# Acceptable extraction from relative clauses in English

Jake W. Vincent
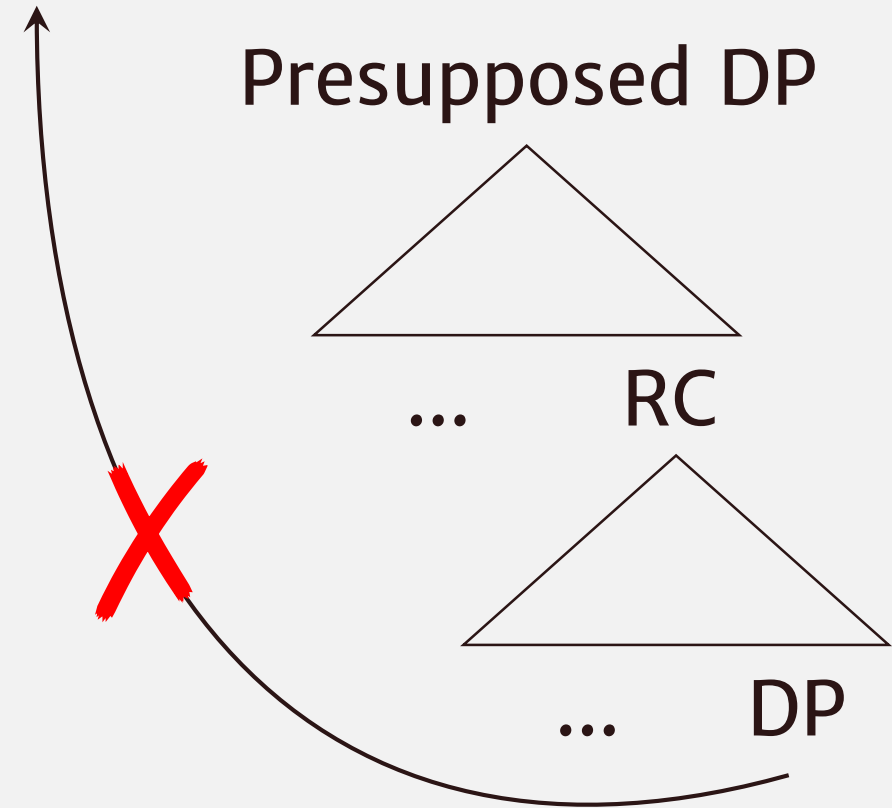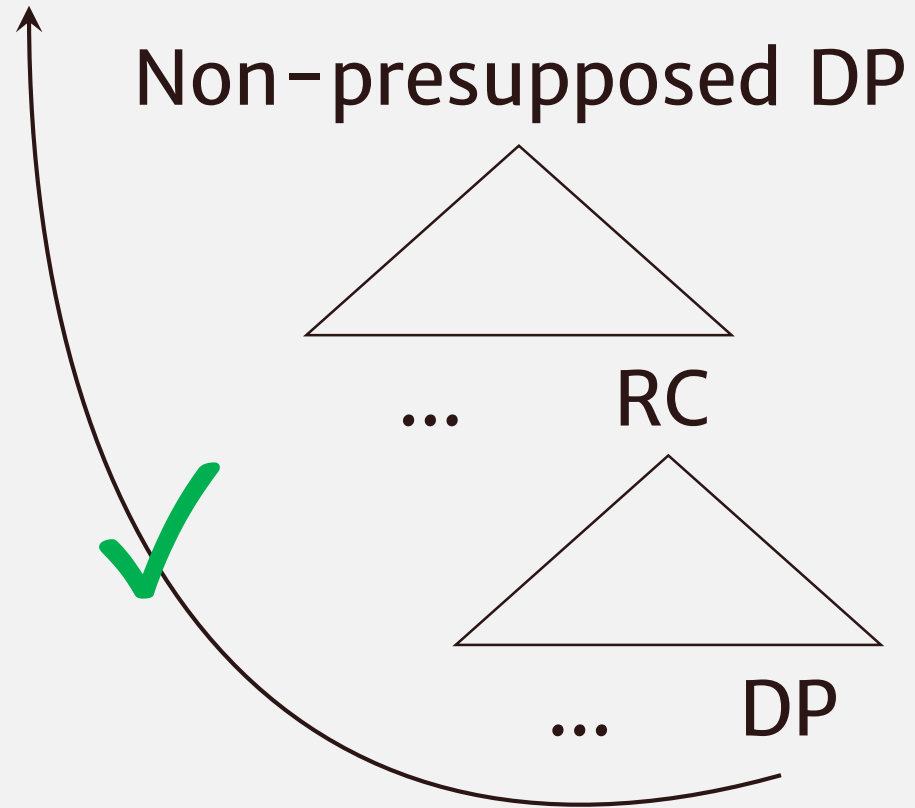
January 25, 2019

# 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.

# Thesis



Non-presupposed DP

... RC

... DP

Presupposed DP

... RC

... 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) Det$_1$ er der mange der kan lide __$_1$.
that are there many who like

'That, there are many who like (it).' Danish [1]

(2) Det språket$_1$ finns det manga som talar __$_1$.
that language exist it many that speak

'That language, there are many who speak (it).'
Swedish [2]

# Canonical existential

(3)  [Al lexem šaxor]$_1$, 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]

# Canonical existential

(4) Ida, [di cui]$_1$ non c'è nessuno che sia
Ida whom not there.is nobody that be
mai stato innamorato __$_1$, …
never in love with
'Ida, whom there is nobody that was ever in love with, …'

# Canonical existential

(5) Ese es un sitio [en el que]$_1$ no hay nadie
this is a place in.which not exist nobody

que querría vivir __$_1$.
who would.want to.live

'This is a place where there is no one that would like to live.' Spanish [6]

- ...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) Det$_1$ har jeg mødt mange der har gjort __$_1$.
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 __$_1$.
that theory know I nobody that believes in

'That theory, I know nobody that believes in (it).'

Swedish [2]

# Non-canonical existentials

(8) Marit$_1$ 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.I  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$_1$, al quale non conosco nessuno che   sarebbe
       Giorgio   whom   not  know.I   nobody  that would.be

       disposto  ad  affidare  i     propri  risparmi, ...
       disposed  to  entrust   the  own    savings

‘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:

$$[\text{FILLER}]_1 \ldots be \; [_{DP} \; [_{RC} \ldots \underline{\quad}_1]$$

(11) Surströmming$_1$       är Fredrik  den  ende       som
     fermented.herring    is Fredrik  the  only.one  who

     tycker om                      $\underline{\quad}_1$.
     likes                                        Swedish [2]

'Fermented herring, Fredrik is the only one who likes (it).'

# Predicate nominals

(12) [Et   ha-toxnit      ha-zot]$_1$, 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).'

# 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. [13]

  - 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

   $[_{RC}$ who owns $\_\__1]$?

- What factors could be influencing ratings that might be given to (16)?

# Measuring island strength

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

[$_{RC}$ who owns __$_1$]?

- 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 $\left\{ \begin{array}{l} [_{DP} \text{ teachers who dislike... }] \\ [_{CP} \text{ that teachers dislike... }] \end{array} \right.$

# 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

| Condition | Costs |
|---|---|
| NON−ISLAND \| SHORT | β |
| NON−ISLAND \| LONG | β + Length |
| ISLAND \| SHORT | β + Complexity |
| ISLAND \| LONG | β + Length + Complexity + Island subextraction |

# Measuring island strength

- Isolating the effect of island subextraction is achieved arithmetically:

$$\text{DIFFERENCE 1 (D1)}$$

$\beta$ + Length  NON-ISL | LONG

− $\beta$ + Length + Complexity + Isl. subextraction  ISL | LONG

_____

= −(Complexity + Isl. subextraction)  D1

# Measuring island strength

$$\text{DIFFERENCE 2 (D2)}$$

| | |
|---|---|
| β | NON-ISLAND | SHORT |
| − β + Complexity | ISLAND | SHORT |
| = −Complexity | D2 |

# Measuring island strength

## DIFFERENCE OF DIFFERENCES (DD)

-(Complexity + Isl. subextraction)     D1

−   -Complexity                                      D2

——————————————————————————

=  -Isl. subextraction                          DD

# RC subextraction in English

1. Experiment 1

# Design

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

- Adding this three-level factor, we end up with a 2×2×3 factorial design (=12 conditions/item)

# Sample item (object conditions)

a. Who$_1$ __$_1$ thinks that Courtney saw that only one art collector bid on this painting?   NON-ISL | SHORT

b. Which painting$_1$ do you think that Courtney saw that only one art collector bid on __$_1$?   NON-ISL | LONG

c. Who$_1$ __$_1$ thinks that Courtney saw the only art collector who bid on this painting?   ISLAND | SHORT

d. Which painting$_1$ do you think that Courtney saw the only art collector who bid on __$_1$?   ISLAND | LONG

# Sample item (predicate conditions)

a. Who$_1$ __$_1$ thinks that Courtney believes that only one art collector bid on this painting?                    NON-ISL | SHORT

b. Which painting$_1$ do you think that Courtney believes that only one art collector bid on __$_1$?                    NON-ISL | LONG

c. Who$_1$ __$_1$ thinks that Courtney believes that she is the only art collector who bid on this painting?                    ISL | SHORT

d. Which painting$_1$ 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
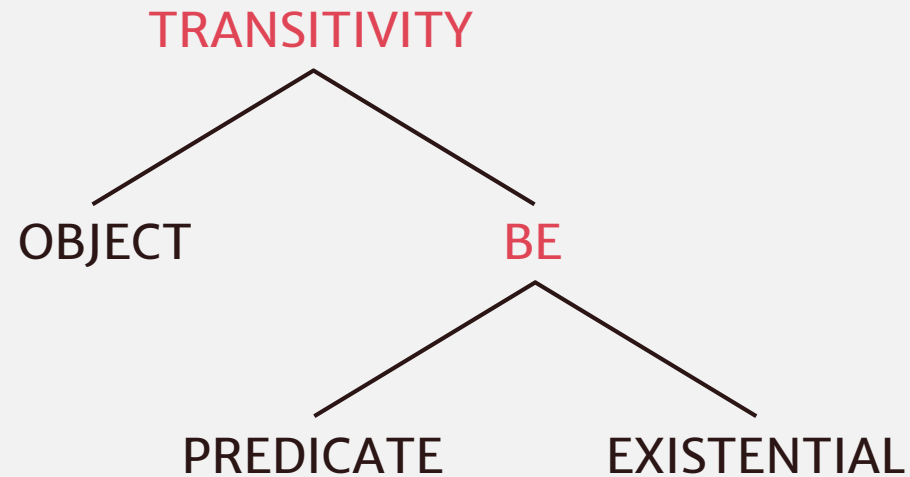
# Predictions

- 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
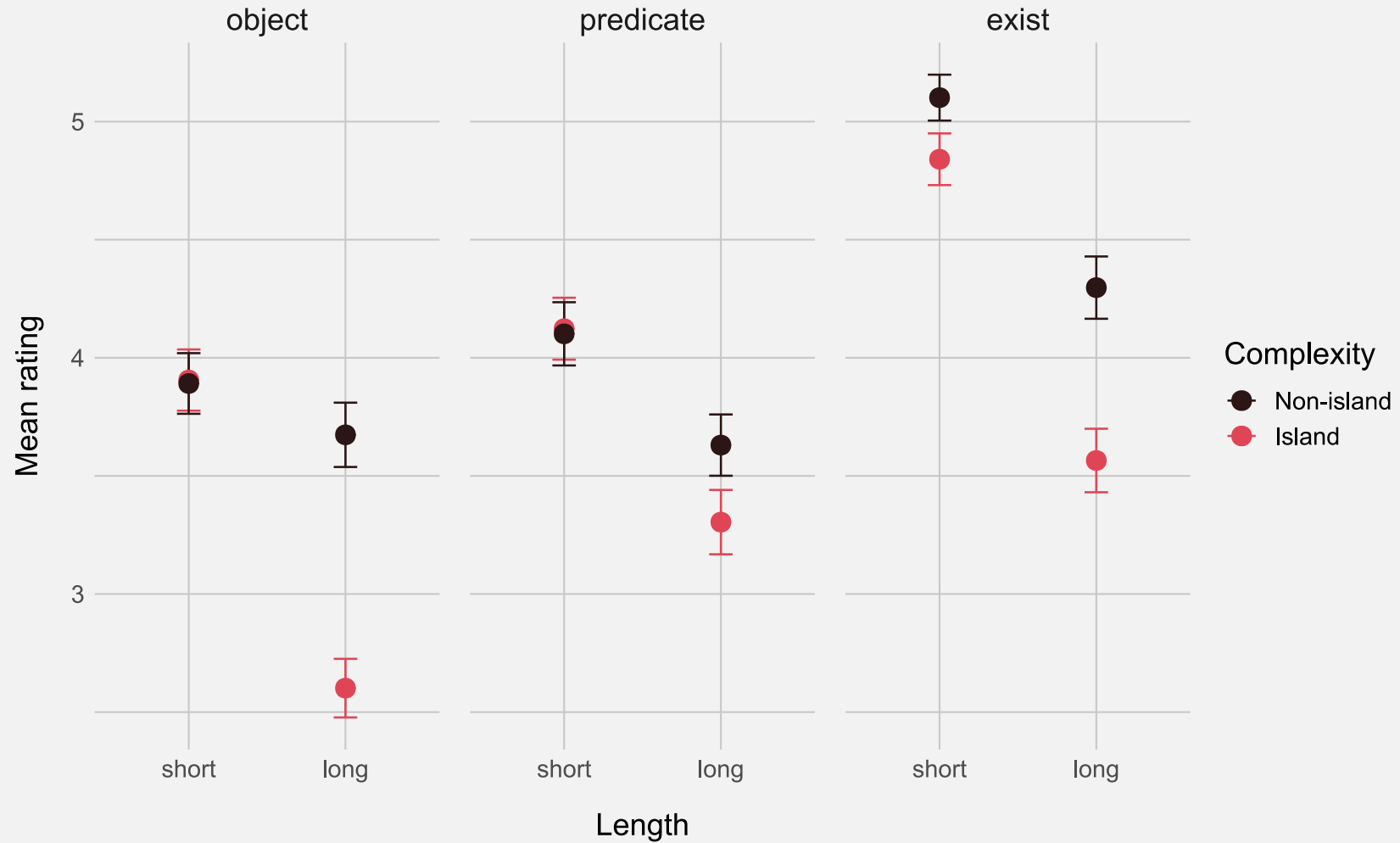
# Results: Main effects

- Main effects for each Environment (object lowest on average, followed by predicate, followed by existential; all $p$s < 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…

| ENVIRONMENT | NON-ISLAND, SHORT | NON-ISLAND, LONG | ISLAND, SHORT | ISLAND, LONG | D1 (COMPLEXITY + ISL. SUBEXT.) | D2 (COMPLEXITY) | DD (ISL. SUBEXT.) |
|---|---|---|---|---|---|---|---|
| OBJECT | 0.17 | 0.09 | 0.17 | −0.53 | 0.62 | 0.00 | **0.62** |
| PREDICATE | 0.30 | 0.04 | 0.29 | −0.13 | 0.18 | 0.02 | **0.16** |
| EXISTENTIAL | 0.85 | 0.42 | 0.71 | 0.02 | 0.40 | 0.14 | **0.26** |

# 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

1. Erteschik-Shir, N., & Lappin, S. (1979). Dominance and the functional explanation of island phenomena. *Theoretical Linguistics*, 6(1), 41--86. https://doi.org/10.1111/j.1467-9841.2007.00319.x

2. Engdahl, E. (1997). Relative Clause Extractions in Context. In *Working Papers in Scandinavian Syntax* (pp. 51–79).

3. Lindahl, F. (2017). *Extraction from relative clauses in Swedish*. University of Gothenburg. Retrieved from http://hdl.handle.net/2077/51985

4. Taraldsen, K. T. (1982). Extraction from relative clauses in Norwegian. In E. Engdahl & E. Ejerhed (Eds.), *Readings on Unbounded Dependencies in Scandinavian Languages* (Vol. 43, pp. 205-221). Stockholm, Sweden: Almqvist & Wiksell International.

5. Sichel, I. (2018). Anatomy of a Counterexample: Extraction from relative clauses. *Linguistic Inquiry*, 49(2).

6. Cinque, G. (2010). On a selective "violation" of the Complex NP Constraint. In J.-W. Zwart & M. de Vries (Eds.), *Structure Preserved: Studies in syntax for Jan Koster* (Vol. 164, pp. 81--89). Amsterdam/Philadelphia: John Benjamins Publishing Company.

7. Allwood, J. (1982). The Complex NP Constraint in Swedish. In E. Engdahl & E. Ejerhed (Eds.), *Readings on Unbounded Dependencies in Scandinavian Languages* (Vol. 43, pp. 15--32). Stockholm, Sweden: Almqvist & Wiksell International.

# References (continued)

8. Rubovitz-Mann, T. (2000). *Extractions from relative clauses: An information-structural account*. Hebrew University of Jerusalem.

9. Taraldsen, K. T. (1981). The Theoretical Interpretation of a Class of "Marked" Extractions. In A. Belletti, L. Brandi, & L. Rizzi (Eds.), *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW Conference* (pp. 475–516). Pisa, Italy: Scuola Normale Superiore di Pisa.

10. Kuno, S. (1976). Subject, Theme, and the Speaker's Empathy--A Reexamination of Relativization Phenomena. In C. N. Li (Ed.), *Subject and Topic* (pp. 417--444). New York, NY: Academic Press, Inc.

11. Kuno, S. (1976). Subject, Theme, and the Speaker's Empathy--A Reexamination of Relativization Phenomena. In C. N. Li (Ed.), *Subject and Topic* (pp. 417--444). New York, NY: Academic Press, Inc.

12. Chung, S., & McCloskey, J. (1983). On the Interpretation of Certain Island Facts in GPSG. *Linguistic Inquiry*, 14(4), 704–713. Retrieved from http://www.jstor.org/stable/4178357

14. Sprouse, J., Wagers, M. W., & Phillips, C. (2012). A test of the relation between working memory capacity and syntactic island effects. *Language*, 88(1), 82–123.

15. Sprouse, J., Schütze, C. T., & Almeida, D. (2013). A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001-2010. *Lingua*, 134, 219–248. https://doi.org/10.1016/j.lingua.2013.07.002