabor et al. 1997)

(1) a. That cheap hotel was clean and comfortable to our surprise.
    b. That cheap hotels were clean and comfortable surprised us.

(2) a. The lawyer insisted that cheap hotel was clean and comfortable.
    b. The lawyer insisted that cheap hotels were clean and comfortable.
Papers
Top-down and Bottom-up Statistics. Gibson 2005

- Different words in the same environment (from Tabor et al. 1997)

(3) a. The lawyer visited that cheap hotel to stay for the night.
   b. The lawyer visited those cheap hotels to stay for the night.
Papers
Top-down and Bottom-up Statistics. Gibson 2005

- Context-dependent category-frequency hypothesis

The lawyer visited ... that cheap hotel ...
Papers
Top-down and Bottom-up Statistics. Gibson 2005

- Context-independent category-frequency hypothesis

The lawyer visited that cheap hotel...
Papers
Top-down and Bottom-up Statistics. Gibson 2005

- Context-independent category-frequency hypothesis

\[
P(\text{“that” being tagged as complementizer in this sentence}) =
\]

\[
P(\text{“that” being a complementizer}) \times P(\text{a complementizer following a verb})
\]

Independent info

Environmental info
Experiment 1: “that” slows down reading independently from the context?

- The lawyer for that skilled surgeon asked for a raise.
- The lawyer for those skilled surgeons asked for a raise.
Experiment 1: Items

(11) a. Prep, that: The lawyer for that skilled surgeon asked for a raise.
    b. Prep, those: The lawyer for those skilled surgeons asked for a raise.
    c. Verb, that: The lawyer visited that skilled surgeon before the hearings began.
    d. Verb, those: The lawyer visited those skilled surgeons before the hearings began.

Tested by Tabor 1997
Papers
Top-down and Bottom-up Statistics. Gibson 2005

- Experiment 1: Results

“that” read slower

![Graph showing reading times for different phrases](image)
Experiment 1: Results

- Significant determiner effect
- No significant category*determiner effect

Table 5
Analysis of Variance results for Experiment 1

<table>
<thead>
<tr>
<th>Source of variance</th>
<th>df</th>
<th>min F* value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Position 3</strong> <em>(“for”</em>/“visited”)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1.47</td>
<td>1.46</td>
</tr>
<tr>
<td>Determiner</td>
<td>1.47</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>1.47</td>
<td>&lt;1</td>
</tr>
<tr>
<td><strong>Position 4</strong> <em>(“that”</em>/“those”)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1.47</td>
<td>4.85*</td>
</tr>
<tr>
<td>Determiner</td>
<td>1.47</td>
<td>1.26</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>1.47</td>
<td>&lt;1</td>
</tr>
<tr>
<td><strong>Position 5</strong> <em>(“skilled”)</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1.47</td>
<td>4.19*</td>
</tr>
<tr>
<td>Determiner</td>
<td>1.47</td>
<td>3.26</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>1.47</td>
<td>&lt;1</td>
</tr>
<tr>
<td><strong>Position 6</strong> <em>(“surgeon(s)”)</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1.47</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Determiner</td>
<td>1.47</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>1.47</td>
<td>&lt;1</td>
</tr>
<tr>
<td><strong>Position 7</strong> <em>(“asked”</em>/“before”)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1.47</td>
<td>2.31</td>
</tr>
<tr>
<td>Determiner</td>
<td>1.47</td>
<td>3.77</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>1.47</td>
<td>2.04</td>
</tr>
<tr>
<td><strong>Positions 8-end of sentence</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1.47</td>
<td>1.89</td>
</tr>
<tr>
<td>Determiner</td>
<td>1.47</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>1.47</td>
<td>1.80</td>
</tr>
<tr>
<td><strong>Positions 4–6</strong> <em>(“that”</em>/“those”)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1.47</td>
<td>3.24</td>
</tr>
<tr>
<td>Determiner</td>
<td>1.47</td>
<td>3.29</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>1.47</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>
Experiment 1: Results

- Small effect of determiner in verbal case
- Verbal “those” condition read slower after critical region
Experiment 2: Replication of experiment 1 to see…

- Even more significant slower reading time for “that” in verbal versions?
- Any similar pattern when comparing with “this”?
Experiment 2: Items

(12) a. Prep, that: The lawyer for that skilled surgeon asked for a raise.
    b. Prep, those: The lawyer for those skilled surgeons asked for a raise.
    c. Prep, this: The lawyer for this skilled surgeon asked for a raise.
    d. Verb, that: The lawyer visited that skilled surgeon before the hearings began.
    e. Verb, those: The lawyer visited those skilled surgeons before the hearings began.
    f. Verb, this: The lawyer visited this skilled surgeon before the hearings began.
### Experiment 2: Results

#### Table 7

Analysis of Variance results for Experiment 2

<table>
<thead>
<tr>
<th>Source of variance</th>
<th>By participants</th>
<th>By items</th>
<th>min F'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>df</td>
<td>F1 value</td>
<td>MSE</td>
</tr>
<tr>
<td><strong>Position 3 (&quot;for&quot;/&quot;i&quot;/visited&quot;)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1,95</td>
<td>3.63</td>
<td>19669</td>
</tr>
<tr>
<td>Determiner</td>
<td>2,190</td>
<td>1.75</td>
<td>15381</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>2,190</td>
<td>&lt;1</td>
<td>20605</td>
</tr>
<tr>
<td><strong>Position 4 (&quot;that&quot;/&quot;i&quot;/these&quot;/this&quot;)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1,95</td>
<td>53.3*</td>
<td>4432</td>
</tr>
<tr>
<td>Determiner</td>
<td>2,190</td>
<td>3.40*</td>
<td>3140</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>2,190</td>
<td>1.44</td>
<td>3790</td>
</tr>
<tr>
<td><strong>Position 5 (&quot;skilled&quot;)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1,95</td>
<td>6.97*</td>
<td>6050</td>
</tr>
<tr>
<td>Determiner</td>
<td>2,190</td>
<td>17.41*</td>
<td>5939</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>2,190</td>
<td>1.03</td>
<td>4966</td>
</tr>
<tr>
<td><strong>Position 6 (&quot;surgeon(s)&quot;)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1,95</td>
<td>5.65*</td>
<td>4556</td>
</tr>
<tr>
<td>Determiner</td>
<td>2,190</td>
<td>5.84*</td>
<td>5418</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>2,190</td>
<td>&lt;1</td>
<td>3851</td>
</tr>
<tr>
<td><strong>Position 7 (&quot;asked&quot;/&quot;i&quot;/before&quot;)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1,95</td>
<td>2.70</td>
<td>5377</td>
</tr>
<tr>
<td>Determiner</td>
<td>2,190</td>
<td>2.77</td>
<td>8866</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>2,190</td>
<td>3.01</td>
<td>6575</td>
</tr>
<tr>
<td><strong>Positions 8-end of sentence</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1,95</td>
<td>&lt;1</td>
<td>1441</td>
</tr>
<tr>
<td>Determiner</td>
<td>2,190</td>
<td>&lt;1</td>
<td>1609</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>2,190</td>
<td>1.93</td>
<td>1611</td>
</tr>
<tr>
<td><strong>Positions 4-6 (&quot;that/this/this skilled surgeon(s)&quot;)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>1,95</td>
<td>33.05*</td>
<td>2449</td>
</tr>
<tr>
<td>Determiner</td>
<td>2,190</td>
<td>15.53*</td>
<td>2302</td>
</tr>
<tr>
<td>Category × Determiner</td>
<td>2,190</td>
<td>1.64</td>
<td>1940</td>
</tr>
</tbody>
</table>

*Note: Significant effects are marked by asterisk.*
Experiment 3: sentence-initial effect of “that” due to parallel processing?

That experienced diplomats would be very helpful made the layer confident.

or a single arbitrary parse will be fast selected among possible ones, and only after reaching a disambiguating point in the sentence we observe a slower reading due to reanalysis?
Experiment 3: Items

(15) a. “that”-determiner: That experienced diplomat would be very helpful to the lawyer.
b. “those”-determiner: Those experienced diplomats would be very helpful to the lawyer.
c. “that”-complementizer: That experienced diplomats would be very helpful made the lawyer confident.
Experiment 3: Results

Longer reading time for “that” (even though it is shorter and more frequent)

Fig. 4. Residual reading times on the word “that”/“those” in sentence-initial context in Experiment 3.
A flash back to Tabor et al. 2004 (exp1):

✓ “tossed” was read slower
✓ Even when the alternative tags were not possible

Fig. 1. Mean residual reading times from Experiment 1. Error bars show one standard error around each data point.
Papers
Top-down and Bottom-up Statistics. Gibson 2005

- A flash back to Tabor et al. 2004 (exp2,3):
  - “transformed” with various behaviors!
  - A new component required in formulation (thematic plausibility)

![Graph showing residual reading times](image)

*Fig. 2. Mean residual reading times from Experiment 2. Error bars show one standard error around each data point.*
Papers
Top-down and Bottom-up Statistics. Gibson 2005

- Conclusions (with respect to incrementality):
  - Context-independent category statistics affect parsing
  - Sentence-initial ambiguity effects suggest parallel processing
  - Parallelism is not in line with strict incrementality
Questions?

Fatemeh Torabi Asr
Spring 2011