ISLAND EXTRACTIONS IN THE WILD

A corpus study of adjunct and relative clause extractions in Danish and English

Christiane Müller & Clara Ulrich Eggers

MINDS, October 29
LONG-DISTANCE DEPENDENCIES

- Languages can have long-distance dependencies (see AS-08, pp. 8-9):

  (1) [Which book]$_i$ did Harry think [$_j$ that Bill bought $_i$]?

- Other types of long-distance dependencies are topicalization (2) and relativization (3):

  (2) [That book]$_i$ Harry thought [$_j$ that Bill should buy $_i$].

  (3) This is [the book]$_i$ that Harry thought [$_j$ that Bill should buy $_i$].
LONG-DISTANCE DEPENDENCIES

- Long-distance dependencies can in principle apply across an indefinite number of clauses:

\[(4)\]

a. What \( i \) did Bill buy \( _i \)?
b. What \( i \) did you force Bill to buy \( _i \)?
c. What \( i \) did Harry say you had forced Bill to buy \( _i \)?
d. What \( i \) was it obvious that Harry said you had forced Bill to buy \( _i \)?

- But certain syntactic environments block the formation of such long distance dependencies!
ISLANDS

- **Relative clauses:**
  
  (5) *[Which book], did John meet [a child who read _]?*

- **Subjects:**
  
  (6) *[Who], did [pictures of _] annoy Bill?*

- **Adjunct clauses:**
  
  (7) *[Who], did John arrive [after Bill kissed _]?*

  - Such domains are termed *islands* (Ross 1967)

  - The same effect arises with other dependencies:
    
    (8) *[This girl], John arrived [after Bill kissed _].*
ISLANDS

- A central question in linguistic theory has been what the source of such island effects is.
- I.e. why is extraction possible e.g. in (9a) but not in (9b) or (9c)?

(9)  

a. Who, did John say [that Bill kissed _]?  
b. *Who, did John arrive [after Bill kissed _]? (adjunct island)  
c. *Who, did John know [a girl that kissed _]? (relative clause island)

- Many accounts assume that extraction from islands is ruled out by syntactic island constraints, and that these island constraints apply universally:
  - The Complex NP Constraint (Ross 1967): Extraction from complex NPs is disallowed.
  - The Condition on Extraction Domains (Huang 1982): Extraction is disallowed from domains that are not properly governed (roughly, non-complements).
ISLAND EXTRCTIONS IN MSC.

- However, the claim that island constraints are universal has been challenged by data from the Mainland Scandinavian (MSc.) languages:

(10) **Relative clause extraction (by topicalization)**

a. [De blommorna], känner jag [en man som säljer _].

   *those flowers*  *know I a man who sells*

   ‘I know a man who sells those flowers.’ (Swedish; Allwood 1982: 24)

b. [Dette biletet], kjener eg [den målaren som har måla _].

   *that picture*  *know I the painter who has painted*

   ‘I know the painter who has painted this picture.’ (Norwegian; Faarlund et al. 1997: 1099)

c. [Suppe], kender jeg [mange der kan lide _].

   *soup*  *know I many who can like*

   ‘I know many people who like soup.’ (Danish; Erteschik-Shir 1973: 67)
Adjunct clause extraction (by topicalization)

a. [SportspegeIn], somnar jag [när jag ser _].
   sports program.the fall.asleep I when I watch
   ‘I fall asleep when I watch the sports program.’ (Swedish; Anward 1982: 74)

b. [Den saka], ventar vi her [mens de ordnar _].
   this thing wait we here while they fix
   ‘We are waiting here while they fix this thing.’ (Norwegian; Faarlund 1992: 117)

c. [Den vase], får du ballade [hvis du taber _].
   this vase get you trouble if you drop
   ‘You are in trouble if you drop this vase.’ (Danish; Hansen & Heltoft 2011: 1814)

- Examples like these violate island constraints that are assumed to apply universally.
- So: Is there cross-linguistic variation in islands?
FORMAL VS. INFORMAL DATA

- The claims that such structures are acceptable in MSc. are based on informal judgments.
- However, formal studies of Scandinavian island extractions yield rather low ratings:
  - Swedish relative clause extractions scored ratings on the lower half of the scale in studies by Müller (2015), Wiklund et al. (2017), and Tutunjian et al. (2017)
  - Kush et al. (2018): extraction from relative clauses and from (conditional) adjunct clauses in Norwegian yielded acceptability scores at the bottom end of the scale
- What is actually possible in MSc.???
ACCEPTABLE EXTRACTION IN ENGLISH?

- Anecdotal evidence of acceptable extractions from finite adjunct clauses (12) and relative clauses (13) in English:

(12) a. This is [the watch] that I got upset [when I lost _]. (Truswell 2011: 175)

b. a stranger, from [that remote and barbarian Isle] which the Imperial Roman shivered [when he named _], paused. (Haegeman 2004: 70)

(13) a. Violence is [something] that there are [many Americans who condone _].
(McCawley 1981: 108)

b. John is [the sort of guy] that I don’t know [a lot of people who think well of _].
(Culicover 1999: 230)

How different are English and MSc. really?
A CORPUS STUDY

- We conduct an exploratory corpus study on:
  - adjunct clause extraction + relative clause extraction
    (Danish & English) (Danish & English)

- Why a corpus study?
  - We want to explore naturally produced examples – this will provide further insights into what is in fact possible in Danish and English
  - If we can find examples in the corpora, this would demonstrate that island extraction is not just a peripheral phenomenon restricted to isolated constructed examples
  - Most acceptability studies of island constructions aim to use test sentences that are modelled on naturally occurring data.
    ➡️ We can uncover potential patterns or trends among the found extractions.
**METHOD**

- Danish corpora: *KorpusDK* (+ *BySoc, Samtalegrammatik*)
- English corpora: *British National Corpus* (BNC),
  *Corpus of Contemporary American English* (COCA)
- complemented by examples retrieved from a Google search

We used a combination of different search strings that can target relative clause and adjunct clause extractions (if the corpus contains any).

- Restriction: Our corpora are not annotated in a way that makes it possible to search for extraction constructions *directly*.  
  This means that we can’t get quantitative data or compare frequencies across constructions or across languages.
- Rather, our results can provide informal insights into what appears to be possible at all, or “common”, among the found extraction instances.
ADJUNCT CLAUSE EXTRACTION - FINDINGS

DANISH:

- > 30 instances of extraction from *hvis* (‘if’) - and *når* (‘when’)-adjunct clauses in Danish
  
  (14) a. *især [den sidste], ville jeg blive sur, [hvis du ikke får _].*
  
  b. *fordi du værdsætter hundenes liv højt <3 og [det], bliver jeg glad [når jeg læser _].*
  
  c. *Det er [et nødvendigt mineral], som vi ville dø, [hvis vi ikke fik _].*
  
  d. *Her er lige [et par billeder af en frederiksboger, en dansk varmblod og nogle fjordheste], som jeg ville blive glad [hvis I gad bedømme _].* (from Google search)

- A search on KorpusDK yielded > 600 instances of extraction from *at*-clauses (not islands; e.g. “[Den hest], er jeg glad for [at du kan lide _].”)

  • Extraction from (some) adjunct clauses appears to be possible, but less frequent than e.g. *at*-extraction.
ADJUNCT CLAUSE EXTRACTION - FINDINGS

ENGLISH:

- \( \approx 50 \) instances of extraction from *if-* and *when-*adjunct clauses (only by relativization)

(15) a. Many of the exercises are [ones], that I would be surprised [if even 1 percent of healthy women can do _]. (COCA)

b. Now, those are [things] that I feel very warm [when I look at _], and I wouldn't want to live in a house that they - a house that didn't have room for those. (COCA)

- Naturally produced examples of adjunct clause extraction were found as easily in English as in Danish.

\[\text{– Potentially challenging the claim that English differs from the Scandinavian languages in never allowing extraction from finite adjunct clauses.}\]
ADJUNCT CLAUSE EXTRACTION - FINDINGS

PATTERNS: ADJUNCT CLAUSE TYPE

- All examples found in English as well as in Danish featured *if*- and *when*-clauses, with extraction from *if*-clauses being the most common.
  - This result matches recent experimental findings that at least *if*-adjuncts do not behave like absolute islands for extraction in English (Nywad et al. forthcoming).

- No examples of extraction from *because*-clauses were found.
  - Matching previous observations that conditional (*if/hvis*) and temporal (*when/når*) adjuncts appear to be more permissive for extraction than causal adjuncts (e.g. Müller 2019; Bondevik et al. 2020).
ADJUNCT CLAUSE EXTRACTION - FINDINGS

PATTERNS: EXTRACTION TYPE

- In Danish, most cases of adjunct clause extraction found involved relativization (and in English, only examples involving relativization were found).
  - Most existing acceptability studies of island structures have tested extraction in the form of topicalization and wh-movement.

Contrast between experimental designs and natural data
RELATIVE CLAUSE EXTRACTION

Preliminary findings – Danish:

- Examples of extraction from relative clauses almost always involve a presentational matrix verb (i.e. "presentational relatives"): 

  (16) a. [De store bøger], er det efterhånden næsten kun [humanister der skriver _].
  b. [Den gave], er vi [mange, der siger tak for _].
  c. Men [den slags spekulationer], kender jeg ikke [nogen, der tør fremsætte offentligt _].

- These findings match the English examples of relative clause extraction reported to be acceptable in the literature (Chaves & Putnam 2020: 67).

Prediction for English:

- Examples will also primarily involve presentational matrix verbs.
- This would be a further indication that English and Danish are possibly more similar with regard to island extractions than previously assumed.
REFERENCES


REFERENCES


## Relative Clause Extraction

### Extraction by topicalization – KorpusDK findings:

<table>
<thead>
<tr>
<th>Extracted element</th>
<th>jeg/du/vi/l/de</th>
<th>han/hun/man</th>
<th>det</th>
<th>der</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>&quot;Heavy&quot; element =</strong> headed by a noun</td>
<td>7 (1) [Den gave], er vi [mange, der siger tak for __].</td>
<td>1 (0) [Den lovovertrædelse], er han næppe [den eneste, der har begået __].</td>
<td>2 (0) [De store bager], er det efterhånden næsten kun [humanister der skriver __].</td>
<td>36 (16) [De der æg], er der da [ingen, der tager skade af __].</td>
</tr>
<tr>
<td>Demonstrative/personal pronoun of the type “den/det/dem”, (En. “that/those”; [ortho=&quot;[Den</td>
<td>Det</td>
<td>Dem]&quot;])</td>
<td>32 (30) [Det], var vi [nogle stykker som var enige om __].</td>
<td>3 (1) [Det], var han [den eneste i familien, der kaldte mig __].</td>
</tr>
<tr>
<td>Personal pronoun of the type “mig/dig/hende/ham/os/jer”, (En. “me/you/her/him/us/you”; [ortho=&quot;[Mig</td>
<td>Dig</td>
<td>Hende</td>
<td>Ham</td>
<td>Os</td>
</tr>
</tbody>
</table>