When data collide:
Traditional judgments vs. formal experiments in sentence acceptability

Grant Goodall
UC San Diego
Two areas of concern in syntax
1. Traditional judgments + formal experiments

• What does it mean when they conflict?
1. Traditional judgments + formal experiments
   • What does it mean when they conflict?

2. Constraints on wh-movement
   • Are some still completely beyond our understanding?
1. Traditional judgments + formal experiments

- What does it mean when they conflict?
Traditional judgments
Traditional judgments  Formal experiments
Traditional judgments

Formal experiments

Generally align very closely
(Cowart 1997, Myers 2009, Sprouse et al. 2013)
Traditional judgments generally align very closely with formal experiments (Cowart 1997, Myers 2009, Sprouse et al. 2013). Occasional discrepancies have been noted (Gibson and Fedorenko 2013).
Traditional judgments: Formally experiments

- Generally align very closely (Cowart 1997, Myers 2009, Sprouse et al. 2013)
- Occasional discrepancies (Gibson and Fedorenko 2013)
- Discussion often misguided:
Traditional judgments generally align very closely with formal experiments (Cowart 1997, Myers 2009, Sprouse et al. 2013). Occasional discrepancies have been noted (Gibson and Fedorenko 2013), but discussion is often misguided: Which method is right?
Traditional judgments

Formal experiments

Generally align very closely
(Cowart 1997, Myers 2009, Sprouse et al. 2013)

Occasional discrepancies
(Gibson and Fedorenko 2013)

Discussion often misguided:

Which method is right?
Traditional judgments

Generally align very closely
(Cowart 1997, Myers 2009, Sprouse et al. 2013)

Occasional discrepancies
(Gibson and Fedorenko 2013)

Discussion often misguided:

Which method is right?  What is each method telling us?
Traditional judgments

Formal experiments

Generally align very closely
(Cowart 1997, Myers 2009, Sprouse et al. 2013)

Occasional discrepancies
(Gibson and Fedorenko 2013)

Discussion often misguided:

Which method is right?

What is each method telling us?

What is the right tool for the job?
One difference we know about

• Formal acceptability experiments are very sensitive to demands on working memory.
• Affects long-distance extraction, for example.
• Sometimes larger than well-known grammatical effects.
Working memory vs. grammar

From Keffala (2011), Keffala & Goodall (2011)
Working memory vs. grammar

From Keffala (2011), Keffala & Goodall (2011)
Working memory vs. grammar

Effect of Embedding

From Keffala (2011), Keffala & Goodall (2011)
Working memory vs. grammar

From Keffala (2011), Keffala & Goodall (2011)
Working memory vs. grammar

- Subject Gap
- Subject Resumptive
- Object Gap
- Object Resumptive

Effect of Embedding

That-trace effect

From Keffala (2011), Keffala & Goodall (2011)
• Are all differences in results attributable to known differences like this?
• What are we to make of differences that aren’t so easily explained?
1. Traditional judgments + formal experiments

- What does it mean when they conflict?
What does it mean when they conflict?

1. Traditional judgments
   + formal experiments

Are some still completely beyond our understanding?

2. Constraints on *wh*-movement
   
   • Are some still completely beyond our understanding?
An embarrassment of riches

• Many \textit{wh}-phenomena have several possible explanations.
  
  CNPC: What do you believe the claim that Mary saw \____? 
  
  \textit{Wh}-island: What do you wonder when Mary saw \____?
An embarrassment of riches

- Many *wh*-phenomena have several possible explanations.
  
  CNPC: What do you believe the claim that Mary saw __?  
  *Wh*-island: What do you wonder when Mary saw __?

- Both are structurally more complex than:
  
  What do you believe that Mary saw __?
A number of possible, plausible explanations

• In terms of:
  grammar
  working memory
A number of possible, plausible explanations

- In terms of:
  - grammar
  - working memory

- Problem becomes figuring out relative role of each.
Not all cases give us this luxury

- For other *wh*-phenomena, we are lucky to come up with *any* plausible explanation.
  
  Who do you think (that) Mary saw __ ?
  
  Who do you think (*that) __ saw Mary?
Not all cases give us this luxury

• For other *wh*-phenomena, we are lucky to come up with any plausible explanation.
  Who do you think (that) Mary saw __ ?
  Who do you think (*that) __ saw Mary?

• Appears to make little sense in terms of processing.
• Also hard to find grammatical reason.
1. Traditional judgments + formal experiments

- What does it mean when they conflict?

2. Constraints on *wh*-movement

- Are some still completely beyond our understanding?
1. Traditional judgments + formal experiments

2. Constraints on wn-movement

- Are some still completely beyond our understanding?

Subextraction from SPEC of CP
Background

• Like many languages, Spanish forbids subextraction from (preverbal) subjects:

*Esta es la autora de la que
   [varias traducciones __ ] han ganado premios internacionales.
`This is the author by whom
   [several translations __ ] have won international awards.'
Subextraction out of SPEC/CP

- But Esther Torrego noticed an amazing fact: subextraction improves when subject is moved to SPEC of CP:

¿De qué autora no sabes [qué traducciones __ ] han ganado premios internacionales?
`By what author don't you know [which translations __ ] have won internat’l awards?'
(Torrego (1985), Chomsky (1986))
Also in Italian

? [Di quale autore] ti domandi [\textsubscript{CP} [quanti libri t]\textsubscript{i} [\textsubscript{TP} siano stati censurati t\textsubscript{j}]]

‘Which author do you wonder how many books by have been censored?’

(from Rizzi 2006: 114)
? Which athletes do you wonder [CP [which pictures of t₁]j Mary bought tₐ ]

? Which athletes do you wonder [CP [which pictures of t₁]j tₐ are on sale]

Lasnik & Saito (1992:111)

? [CP Who can’t you decide [CP [how many pictures of t₁]z to buy tₚ for your kids]]?

Kayne (1984:192)
Why is this amazing?
Why is this amazing?

I. Syntax
   – Freezing Principle
   – Criterial Freezing
   – Chain Uniformity
Why is this amazing?

I. Syntax
- Freezing Principle
- Criterial Freezing
- Chain Uniformity

\[ [\ldots]_k \ldots [\ldots \underline{\kappa} ]_j \ldots \underline{\ell}_j \]
Why is this amazing?

I. Syntax
   – Freezing Principle
   – Criterial Freezing
   – Chain Uniformity

\[ [\ldots]_k \ldots [\ldots \underline{\ldots}_k]_j \ldots \underline{\ldots}_j \]
Why is this amazing?

I. Syntax
   – Freezing Principle
   – Criterial Freezing
   – Chain Uniformity

\[ \ldots \]_k \ldots \[\ldots \_k \]_j \ldots \_j \]
Freezing Principle at work

You gave [a book about Tom] to your niece.
You gave __ to your niece [a book about Tom].
Freezing Principle at work

You gave [a book about Tom] to your niece.
You gave __ to your niece [a book about Tom].

Who did you give [a book about __] to your niece]?
?*Who did you give __ to your niece [a book about __]?
Criterial Freezing

An element moved to a position dedicated to some scope-discourse interpretive property, a criterial position, is frozen in place.

Rizzi (2004)

*$_{\text{CP}}$ [Which book] does Bill wonder $[_{\text{CP}} t$ [she read $t$] ] ]?
Why is this amazing?

II. Sentence processing

[Which author] don’t you know

[which translations of ___ ] won prizes?
Why is this amazing?

II. Sentence processing

[Which author] don’t you know

[which translations of ___ ] won prizes?

Requires:
• Positing filler (hard)
• Positing filler before other dependency resolved (really hard)
• Positing gap (hard)
• Doing all of this at the same time (really hard!)
So... many reasons to be amazed by subextraction from SPEC/CP

[Which author] don’t you know

[which translations of ___ ] won prizes?

We would expect it to be worse, not better, than subextraction from subject.
But does it happen?

• Some have claimed that it does not (e.g., G. Müller 2010, Gallego 2010).

• Gallego (2010) suggests *de qué autora* is argument of *sabes*:

  ¿De qué autora no sabes __
  [qué traducciones] han ganado premios internacionales?
  `By what author don't you know __
  [which translations] have won international awards?`
  – This shouldn’t be possible in English.
English is a great test case

• Because of preposition stranding, we can be more certain where the gap is.
English is a great test case

- Because of preposition stranding, we can be more certain where the gap is.
- This is not possible in Spanish.
English is a great test case

• Because of preposition stranding, we can be more certain where the gap is.
• This is not possible in Spanish.
• A formal experiment makes sense here.
  – Phenomenon is subtle at best.
  – Contrast has been called into question.
Method

- 48 participants
- 7-point scale (1 = “very bad”, 7 = “very good”)
Materials: design

• 2 x 3 x 2 design:
  – Grammatical function of affected constituent:
    • Subject vs. Object
  – Location of affected constituent:
    • SPEC/CP vs. Embedded clause vs. Matrix clause
  – Type of *wh*-movement:
    • Preposition-stranding vs. Pied-Piping
Materials: lists

• 4 tokens of each condition: Subjects see 48 experimental items
• 57 fillers (1.2 : 1 filler/experimental ratio)
• 12 lists: counterbalanced (Latin square) and pseudo-randomized
• 12 additional lists with reverse order of items
• 2 subjects randomly assigned to each list
• Screening out of outlier subjects based on fillers
<table>
<thead>
<tr>
<th>Phrase with gap</th>
<th>Location</th>
<th>Function</th>
<th>Sample stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>matrix</td>
<td><strong>Subject</strong></td>
<td>[Which animal] will [several movies about __] be shown to the visitors?</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Object</strong></td>
<td>[Which animal] will they show [several movies about __] to the visitors?</td>
<td></td>
</tr>
<tr>
<td>Phrase with gap</td>
<td>Location</td>
<td>Function</td>
<td>Sample stimuli</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
<td>----------</td>
<td>----------------</td>
</tr>
<tr>
<td></td>
<td><strong>matrix</strong></td>
<td>Subject</td>
<td>[Which animal] will [several movies about __] be shown to the visitors?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Object</td>
<td>[Which animal] will they show [several movies about __] to the visitors?</td>
</tr>
<tr>
<td></td>
<td><strong>embedded</strong></td>
<td>Subject</td>
<td>[Which animal] do you wonder whether [several movies about __] will be shown to the visitors?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Object</td>
<td>[Which animal] do you wonder whether they will show [several movies about __] to the visitors?</td>
</tr>
<tr>
<td>Phrase with gap</td>
<td>Location</td>
<td>Function</td>
<td>Sample stimuli</td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
<td>----------</td>
<td>----------------</td>
</tr>
<tr>
<td>matrix</td>
<td>Subject</td>
<td>[Which animal] will [several movies about __] be shown to the visitors?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Object</td>
<td>[Which animal] will they show [several movies about __] to the visitors?</td>
<td></td>
</tr>
<tr>
<td>embedded</td>
<td>Subject</td>
<td>[Which animal] do you wonder whether [several movies about __] will be shown to the visitors?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Object</td>
<td>[Which animal] do you wonder whether they will show [several movies about __] to the visitors?</td>
<td></td>
</tr>
<tr>
<td>embedded SPEC/CP</td>
<td>Subject</td>
<td>[Which animal] do you wonder [how many movies about __] will be shown to the visitors?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Object</td>
<td>[Which animal] do you wonder [how many movies about __] they will show to the visitors?</td>
<td></td>
</tr>
</tbody>
</table>
With preposition-stranding

Graph showing the relationship between preposition-stranding and SPEC/CP with error bars for both GapInSubject-Stranding and GapInObject-Stranding.
With preposition-stranding

![Graph showing the comparison between matrix, embedded, and SPEC/CP with preposition-stranding. The graph illustrates the relationship between the categories and their corresponding values with error bars. The categories matrix, embedded, and SPEC/CP are plotted along the x-axis, and the values range from -1 to 0.6 along the y-axis. The graph includes two lines: one for GapInSubject-Stranding (blue) and one for GapInObject-Stranding (red).]
With preposition-stranding

Subject island effect

GapInSubject-Stranding

GapInObject-Stranding
With preposition-stranding

- Subject island effect
- Whether island effect
With preposition-stranding

matrix | embedded | SPEC/CP

GapInSubject-Stranding

GapInObject-Stranding

Subject island effect

Whether island effect
With preposition-stranding

Subject island effect

Whether island effect

Subject island effect
With preposition-stranding

Whether-island effect disappears.
With preposition-stranding

Whether-island effect disappears.

SPEC/CP amelioration doesn’t materialize.
Conclusion so far: effect does not exist

- Severe degradation with subextraction from subject.
- Even worse with subextraction from SPEC/CP.
Conclusion so far: effect does not exist

• Severe degradation with subextraction from subject.
• Even worse with subextraction from SPEC/CP.

• So far: Preposition-stranding, where gap position is clear.
• What about with pied-piping?
<table>
<thead>
<tr>
<th>Phrase with gap</th>
<th>Location</th>
<th>Function</th>
<th>Sample stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>matrix</td>
<td>Subject</td>
<td>[About which animal] will [several movies __] be shown to the visitors?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Object</td>
<td>[About which animal] will they show [several movies __] to the visitors?</td>
</tr>
<tr>
<td></td>
<td>embedded</td>
<td>Subject</td>
<td>[About which animal] do you wonder whether [several movies __] will be shown to the visitors?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Object</td>
<td>[About which animal] do you wonder whether they will show [several movies __] to the visitors?</td>
</tr>
<tr>
<td></td>
<td>embedded SPEC/CP</td>
<td>Subject</td>
<td>[About which animal] do you wonder [how many movies __] will be shown to the visitors?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Object</td>
<td>[About which animal] do you wonder [how many movies __] they will show to the visitors?</td>
</tr>
</tbody>
</table>
With pied-piping

matrix    embedded    SPEC/CP

GapInSubject-PPing

GapInObject-PPing
With pied-piping

No Subject island effects.
With pied-piping

No Subject island effects.

Lack of subject island effects suggests participants are not positing gap within subject.
P-stranding and pied-piping together

matrix  embedded  SPEC/CP

GapInSubject-Stranding
GapInObject-Stranding
GapInSubject-PPing
GapInObject-PPing
P-stranding and pied-piping together

Subject island effect

Graph showing the relationship between matrix, embedded, and SPEC/CP with the y-axis representing various values from 0.6 to -1 and the x-axis representing different conditions.
P-stranding and pied-piping together

matrix  
embedding SPEC/CP

Subject island effect
SPEC/CP effect
Original subextraction claims?

- GapInSubject-Stranding
- GapInObject-Stranding
- GapInSubject-PPing
- GapInObject-PPing

Matrix | Embedded | SPEC/CP
---|---|---
-1 | -0.8 | -0.6
-0.4 | 0.2 | 0.4
Original subextraction claims?

[By which author] did [various translations __ ] win awards?
Original subextraction claims?

[By which author] did [various translations __ ] win awards?

Position of gap is clear.
Original subextraction claims?

By which author did [various translations __ ] win awards?

By which author don’t you know [how many translations __ ] won awards?

Position of gap is clear.
Original subextraction claims?

- Position of gap is unclear.
- Position of gap is clear.

Graph showing the relationship between matrix, embedded, and SPEC/CP with different gap types: GapInSubject-Stranding, GapInObject-Stranding, GapInSubject-PPing, and GapInObject-PPing.

Questions:
- [By which author] did [various translations __ ] win awards?
- [By which author] don’t you know [how many translations __ ] won awards?
Original subextraction claims?

Position of gap is unclear.

Position of gap is clear.

This could be the source of the claimed contrast.

[By which author] did [various translations __ ] win awards?

[By which author] don’t you know [how many translations __ ] won awards?
Conclusion

• One can detect a contrast, but only if one ignores a confound.
Conclusion

• One can detect a contrast, but only if one ignores a confound.

• Original claim:
  gap in subject  <  gap in SPEC/CP
Conclusion

• One can detect a contrast, but only if one ignores a confound.

• Original claim:
  
gap in subject < gap in SPEC/CP

Could also be “not clear where gap is”
Conclusion

• One can detect a contrast, but only if one ignores a confound.

• Original claim:
  gap in subject < gap in SPEC/CP

• Once we distenthagle confound, situation becomes clear:
  gap in SPEC/CP < gap in subject < unclear gap position

Could also be “not clear where gap is”
Moral of the story

• Traditional judgments need to be done carefully too, with due concern for true minimal pairs, etc.
Moral of the story

• Traditional judgments need to be done carefully too, with due concern for true minimal pairs, etc.
• Formal experiments can be useful in adjudicating disputes.
Moral of the story

• Traditional judgments need to be done carefully too, with due concern for true minimal pairs, etc.
• Formal experiments can be useful in adjudicating disputes.
• Neither method is unassailable or always done flawlessly. Choose the best tool for the job, and use it correctly.
1. Traditional judgments + formal experiments

- What does it mean when they conflict?
1. Traditional judgments + formal experiments

- What does it mean when they conflict?

• Constraints on \[\text{wh-}\]movement

These cases deserve close scrutiny.

We have eliminated one problematic case.
1. Traditional judgments + formal experiments
   • What does it mean when they conflict?

2. Constraints on wh-movement
   • Are some still completely beyond our understanding?

These cases deserve close scrutiny.
1. Traditional judgments + formal experiments
   - What does it mean when they conflict?

2. Constraints on *wh*-movement
   - Are some still completely beyond our understanding?

These cases deserve close scrutiny.

We have eliminated one problematic case.
Thank you!
grammar.ucsd.edu/syntaxlab