A processing model for ungrammatical VP ellipsis

Gregory M. Kobele

Christina S. Kim

John T. Hale

Jeffrey T. Runner
Peter invited Dan and Lyn did, too.
VP Ellipsis

Peter invited Dan and Lyn did, too.

\[ \text{false in worlds where Lyn is not a Dan-inviter} \]
VP Ellipsis

Peter \( [_{VP} \text{invited Dan}] \), and Lyn did \( [_{VP} \text{invite Dan}] \), too
VP Ellipsis

Peter \([VP \text{ invited Dan}]\), and Lyn did \([VP \text{ invite Dan}]\), too.

Peter invited Dan, and Lyn did \(\triangle\), too.
VP Ellipsis is resolved...
VP Ellipsis is resolved...

At a SYNTACTIC level

VP Ellipsis is resolved...

At a **SYNTACTIC level**

At a **SEMANTIC level**
(Dalrymple et al 1991; Hardt 1993)
VP Ellipsis is resolved...

At a **SYNTACTIC level**

At a **SEMANTIC level**
(Dalrymple et al 1991; Hardt 1993)

It depends on the **DISCOURSE RELATION**
Resemblance=**SYNTACTIC**, Cause/Effect=**SEMANTIC**
(Kehler 2000, 2002)
Tanenhaus & Carlson 90 supports SYNTACTIC accounts

✔ Someone had to **take out** the garbage. But Bill refused to **do it**.

✔ The garbage had to **be taken out**. But Bill refused to **do it**.

`Do it`
Tanenhaus & Carlson 90 supports syntactic accounts

✔ Someone had to take out the garbage. But Bill refused to do it.

✔ The garbage had to be taken out. But Bill refused to do it.

✔ Someone had to take out the garbage. But Bill refused to \(\Delta\).

✘ The garbage had to be taken out. But Bill refused to \(\Delta\).
Perhaps voice mismatches undermine syntactic accounts

This information could have been released by Gorbachov, but he chose not to.

*Daniel Shorr, NPR, 10/17/92, from Hardt 1993*

In March, four fireworks manufacturers asked that the decision be reversed, and on Monday the ICC did.

*from Rosenthal 1988, cited in Dalrymple 1991*

I am America (and so can you!)

*Steven Colbert 2007*
Syntactic Approaches

Arregui et al 2006 JML

> VPE licensed under syntactic identity
> mismatches are ungrammatical
> reanalysis rules “recycle” bits and pieces of mismatched antecedents
Syntactic Approaches

Arregui et al 2006 JML

- VPE licensed under syntactic identity
- mismatches are ungrammatical
- reanalysis rules “recycle” bits and pieces of mismatched antecedents

Kobele/Kim/Hale/Runner

- VPE licensed under syntactic identity
  - mismatches are *grammatical*
  - parsing preferences render mismatches *harder to analyze*
The problem with VPE that needs solving

• Our proposal in a nutshell
• Experiment #1: Voice and Category mismatches in VPE
• Experiment #2: Discourse Coherence modulates syntactic effect
• A parsing algorithm
Varieties of dependency

movement

hypothetical reasoning

Which letter did you send to John?

Varieties of dependency
Parsing preferences

- Disprefer antecedents with hypotheses in them

- Expect “large” antecedents
• The problem with VPE that needs solving

• Our proposal in a nutshell

• Experiment #1: Voice and Category mismatches in VPE

• Experiment #2: Discourse Coherence modulates syntactic effect

• A parsing algorithm
### Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ellipsis</td>
<td>Voice</td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ellipsis</td>
<td>Voice</td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Ellipsis-Mismatch

*Ann was invited by Peter, and Sandy did, too.*

*The criticism of Roy was harsh, but Kate didn’t.*
### Experiment 1

<table>
<thead>
<tr>
<th>Ellipsis-Mismatch</th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ellipsis</td>
<td>Voice</td>
<td></td>
</tr>
<tr>
<td>Ellipsis</td>
<td>Voice</td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ellipsis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ellipsis-Mismatch**  
Ann was invited by Peter, and Sandy did, too.  
*The criticism of Roy was harsh, but Kate didn’t.*

**Ellipsis-Match**  
Ann was invited by Peter, and Sandy was, too.  
*The article criticized Roy, but Kate didn’t.*
## Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No ellipsis</strong></td>
<td>Voice</td>
<td></td>
</tr>
<tr>
<td><strong>Category</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Ellipsis</strong></td>
<td>Voice</td>
<td></td>
</tr>
<tr>
<td><strong>Category</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ellipsis-Mismatch**  
*Ann was invited by Peter, and Sandy did, too.*  
*The criticism of Roy was harsh, but Kate didn’t.*

**Ellipsis-Match**  
*Ann was invited by Peter, and Sandy was, too.*  
*The article criticized Roy, but Kate didn’t.*

**NoEllipsis-Mismatch**  
*Ann was invited by Peter, and Sandy invited her, too.*  
*The criticism of Roy was harsh, but Kate didn’t criticize him.*
## Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ellipsis</td>
<td>Voice</td>
<td></td>
</tr>
<tr>
<td>Cat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ellipsis</td>
<td>Voice</td>
<td></td>
</tr>
<tr>
<td>Cat</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Ellipsis-Mismatch**

Ann was invited by Peter, and Sandy did, too.  
*The criticism of Roy was harsh, but Kate didn’t.*

**Ellipsis-Match**

Ann was invited by Peter, and Sandy was, too.  
*The article criticized Roy, but Kate didn’t.*

**NoEllipsis-Mismatch**

Ann was invited by Peter, and Sandy invited her, too.  
*The criticism of Roy was harsh, but Kate didn’t criticize him.*

**NoEllipsis-Match**

Ann was invited by Peter, and Sandy was invited by him, too.  
*The article criticized Roy, but Kate didn’t criticize him.*
Experiment 1

MAGNITUDE ESTIMATION OF LINGUISTIC ACCEPTABILITY

Ellen Gurman Bard, Dan Robertson, Antonella Sorace

University of Edinburgh

Judgments of linguistic acceptability constitute an important source of evidence for theoretical and applied linguistics, but are typically elicited and represented in ways which limit their utility. This paper describes how magnitude estimation, a technique used in psychophysics, can be adapted for eliciting acceptability judgments. Magnitude estimation of linguistic acceptability is shown to solve the measurement scale problems which plague conventional techniques; to provide data which make fine distinctions robustly enough to yield statistically significant results of linguistic interest; to be usable in a consistent way by linguistically naive speaker-hearers, and to allow replication across groups of subjects. Methodological pitfalls are discussed and suggestions are offered for new approaches to the analysis and measurement of linguistic acceptability.*

Language, Vol. 72, No. 1. (Mar., 1996), pp. 32-68
Experiment 1 procedure

- On each trial, participants assign a number to a sentence that indicates its acceptability as compared to a modulus sentence they rated at the beginning of the experiment.

1) Practice line length estimation

2) Practice estimating sentence acceptability

3) 48 experimental trials
Give this sentence a number. Then give each sentence after it a number based on how good it sounds compared to the first sentence.

The kids were amused by the cartoon, but their parents weren't

Press <SPACE> to continue
The kids were amused by the cartoon, but their parents weren't

Sam recommended the New York Times bestseller to Meg, and Todd recommended it to her, too
The kids were amused by the cartoon, but their parents weren't.

The essay was copied by some students, but Jim didn't.
Experiment 1 results

Mismatch less acceptable than Match
Ellipsis x Mismatch interaction:
Mismatch is worse than match, but only in Ellipsis
### Experiment 1: Voice mismatch in Ellipsis

<table>
<thead>
<tr>
<th>Voice Mismatch</th>
<th>Mean Log Acceptability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active-Active</td>
<td>-0.002</td>
</tr>
<tr>
<td>Passive-Passive</td>
<td>-0.285</td>
</tr>
<tr>
<td>Passive-Active</td>
<td>-0.616</td>
</tr>
<tr>
<td>Active-Passive</td>
<td>-0.697</td>
</tr>
</tbody>
</table>

**Ranking:** $AA > PP > \{PA, AP\}$
Peter invited Dan, and Lyn did \( \Delta \), too.
Peter invited Dan, and Lyn did $\Delta$, too.
Peter invited Dan, and Lyn did △, too.
Peter invited Dan, and Lyn did △, too.
Peter invited Dan, and Lyn did △, too.
Josh was invited and Dan was $\triangle$, too
Passive-Passive VPE uses hypotheses

Josh was invited and Dan was $\triangle$, too
Peter invited Josh, and Matt was △, too.
## Interim summary

<table>
<thead>
<tr>
<th>condition</th>
<th>observed mean log acceptability</th>
<th>hypothesis in antecedent?</th>
<th>antecedent derivation size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active-Active</td>
<td>-0.002</td>
<td>¬</td>
<td>8</td>
</tr>
<tr>
<td>Passive-Passive</td>
<td>-0.285</td>
<td>✓</td>
<td>7</td>
</tr>
<tr>
<td>Passive-Active</td>
<td>-0.616</td>
<td>✓</td>
<td>5</td>
</tr>
<tr>
<td>Active-Passive</td>
<td>-0.697</td>
<td>✓</td>
<td>5</td>
</tr>
</tbody>
</table>
### Experiment 1: Category mismatch in Ellipsis

<table>
<thead>
<tr>
<th></th>
<th>mean log acceptability</th>
</tr>
</thead>
<tbody>
<tr>
<td>V-V</td>
<td>0.486</td>
</tr>
<tr>
<td>N-V</td>
<td>-0.690</td>
</tr>
<tr>
<td>Adj-V</td>
<td>-0.981</td>
</tr>
</tbody>
</table>

**V-V > N-V > Adj-V**
Experiment 1: Category mismatch in Ellipsis

<table>
<thead>
<tr>
<th>Syntax</th>
<th>Mean Log Acceptability</th>
</tr>
</thead>
<tbody>
<tr>
<td>V-V</td>
<td>0.486</td>
</tr>
<tr>
<td>N-V</td>
<td>-0.690</td>
</tr>
<tr>
<td>Adj-V</td>
<td>-0.981</td>
</tr>
</tbody>
</table>

V-V > N-V > Adj-V

same syntax as Active-Active
The criticism of Roy was harsh but Matt didn’t $\Delta$. 

N-V mismatching

VP ellipsis

no hypotheses in antecedent

ellipsis at rootP level, 3 nodes

The criticism of Roy was harsh but Matt didn’t $\Delta$. 
Adjectivization

dP(0)
   \[\text{the}\]
      \[\text{mistake}\]

\[
t' \quad \text{t}
   \quad \text{d'}
   \quad \text{d}
   \quad \text{nP}
   \quad \text{perf}
   \quad \text{perf'}
   \quad \text{progP}
   \quad \text{prog'}
   \quad \text{adjP}
   \quad \text{degP}
   \quad \text{deg'}
   \quad \text{deg}
   \quad \text{adj}
   \quad \text{rootP}
   \quad \text{t}
   \quad \text{t(0)}
\]
The mistake was hardly excusable but the director did not.
## Interim summary (reprise)

<table>
<thead>
<tr>
<th>condition</th>
<th>observed mean log acceptability</th>
<th>hypothesis in antecedent?</th>
<th>antecedent derivation size</th>
<th>predicted m.l.a.</th>
</tr>
</thead>
<tbody>
<tr>
<td>V-V</td>
<td>0.047</td>
<td>¬</td>
<td>8</td>
<td>1.449</td>
</tr>
<tr>
<td>Active-Active</td>
<td>-0.002</td>
<td>¬</td>
<td>8</td>
<td>1.449</td>
</tr>
<tr>
<td>Passive-Passive</td>
<td>-0.285</td>
<td>✓</td>
<td>7</td>
<td>0.939</td>
</tr>
<tr>
<td>Passive-Active</td>
<td>-0.616</td>
<td>✓</td>
<td>5</td>
<td>0.577</td>
</tr>
<tr>
<td>N-V</td>
<td>-0.690</td>
<td>¬</td>
<td>3</td>
<td>0.543</td>
</tr>
<tr>
<td>Active-Passive</td>
<td>-0.697</td>
<td>✓</td>
<td>5</td>
<td>0.577</td>
</tr>
<tr>
<td>Adj-V</td>
<td>-0.981</td>
<td>✓</td>
<td>3</td>
<td>0.215</td>
</tr>
</tbody>
</table>

hypothesis demerit=0.3285
derivation-size credit=0.1811
The problem with VPE that needs solving

Our proposal in a nutshell

Experiment #1: Voice and Category mismatches in VPE

Experiment #2: Discourse Coherence modulates syntactic effect

A parsing algorithm
Discourse Coherence

Resemblance

Ellipsis resolution depends on syntactic alignment
→ presence of syntactic effects in VPE

Cause/Effect

Ellipsis only has to recover semantic content
→ absence of syntactic effects in VPE

Kehler 2000, 2002
## Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ellipsis</td>
<td>Resemblance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cause-Effect</td>
<td></td>
</tr>
<tr>
<td>Ellipsis</td>
<td>Resemblance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cause-Effect</td>
<td></td>
</tr>
</tbody>
</table>

**Ellipsis-Mismatch**

*Abby was invited by Bill, and Matt did, too.*
# Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ellipsis</td>
<td>Resemblance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cause-Effect</td>
<td></td>
</tr>
<tr>
<td>Ellipsis</td>
<td>Resemblance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cause-Effect</td>
<td></td>
</tr>
</tbody>
</table>

**Ellipsis-Mismatch**

Abby was invited by Bill, and Matt did, too.

Abby was invited by Bill, so Matt did, too.
## Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ellipsis</td>
<td>Resemblance</td>
<td>Cause-Effect</td>
</tr>
<tr>
<td>Ellipsis</td>
<td>Resemblance</td>
<td>Cause-Effect</td>
</tr>
</tbody>
</table>

### Ellipsis-Mismatch

- Abby was invited by Bill, and Matt did, too.
- Abby was invited by Bill, so Matt did, too.

### Ellipsis-Match

- Abby was invited by Bill, and Matt was, too.
- Abby was invited by Bill, so Matt was, too.
**Experiment 2**

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>No ellipsis</td>
<td>Resemblance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cause-Effect</td>
<td></td>
</tr>
<tr>
<td>Ellipsis</td>
<td>Resemblance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cause-Effect</td>
<td></td>
</tr>
</tbody>
</table>

**Ellipsis-Mismatch**

- Abby was invited by Bill, and Matt did, too.
- Abby was invited by Bill, so Matt did, too.

**Ellipsis-Match**

- Abby was invited by Bill, and Matt was, too.
- Abby was invited by Bill, so Matt was, too.

**NoEllipsis-Mismatch**

- Abby was invited by Bill, and Matt invited her, too.
- Abby was invited by Bill, so Matt invited her, too.
# Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No ellipsis</strong></td>
<td>Resemblance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cause-Effect</td>
<td></td>
</tr>
<tr>
<td><strong>Ellipsis</strong></td>
<td>Resemblance</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cause-Effect</td>
<td></td>
</tr>
</tbody>
</table>

**Ellipsis-Mismatch**

- Abby was invited by Bill, and Matt did, too.
- Abby was invited by Bill, so Matt did, too.

**Ellipsis-Match**

- Abby was invited by Bill, and Matt was, too.
- Abby was invited by Bill, so Matt was, too.

**NoEllipsis-Mismatch**

- Abby was invited by Bill, and Matt invited her, too.
- Abby was invited by Bill, so Matt invited her, too.

**NoEllipsis-Match**

- Abby was invited by Bill, and Matt was invited by her, too.
- Abby was invited by Bill, so Matt was invited by her, too.
Experiment 2 results

Mismatch effect in both Discourse Relations
Experiment 2 results

Ellipsis x Mismatch x DR interaction: Stronger in Resemblance than Cause/Effect
Hypothetical antecedents dispreferred less in Cause-Effect

<table>
<thead>
<tr>
<th>condition</th>
<th>prediction demeriting on</th>
<th>prediction demeriting off</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active-Active</td>
<td>1.449</td>
<td>1.449</td>
</tr>
<tr>
<td>Passive-Passive</td>
<td>0.939</td>
<td>1.268</td>
</tr>
<tr>
<td>Passive-Active</td>
<td>0.577</td>
<td>0.906</td>
</tr>
<tr>
<td>Active-Passive</td>
<td>0.577</td>
<td>0.906</td>
</tr>
</tbody>
</table>

hypothesis demerit=0.0
derivation-size credit=0.181
• The problem with VPE that needs solving
• Our proposal in a nutshell
• Experiment #1: Voice and Category mismatches in VPE
• Experiment #2: Discourse Coherence modulates syntactic effect
• A parsing algorithm
Parsing algorithm

1. Find likely ellipsis site $\triangle$ using surface cues
2. Search for $\triangle$’s antecedent among all previously-found constituents that meet a condition $C$.

$C_{LappinMcCord90} =$ antecedent does not itself contain $\triangle$

$C_{NotCrossing} =$ such that antecedent/$\triangle$ dependencies are nested

$C_{Recoverability} =$ such that antecedent/$\triangle$ dependencies are not cyclical
Conclusion

Syntactic identity can derive mismatching VPE by composing word subparts in the derivation.

Two parser search heuristics combine to make the less-acceptable items *harder to analyze*. 
Bonus Slides
Experiment 1 results

Main effects:  Ellipsis \( (F(1,19)=51.0, p<.0001) \)
Mismatch \( (F(1,19)=99.2, p<.0001) \)
Mismatch Type \( (F(1,19)=61.7, p<0.0001) \)
Experiment 1 results

Interactions:
- Ellipsis x Mismatch $(F(1,19)=61.7, \ p<0.0001)$
- Ellipsis x Mismatch Type $(F(1,19)=5.0, \ p<.05)$
- Mismatch x Mismatch Type $(F(1,19)=33.8, \ p<.0001)$
- Ellipsis x Mismatch x Mismatch Type $(F(1,19)=16.2, \ p<.001)$
Experiment 1 results

Interaction of interest:

Ellipsis x Mismatch interaction \((F(1,19)=61.7, p<0.0001)\)

→ Mismatch is worse than match, but only in Ellipsis
### Experiment 1: Voice mismatch in Ellipsis

<table>
<thead>
<tr>
<th></th>
<th>mean log acceptability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active-Active</td>
<td>-0.002</td>
</tr>
<tr>
<td>Passive-Passive</td>
<td>-0.285</td>
</tr>
<tr>
<td>Passive-Active</td>
<td>-0.616</td>
</tr>
<tr>
<td>Active-Passive</td>
<td>-0.697</td>
</tr>
</tbody>
</table>

**AA > PP > {PA,AP}**

corrected $p < 0.0001$ for all contrasts except PA-AP, $p = .29$
Experiment 1: Category mismatch in Ellipsis

<table>
<thead>
<tr>
<th></th>
<th>mean log acceptability</th>
</tr>
</thead>
<tbody>
<tr>
<td>V-V</td>
<td>0.486</td>
</tr>
<tr>
<td>N-V</td>
<td>-0.690</td>
</tr>
<tr>
<td>Adj-V</td>
<td>-0.981</td>
</tr>
</tbody>
</table>

V-V > N-V > Adj-V

all corrected p ≤ 0.01
Main effects: 

Ellipsis \((F(1,23) = 119.9, \ p < .0001)\)

Mismatch \((F(1,23) = 73.0, \ p < .0001)\)

Discourse Relation Type \((F(1,23) = 22.7, \ p < .0001)\)
Interactions: 

Ellipsis x Mismatch (F(1,23)=102.9, p<.0001)
Mismatch x Discourse Relation Type (F(1,23)=4.9, p<.03)
Ellipsis x Mismatch x DR Type (F(1,23)=4.0, p<.05)
**Experiment 2 results**

Interaction of interest:

Ellipsis x Mismatch x DR type \( (F(1, 23) = 4.0, p < .05) \)

→ Stronger Ellipsis-Mismatch interaction in Resemblance than Cause-Effect
THE GRAMMAR OF ENGLISH NOMINALIZATIONS

By Robert B. Lees
Both among the rules to derive new sentences and among the morphophonemic mappings it will be necessary to include many rules which transcend in power the limitations which were imposed upon the phrase-structure rules. These will no longer be constrained to the conversion of only a single symbol at a time, nor will it be possible to specify the domain of a rule of this kind merely by the shape of the strings upon which it is to operate, for often the applicability of such a rule will depend upon the phrase-structure history of a string as given by its tree of derivation (and perhaps also any other preceding rules which may have been applied). This is to say that, without some special provisions, these rules will no longer automatically preserve a recoverable constituent-analysis for strings, and yet these rules require for their very application just such a derivation history.

As an example of how this comes about we might note that we have already permitted the constituent-structure grammar to generate at least one string with two different derivational histories, and at the same time we desire that this string be permitted to undergo a certain nominalization transformation only in the case where it has resulted in one of the two convergent ways.
Passive precludes simple identity

Josh was invited and Dan was, \(\Delta\) too.