Acceptable extraction from relative clauses in English

Jake W. Vincent
January 25, 2019
Thesis

• Relative clause subextraction in English is tolerated in environments that facilitate non-presuppositional interpretation of the DP host.

• The resistance of the relative clause to extraction in such environments is low, close to zero.
Non-presupposed DP

Presupposed DP

✓

✗
Why?

- Does acceptable RC subextraction pose a challenge to structural accounts of islands?
- Discussion on RC subextraction in English is limited.
- Studying a broad variety of languages is important and can help highlight patterns that are more subtle in other languages.
Outline

1. Selective RC subextraction in other languages
2. Why should we look more closely at English?
3. How to measure the strength of an island
4. Looking at English RC subextraction
5. Discussion and conclusion
Selective RC subextraction in other languages

1. Languages
2. What facilitates RC subextraction?
3. Interim summary
Languages

• Scandinavian languages
  ▪ Danish [1]
  ▪ Swedish [2,3]
  ▪ Norwegian [4]

• Hebrew [5]

• Italian (and other Romance languages) [6]

• English?
Scandinavian languages

• Danish, Swedish, and Norwegian are well-known for having more “porous” relative clauses. [2]

• The phenomenon of RC subextraction is robust enough that traditional grammarians have warned against using such configurations in written language, and have received the name *satsfläta* ‘sentence braid’ in Swedish, *sætningsknude* ‘sentence knot’ in Danish, and *knutesætning* ‘sentence knot’. [2,7]
What facilitates RC subextraction?

- RC subextraction impossible in most environments
- Overarching patterns: RC subextraction is facilitated when...
  - The matrix clause is a canonical existential
  - The matrix predicate serves to introduce the referent of the DP host into the discourse (*see, know,...*)
  - The DP host is the non-verbal predicate of a clause
  - No other impeding factors are present
Canonical existential

• The prototypical way the language asserts or denies existence

(1) Det₁ er der mange der kan lide __₁.
that there many who like
‘That, there are many who like (it).’

(2) Det språket₁ finnas det manga som talar __₁.
that language exist it many that speak
‘That language, there are many who speak (it).’
Canonical existential

(3) [Al lexem šaxor], yeš rak gvina axat on black bread BE only cheese one

še-keday limroax _1.
that-worth to.spread

‘On black bread, there is only one cheese that’s worth spreading.’ Hebrew [5]
(4) Ida, [di cui]₁ non c’è nessuno che sia mai stato innamorato, ...

‘Ida, whom there is nobody that was ever in love with, …’

Italian [6]
(5) Ese es un sitio [en el que]₁ no hay nadie
    this is a place in which nobody
    que querría vivir ₁.
    who would want to live

    ‘This is a place where there is no one that would like
    to live.’

• ...and French.
Non-canonical existentials

• “Non-canonical existentials” [5]
• “Evidential existentials” [8]

• Usually with a first person subject, used to...
  ▪ assert/deny the existence of the referent of the DP host
  ▪ state how evidence of existence was acquired
Non-canonical existentials

(6) Det₃ har jeg mødt mange der har gjort __.  
that have I met many who have done 
‘That, I have met many who have done (it).’ Danish [1]

(7) [Den teorin]₁ känner jag ingen som tror på __.  
that theory know I nobody that believes in 
‘That theory, I know nobody that believes in (it).’  
Swedish [2]
Non-canonical existentials

(8) Marit₁ har vi endelig funnet en gutt som Mary have we finally found a boy that kan hamle opp med __₁. can handle

‘Mary, we have finally found a boy that can handle (her).’ Norwegian [9]
Non-canonical existentials

(9) [Miškafayim yerukot ka-ele]₁, ra’iti kan eyeglasses green like-that saw.l here etmol mišehu še-moxer. yesterday someone that-sells ‘That kind of green eyeglasses, I saw here yesterday someone who sells (them).’ Hebrew [5]
Non-canonical existentials

(10) Giorgio, al quale non conosco nessuno che sarebbe disposto ad affidare i propri risparmi, ...

‘Giorgio, whom I don’t know anybody that would be ready to entrust with their savings, ...’

*Italian [6]*
Predicate nominals

• DP host functions as the predicate of a clause

• The configuration:

\[ [\text{FILLER}]_1 \ldots \text{be} [\text{DP} [\text{RC} \ldots \_]_1] \]

(11) Surströmming₁ är Fredrik den ende som fermented.herring is Fredrik the only.one who
tycker om \_1.
likes \_1. \text{Swedish} [2]

‘Fermented herring, Fredrik is the only one who likes (it).’
Predicate nominals

(12) [Et ha-toxnit ha-zot], ata ha-yaxid ACC the-program the-this you the-single še-ro’e —1.
that-watches

‘This program, you’re the only one who watches (it).’

Hebrew [5]
Interim summary

• The DP host sits in a matrix clause that is...
  ▪ A canonical existential
  ▪ A non-canonical existential
  ▪ A copular clause, where the DP host is the main predicate

• Examples in which the matrix predicate is an ordinary transitive verb are ungrammatical.
Why should we look more closely at English?
Why look at English?

• Previous theoretical work has pointed out relatively acceptable examples [10,11,12]

(13) This is the child who_1 there is nobody who is willing to accept __1.          [10]

(14) Then you look at what happens in languages that you know and languages_1 that you have a friend who knows __1.          [11]

(15) Isn’t that the song_1 that Paul and Stevie were the only ones who wanted to record __1?          [12]
Why look at English?

• Kush et al. (2013) present experimental evidence that island effects are attenuated in English in certain environments that facilitate RC subextraction in Swedish.

  ▪ Grammatical illusion?

• In light of findings on Scandinavian languages, Hebrew, and Romance languages, we might have missed something.
How to measure the strength of an island

1. The length by complexity design
Measuring island strength

• “Length by complexity” design for acceptability judgment experiments [14]

• Factorial design that allows one to isolate island violation effects from other potentially confounding factors

(16) *Which dog$_1$ did your toddler bite the neighbor

\[ [\text{RC who owns } \_1] \]?

• What factors could be influencing ratings that might be given to (16)?
Measuring island strength

(16) *Which dog did your toddler bite the neighbor [RC who owns ___]?

- Island violation effect
- Length of extraction
- Complexity associated with the relative clause
- Baseline costs associated with lexical items
- Factors: extraction length, complexity of embedded clause
Measuring island strength

- Baseline sentence
  - Keep lexical items relatively constant
  - No length cost (matrix subject extraction)
  - No RC-related complexity cost (embedded clause ≠ A-bar CP)

- The matrix verb: select for either a DP or a CP

(17) Who understands \[
\begin{cases}
\text{DP teachers who dislike... } \\
\text{CP that teachers dislike... }
\end{cases}
\]
Measuring island strength

(18) a. Who$_1$ __$_1$ understands that teachers dislike unstapled papers?
    NON-ISL | SHORT

b. What$_1$ does Lorena understand that teachers dislike __$_1$?
    NON-ISL | LONG

c. Who$_1$ __$_1$ understands teachers who dislike unstapled papers?
    ISLAND | SHORT

d. What$_1$ does Lorena understand teachers who dislike __$_1$?
    ISLAND | LONG
Measuring island strength

<table>
<thead>
<tr>
<th>Condition</th>
<th>Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>NON-ISLAND</td>
<td>SHORT</td>
</tr>
<tr>
<td>NON-ISLAND</td>
<td>LONG</td>
</tr>
<tr>
<td>ISLAND</td>
<td>SHORT</td>
</tr>
<tr>
<td>ISLAND</td>
<td>LONG</td>
</tr>
</tbody>
</table>
Measuring island strength

• Isolating the effect of island subextraction is achieved arithmetically:

\[
\text{DIFFERENCE 1 (D1)} = \beta + \text{Length} \quad \text{NON-ISL | LONG} \\
\quad - \beta + \text{Length} + \text{Complexity} + \text{Isl. subextraction} \quad \text{ISL | LONG} \\
\quad = -(\text{Complexity} + \text{Isl. subextraction}) \quad D1
\]
Measuring island strength

DIFFERENCE 2 (D2)

\[ \beta - \beta + \text{Complexity} \]

= −Complexity

\[ \text{NON-ISLAND | SHORT} \]

\[ \text{ISLAND | SHORT} \]

\[ \text{D2} \]
Measuring island strength

DIFFERENCE OF DIFFERENCES (DD)

\[-(\text{Complexity} + \text{Isl. subextraction})\]

\[-\text{Complexity}\]

\[= -\text{Isl. subextraction}\]
RC subextraction in English

1. Experiment 1
Design

• Goal: compare RC resistance to extraction in three different environments:
  ▪ Transitive objects (uncontroversially bad) \( \mapsto \) OBJECT
  ▪ Existential \( \mapsto \) EXIST(ENTIAL)
  ▪ Predicate nominal \( \mapsto \) PRED(ICATE)

• Adding this three-level factor, we end up with a \( 2 \times 2 \times 3 \) factorial design (=12 conditions/item)
Sample item (object conditions)

a. Who thinks that Courtney saw that only one art collector bid on this painting?  
   NON-ISL | SHORT

b. Which painting do you think that Courtney saw that only one art collector bid on?  
   NON-ISL | LONG

c. Who thinks that Courtney saw the only art collector who bid on this painting?  
   ISLAND | SHORT

d. Which painting do you think that Courtney saw the only art collector who bid on?  
   ISLAND | LONG
Sample item (predicate conditions)

a. Who₁ thinks that Courtney believes that only one art collector bid on this painting?  
   NON-ISL | SHORT

b. Which painting₁ do you think that Courtney believes that only one art collector bid on __₁?  
   NON-ISL | LONG

c. Who₁ thinks that Courtney believes that she is the only art collector who bid on this painting?  
   ISL | SHORT

d. Which painting₁ do you think that Courtney believes that she is the only art collector who bid on __₁?  
   ISL | SHORT
Sample item (existential conditions)

a. Who$_1$$_1$ thinks that there is only one art collector bidding on this painting?

b. Which painting$_1$ do you think that there is only one art collector bidding on $_1$?

c. Who$_1$$_1$ thinks that there is only one art collector who bid on this painting?

d. Which painting$_1$ do you think that there is only one art collector who bid on $_1$?
Materials

• 36 items separated into 12 lists via Latin Square
  ▪ 3 observations/participant/condition
  ▪ 36 experimental sentences per participant

• 72 filler sentences adapted from Sprouse et al. (2013) study [15]
  
• Filler sentences adjusted so that half of all sentences seen by the participant were definitely grammatical.

• Sentence types: half declaratives, half WH-questions

• Half of fillers contained only since all of experimental sentences did.
Procedure and participants

• Acceptability judgment task run on IBEX
• 1–6 Likert scale
• 48 individuals recruited on MTurk/TurkPrime
  ▪ Paid $5.00 for their participation
  ▪ Native English speakers only
• 2 participants’ data was excluded because ungrammatical fillers were rated higher on average than grammatical fillers.
Analysis

• Mixed effects ordinal regression w/ cumulative link
• Dependent variable = rating
• Environment, Complexity, and Length factors and their interactions set as fixed effects
• Maximal random effects structure
• Helmert contrast coding
  ▪ Predicate and Existential conditions compared directly (“BE” comparison)
  ▪ Object conditions compared to the mean of the Predicate and Existential conditions (“TRANSITIVITY” comparison)
Analysis

• Helmert contrast coding–enabled comparisons:

```
TRANSITIVITY
  OBJECT
  BE
    PREDICATE
    EXISTENTIAL
```
Predictions

• Main effects of both length and complexity

• General island effect
  ▪ (interaction between Length and Complexity)

• Assuming Existential and Predicate environments facilitate RC subextraction, attenuated island effect in these conditions
  ▪ (interaction between Length, Complexity, and TRANSITIVITY)
Results

The graphs display the mean rating of 'object', 'predicate', and 'exist' across different lengths (short and long) and complexity (Non-island and Island). The x-axis represents the length, with 'short' on the left and 'long' on the right. The y-axis shows the mean rating, ranging from 3 to 5.
Results: Main effects

• Main effects for each Environment (object lowest on average, followed by predicate, followed by existential; all $ps < 0.001$)

• Main effect of Length ($p < 0.001$)

• Main effect of Complexity ($p < 0.001$)
Results: Interactions

• Interaction of Length and Complexity ($p < 0.001$)
  - General island effect (across environments)

• Interaction between Length, Complexity, and TRANSITIVITY ($p = 0.031$)
  - Object RC subextraction conditions significantly less acceptable than both Predicate and Existential subextraction conditions

• Interaction between Length, Complexity, and BE was not significant ($p = 0.558$)
Island strength (DD scores)

• Using z-scored ratings...

<table>
<thead>
<tr>
<th>ENVIRONMENT</th>
<th>NON–ISLAND, SHORT</th>
<th>NON–ISLAND, LONG</th>
<th>ISLAND, SHORT</th>
<th>ISLAND, LONG</th>
<th>D1 (COMPLEXITY + ISL. SUBEXT.)</th>
<th>D2 (COMPLEXITY)</th>
<th>DD (ISL. SUBEXT.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBJECT</td>
<td>0.17</td>
<td>0.09</td>
<td>0.17</td>
<td>−0.53</td>
<td>0.62</td>
<td>0.00</td>
<td><strong>0.62</strong></td>
</tr>
<tr>
<td>PREDICATE</td>
<td>0.30</td>
<td>0.04</td>
<td>0.29</td>
<td>−0.13</td>
<td>0.18</td>
<td>0.02</td>
<td><strong>0.16</strong></td>
</tr>
<tr>
<td>EXISTENTIAL</td>
<td>0.85</td>
<td>0.42</td>
<td>0.71</td>
<td>0.02</td>
<td>0.40</td>
<td>0.14</td>
<td><strong>0.26</strong></td>
</tr>
</tbody>
</table>
Discussion

1. Interpretation of results
2. Why would this have gone largely unnoticed?
Discussion

- Extracting from RCs in English is not unacceptable across the board

- When the DP that hosts the RC is in an Existential or Predicate nominal environment, RCs are substantially more transparent to extraction
  - DPs in situ and non-presupposed (no freezing effects)

- RC subextraction may be grammatical in English in limited circumstances
Discussion

• Why would this have gone relatively unnoticed in English?

• Scandinavian languages use fronting for multiple information-structural purposes
  ▪ More situations in which long-distance fronting is employed

• English has fewer long-distance fronting strategies and Topicalization is relatively marked, so there are fewer opportunities to observe acceptable RC subextraction
Discussion

• Affirms the importance of studying a variety of languages
  ▪ Even patterns less readily observable in some languages can be observed with sufficient insight from other languages

• Affirms the importance of more controlled experimentation
  ▪ Judgments are subtle and easily called into question, but experimentation provides a more precise measurement of the effect
Thank you!
References


