The syntactic complexity of Russian relative clauses

Roger Levy    Evelina Fedorenko   Ted Gibson
UC San Diego   MIT               MIT

CUNY 2007, 31 March
Big picture

- Why is online sentence comprehension interesting?
Big picture

- Why is online sentence comprehension interesting?

- Finite set of knowledge must be deployed in real time to process potentially unbounded set of inputs
Big picture

- Why is online sentence comprehension interesting?

- Finite set of knowledge must be deployed in real time to process potentially unbounded set of inputs

- What are the major cognitive constraints governing the deployment of this knowledge?
Big picture

- Why is online sentence comprehension interesting?
- Finite set of knowledge must be deployed in real time to process potentially unbounded set of inputs
- What are the major cognitive constraints governing the deployment of this knowledge?
- *Relative clauses* have played a crucial role in the study of this problem
Overview

- What determines difficulty of an unambiguous relative clause?
  - Locality? (Gibson 2000)
  - Time-based decay? (Lewis et al., 2006)
  - Word order? (MacDonald & Christiansen 2002)
  - Entropy reduction hypothesis (ERH)? (Hale 2003)
  - Perspective shift? (MacWhinney & Pleh, 1988)
  - Structural asymmetry? (Lin & Bever 2006)
  - Similarity-based interference (SBI)? (Lewis 1996)
- These theories difficult to distinguish using English data alone
- Today: add to the existing cross-linguistic picture with two experiments on Russian relative clauses
The classic result & explanations (1)

Subject RC (SRC): the reporter that **attacked** the senator    EASY

Object RC (ORC): the reporter that the senator **attacked**    HARD
The classic result & explanations (1)

Subject RC (SRC): the reporter that **attacked** the senator EASY

Object RC (ORC): the reporter that the senator **attacked** HARD

- Locality/integration complexity
The classic result & explanations (1)

Subject RC (SRC): the reporter that attacked the senator EASY

Object RC (ORC): the reporter that the senator attacked HARD

- Locality/integration complexity
  - SRC dependencies get constructed one at a time
  - two ORC dependencies must get constructed at once
The classic result & explanations (1)

Subject RC (SRC): the reporter that **attacked** the senator EASY

Object RC (ORC): the reporter that the senator **attacked** HARD

- Locality/integration complexity
  - SRC dependencies get constructed one at a time
  - two ORC dependencies must get constructed at once
The classic result & explanations (1)

Subject RC (SRC): \textit{the reporter that attacked the senator} \hspace{1cm} \text{EASY}

Object RC (ORC): \textit{the reporter that the senator attacked} \hspace{1cm} \text{HARD}

- Locality/integration complexity
  - SRC dependencies get constructed one at a time
  - two ORC dependencies must get constructed at once
- Time-based decay:
  - the distance from an early dependent is greater in ORC
### The classic result & explanations (2)

<table>
<thead>
<tr>
<th>Subject RC (SRC):</th>
<th>the reporter that attacked the senator</th>
<th>EASY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object RC (ORC):</td>
<td>the reporter that the senator attacked</td>
<td>HARD</td>
</tr>
</tbody>
</table>
The classic result & explanations (2)

Subject RC (SRC): the reporter that attacked the senator EASY

Object RC (ORC): the reporter that the senator attacked HARD

- Word order theories
  - SRCs preserve SVO word order, common in English
  - ORCs are OSV
The classic result & explanations (2)

**Subject RC (SRC):**  the reporter that attacked the senator  EASY

**Object RC (ORC):**  the reporter that the senator attacked  HARD

- Word order theories
  - SRCs preserve SVO word order, common in English
  - ORCs are OSV
- Expectations/surprisal
  - Given the reporter that..., conditional probability of SRC is higher than probability of ORC
The classic result & explanations (3)

Subject RC (SRC):  the reporter that attacked the senator  EASY

Object RC (ORC):  the reporter that the senator attacked  HARD
The classic result & explanations (3)

Subject RC (SRC):  the reporter that attacked the senator  EASY

Object RC (ORC):  the reporter that the senator attacked  HARD

- Entropy reduction
  - Crossing a possibly recursive category (postmodifiers of an NP) leads to a big drop in entropy
The classic result & explanations (3)

Subject RC (SRC): the reporter that **attacked** the senator EASY

Object RC (ORC): the reporter that the senator **attacked** HARD

- Entropy reduction
- Crossing a possibly recursive category (postmodifiers of an NP) leads to a big drop in entropy
The classic result & explanations (3)

Subject RC (SRC): the reporter that attacked the senator EASY

Object RC (ORC): the reporter that the senator attacked HARD

- Entropy reduction
  - Crossing a possibly recursive category (postmodifiers of an NP) leads to a big drop in entropy
The classic result & explanations (3)

Subject RC (SRC): the reporter that attacked the senator EASY

Object RC (ORC): the reporter that the senator attacked HARD

- Entropy reduction
- Crossing a possibly recursive category (postmodifiers of an NP) leads to a big drop in entropy
The classic result & explanations (3)

Subject RC (SRC): the reporter that attacked the senator EASY

Object RC (ORC): the reporter that the senator attacked HARD

- Entropy reduction
  - Crossing a possibly recursive category (postmodifiers of an NP) leads to a big drop in entropy

- Perspective shift
  - Subject/subject match of grammatical function in commonly investigated SRCs
The classic result & explanations (4)

Subject RC (SRC): the reporter that attacked the senator EASY

Object RC (ORC): the reporter that the senator attacked HARD

- Structural asymmetry
  - Subject RCs are universally simpler and hence should be easier to process
The classic result & explanations (4)

Subject RC (SRC): the reporter that **attacked** the senator  EASY

Object RC (ORC): the reporter that the senator **attacked**  HARD

- Structural asymmetry
  - Subject RCs are universally simpler and hence should be easier to process
- Similarity-based interference
  - the two preceding NPs in an ORC are confusible, creating integration difficulty at the verb
The classic result & explanations (4)

Subject RC (SRC): the reporter that attacked the senator  EASY

Object RC (ORC): the reporter that the senator attacked  HARD

- Structural asymmetry
  - Subject RCs are universally simpler and hence should be easier to process
- Similarity-based interference
  - the two preceding NPs in an ORC are confusable, creating integration difficulty at the verb
Today

- Use Russian (free word order, SVO-dominant, case marking) to distinguish among these theories
- Two experiments on Russian relative clauses
- Experiment 1:
  - Two factors: subject versus object RC and locality/scrambledness of word order
- Experiment 2:
  - Parametric manipulation of word order in subject RCs
  - Crossed with argument/adjunct manipulation
- Both experiments are self-paced reading studies conducted in Russia
Experiment 1

- Investigate *subject* versus *object* relative clauses
- Cross w/ “scrambledness” of remaining verb+argument pair
  - also interpretable as *locality* of relative pronoun/verb relationship
- Word-by-word self-paced reading
- 40 native Russian speakers (Volgograd, Russia; 5 removed due to low question-answering accuracy)
Subject RCs

- Unscrambled (Verb+RelPro local):

  Subj
  Diktator kotoryj nenavidel dissidenta proiznes rech…
  dictator who-Nom hated dissident - Acc gave speech…

  Verb
  ‘The dictator who hated the dissident gave a speech…’

  Obj
Subject RCs

- Unscrambled (Verb+RelPro \textit{local}):

\[
\begin{array}{l}
\text{Subj} & \text{Verb} & \text{Obj} \\
\text{Diktator kotoryj nena\v{d}idel dissidenta proiznes rech...} \\
\text{dictator who-Nom hated dissident-Acc gave speech...} \\
\text{‘The dictator who hated the dissident gave a speech...’}
\end{array}
\]

- Scrambled (Verb+RelPro \textit{non-local}):

\[
\begin{array}{l}
\text{Subj} & \text{Obj} & \text{Verb} \\
\text{Diktator kotoryj dissidenta nena\v{d}idel proiznes rech...} \\
\text{dictator who-Nom dissident-Acc hated gave speech...} \\
\text{‘The dictator who hated the dissident gave a speech...’}
\end{array}
\]
Object RCs

- Unscrambled (Verb+RelPro *non-local)*:

  Diktator  kotorogo  dissident  nenavidel  proiznes rech...
dictator  who-Acc  dissident-Nom  hated  gave speech...

  ‘The dictator who the dissident hated gave a speech...’
Object RCs

- Unscrambled (Verb+RelPro non-local):
  
  Obj  Subj  Verb
  Diktator  kotorogo  dissident  nenavidel  proiznes  rech...
dictator  who-Acc  dissident-Nom  hated  gave  speech...
  ‘The dictator who the dissident hated gave a speech...’

- Scrambled (Verb+RelPro local):
  
  Obj  Verb  Subj
  Diktator  kotorogo  nenavidel  dissident  proiznes  rech...
dictator  who-Acc  hated  dissident-Nom  gave  speech...
  ‘The dictator who the dissident hated gave a speech...’
<table>
<thead>
<tr>
<th></th>
<th>“Non-scrambled”</th>
<th>“Scrambled”</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRC</td>
<td>Dictator, RelPro[subj] V O</td>
<td>Dictator, RelPro[subj] O V</td>
</tr>
<tr>
<td>ORC</td>
<td>Dictator, RelPro[subj] S V</td>
<td>Dictator, RelPro[subj] V S</td>
</tr>
</tbody>
</table>
## Summary of conditions

<table>
<thead>
<tr>
<th></th>
<th>“Non-scrambled”</th>
<th>“Scrambled”</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRC</td>
<td>Dictator, RelPro[subj] V O</td>
<td>Dictator, RelPro[subj] O V</td>
</tr>
<tr>
<td>ORO</td>
<td>Dictator, RelPro[obj] S V</td>
<td>Dictator, RelPro[obj] V S</td>
</tr>
</tbody>
</table>

*local conditions*
## Summary of conditions

<table>
<thead>
<tr>
<th>SRC</th>
<th>“Non-scrambled”</th>
<th>“Scrambled”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dictator, RelPro[subj] V O</td>
<td>Dictator, RelPro[subj] O V</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ORC</th>
<th>“Non-scrambled”</th>
<th>“Scrambled”</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dictator, RelPro[obj] S V</td>
<td>Dictator, RelPro[obj] V S</td>
<td></td>
</tr>
</tbody>
</table>

**local conditions**

- For word-order & surprisal: small corpus study shows
  - $\text{Freq}(\text{SRC}) > \text{Freq}(\text{ORC})$  $(p<<0.01)$
  - $\text{Freq}(\text{SVO}) > \text{Freq}(\text{SOV})$  $(p<<0.01)$
  - $\text{Freq}(\text{OSV}) > \text{Freq}(\text{OVS})$  $(p<0.05, \text{ but } n=22)$
Locus of RT differences: embedded verb

- Significant main effect of extraction type (SRC vs. ORC; $p<0.05$)
- An interaction between scrambling and extraction type
- Looked at another way: a big main effect of locality
Locus of RT differences: embedded verb

- Significant main effect of extraction type (SRC vs. ORC; \( p<0.05 \))
- An interaction between scrambling and extraction type
- Looked at another way: a big main effect of locality

**OSV order hardest**
## How do different theories fare?

<table>
<thead>
<tr>
<th>Expt1</th>
<th>Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCType</td>
<td>Word order</td>
</tr>
<tr>
<td>?</td>
<td>✓</td>
</tr>
<tr>
<td>?</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>?</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>✓</td>
<td>×</td>
</tr>
<tr>
<td>×</td>
<td>✓</td>
</tr>
</tbody>
</table>
Experiment 1 summary
Experiment 1 summary

- Locus of differential processing difficulty is the RC verb
Experiment 1 summary

- Locus of differential processing difficulty is the RC verb
- Effect of RCType consistent with other langs
  - many theories get it right
Experiment 1 summary

- Locus of differential processing difficulty is the RC verb
- Effect of RCType consistent with other langs
  - many theories get it right
- “Scrambling” result problematic for word-order theories
  - OVS easier but less frequent than OSV
  - Proximity of verb to relative pronoun predicts difficulty
  - Results consistent with integration cost, time-based decay, entropy reduction, and SBI
Experiment 1 summary

• Locus of differential processing difficulty is the RC verb
• Effect of RCType consistent with other langs
  • many theories get it right
• “Scrambling” result problematic for word-order theories
  • OVS easier but less frequent than OSV
  • Proximity of verb to relative pronoun predicts difficulty
  • Results consistent with integration cost, time-based decay, entropy reduction, and SBI
• Some limitations of this experiment:
  • NPs were single words; spillover a possible problem
  • Proximity manipulation between RelPro and verb not parametric
Experiment 2: subject RCs

- We can further distinguish among these theories by testing exclusively subject RCs.
- Russian’s free word order allows us to parametrically vary the amount of intervening material while keeping extraction type constant.
- 0, 1, or 2 interveners between RelPro and verb.
- Interveners can be arguments or adjuncts.
- Region-by-region self-paced reading (DMDX).
- 20 items, 40 participants.
- Administered to native Russian speakers in Russia (Kazan, Moscow, St. Petersburg).
The chef was praised for his mastery early in the evening near six o’clock, but the waiter, who forgot to bring the dish of veal to the customer in the black suit on time, didn’t receive a tip after dinner.

`forgot to bring`  
the dish of veal  
to the customer in the black suit  
early in the evening  
near six o’clock  
Verb complex  
NP.Acc  
NP.Dat  
Tmp1  
Tmp2
“The chef was praised for his mastery early in the evening near six o’clock, but the waiter, who forgot to bring the dish of veal to the customer in the black suit on time, didn’t receive a tip after dinner.”
“The chef was praised for his mastery early in the evening near six o’clock, but the waiter, who forgot to bring the dish of veal to the customer in the black suit on time, didn’t receive a tip after dinner.”
Experiment 2: argument manipulation

- Argument manipulation:
  
  # interveners
  
  0  ...who forgot to bring dish of veal customer in black suit...
Experiment 2: *argument* manipulation

- *Argument* manipulation:

  # interveners

  0  ...who **forgot** *to* _bring_ dish of veal customer in black suit...

  1  ...who dish of veal **forgot** *to* _bring_ customer in black suit...
Experiment 2: *argument* manipulation

- *Argument* manipulation:
  - # interveners
    - 0 ...who *forgot* to _bring_ dish of veal *customer in black suit*...
    - 1 ...who dish of veal *forgot* to _bring_ customer in black suit...
    - 2 ...who dish of veal *customer in black suit forgot* to _bring_...
Experiment 2: *adjunct* manipulation

- We control for sentence position by ensuring the temporal adjuncts occur *somewhere* before the verb

\[ \text{the chef was praised ... , but the waiter, who } \textbf{forgot to bring} \]
Experiment 2: *adjunct* manipulation

- We control for sentence position by ensuring the temporal adjuncts occur *somewhere* before the verb.

```
the chef was praised ..., but the waiter, who   forgot to bring
```

# interveners
0 ...early in the evening near six o’clock...who   forgot to bring...
Experiment 2: *adjunct* manipulation

- We control for sentence position by ensuring the temporal adjuncts occur *somewhere* before the verb.

\[
\text{the chef was praised} \ldots, \text{ but the waiter, who } \overset{\text{Adjunct}}{\text{forgot to bring}}
\]

# interveners
0  ...early in the evening near six o’clock...who forgot to bring...
1  ...near six o’clock...who early in the evening forgot to bring...
Experiment 2: *adjunct* manipulation

- We control for sentence position by ensuring the temporal adjuncts occur *somewhere* before the verb.

\[
\text{the chef was praised . . . , but the waiter, who } \overset{\text{Adjunct}}{\overset{\text{forgot to bring}}} \]

# interveners
0 ...early in the evening near six o’clock...who forgot to bring...
1 ...near six o’clock...who early in the evening forgot to bring...
2 ......who early in the evening near six o’clock forgot to bring...
Experiment 2: corpus study

- Are syntactic expectations really in opposition to locality?
Experiment 2: corpus study

- Are syntactic expectations really in opposition to locality?
- Small study of *kotoryj* (‘who’) using the Uppsala corpus confirms:

<table>
<thead>
<tr>
<th>Conditional Probability</th>
<th>Support</th>
<th>Event prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P( V \mid kotoryj)$</td>
<td>216</td>
<td>0.50</td>
</tr>
<tr>
<td>$P( V \mid kotoryj, NP)$</td>
<td>13</td>
<td>0.85</td>
</tr>
<tr>
<td>$P( V \mid kotoryj, TMP)$</td>
<td>19</td>
<td>0.79</td>
</tr>
</tbody>
</table>
Experiment 2: corpus study

- Are syntactic expectations really in opposition to locality?
- Small study of *kotoryj* (‘who’) using the Uppsala corpus confirms:

<table>
<thead>
<tr>
<th>Conditional Probability</th>
<th>Support</th>
<th>Event prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P( V \mid kotoryj)$</td>
<td>216</td>
<td>0.50</td>
</tr>
<tr>
<td>$P( V \mid kotoryj, NP)$</td>
<td>13</td>
<td>0.85</td>
</tr>
<tr>
<td>$P( V \mid kotoryj, TMP)$</td>
<td>19</td>
<td>0.79</td>
</tr>
</tbody>
</table>

*Intervening constituent strengthens expectation for upcoming verb*
Experiment 2 results

RT at finite verb

- Zero
- One
- Two

RT at non-finite (main) verb

- Zero
- One
- Two

- Arguments
- Adjuncts

# Interveners
Experiment 2 results

RT at finite verb

- Arguments
- Adjuncts

RT at non-finite (main) verb

- Arguments
- Adjuncts

Verbal RTs increase as locality decreases (all 1×3 ANOVA p’s < 0.05)
Experiment 2 results

**RT at finite verb**

- **Arguments**
- **Adjuncts**

This contrast is bad for ERH

**RT at non-finite (main) verb**

- **Arguments**
- **Adjuncts**

Verbal RTs increase as locality decreases (all 1x3 ANOVA p’s < 0.05)
Discussion

- RTs in embedded verb complex increase monotonically with # of intervening constituents
Discussion

- RTs in embedded verb complex increase monotonically with # of intervening constituents
- Intervening arguments and adjuncts have similar effect of slowing RTs at verb
Discussion

- RTs in embedded verb complex increase monotonically with # of intervening constituents
- Intervening arguments and adjuncts have similar effect of slowing RTs at verb
- A fairly clearly locality-based pattern
Discussion

- RTs in embedded verb complex increase monotonically with # of intervening constituents
- Intervening arguments and adjuncts have similar effect of slowing RTs at verb
- A fairly clearly locality-based pattern
  - Not expectations or ERH because interveners should increase certainty; 1v2 intervener results bad for ERH
Discussion

• RTs in embedded verb complex increase monotonically with # of intervening constituents
• Intervening arguments and adjuncts have similar effect of slowing RTs at verb
• A fairly clearly locality-based pattern
  • Not expectations or ERH because interveners should increase certainty; 1v2 intervener results bad for ERH
  • Not structural asymmetry because all conds are SRC
Discussion

- RTs in embedded verb complex increase monotonically with # of intervening constituents
- Intervening arguments and adjuncts have similar effect of slowing RTs at verb
- A fairly clearly locality-based pattern
  - Not expectations or ERH because interveners should increase certainty; 1v2 interveners results bad for ERH
  - Not structural asymmetry because all conds are SRC
  - Not likely to be SBI because
    - arguments are case-marked (c.f. Lewis & Nakayama 2002)
    - object NP is inanimate; subject NP is animate
    - arguments & adjuncts affect verbal RTs similarly
How do different theories fare?

<table>
<thead>
<tr>
<th>Expt1</th>
<th>Expt2</th>
<th>Theory</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCType</td>
<td>WO</td>
<td></td>
</tr>
<tr>
<td>?</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>?</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>?</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>✓</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>✗</td>
<td>✓</td>
<td>✗</td>
</tr>
</tbody>
</table>

Note: ✗ doesn’t mean that a theory is wrong, only that *it can’t explain the results found here*
General Discussion

- Processing verbs in Russian relative clauses involves a locality-based or decay-based retrieval component.
- This component cannot be explained by many other theories advanced to explain SRC/ORC asymmetry.
- Of course, not all verbs in all languages exhibit this pattern of processing difficulty!
  - clause-final verbs in German (Konieczny 2000; Levy & Keller 2007), Japanese (Nakatani & Gibson 2005), Hindi (Vasishth & Lewis 2006)
  - unrelativized verbs in English (Jaeger et al. 2005)
  - these examples show an *expectation-based* pattern.
General Discussion & Prospectus

- So what determines the general pattern of processing difficulty for a verb?
  - Not morphology (contrast German & Russian)
  - Not just SBI
  - Relativization? (but see Vasishth & Lewis 2006)
  - Facts about (dominant) word order in the language?
- More cross-linguistic work in online processing of verbs still needed
- Helps us answer big-picture questions about the role and prominence of different cognitive constraints
Acknowledgments

- UK Economic & Social Research Council
- Frank Keller
- Natalia Roudakova
- Irina Ryabkova
- Natalia Kondratyeva

Thank you!

http://ling.ucsd.edu/~rlevy
http://web.mit.edu/evelina9/www/
http://web.mit.edu/bcs/people/gibson.shtml
Main-verb effects (experiment 2)

- Not much seems to be going on at the matrix verb
- The 2-intervener argument condition seems a bit slower, but this is not significant
Spillover from DO and IO?

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>RelPro Verb Acc Dat</td>
<td>RelPro Acc Verb Dat</td>
<td>RelPro Acc Dat Verb</td>
<td></td>
</tr>
</tbody>
</table>

- For the accusative NP, it seems like spillover could explain the RT differentials at the verb complex…
Spillover from DO and IO? (2)

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>RelPro Verb Acc Dat</td>
<td>RelPro Acc Verb Dat</td>
<td>RelPro Acc Dat Verb</td>
</tr>
</tbody>
</table>

- ...but there seems to be no such sign for spillover from the dative NP
- **CAVEAT**: the dative and accusative NPs sometimes cross lines (we’re re-running to fix this)