

Topicalization: The IO/DO Asymmetry in Icelandic

Elena Callegari & Anton Karl Ingason

1.

Double-Object Constructions in Icelandic

The Basics

Double-Object Constructions

In principle, two word order options:

1. IO < DO

Ég gaf Elínu bókina
I gave Eileen(DAT) book.the(ACC)
'I gave Eileen the book'

2. DO < IO

Ég gaf bókina Elínu
I gave book.the(ACC) Eileen(DAT)
'I gave Eileen the book'

Double-Object Constructions

IO < DO : the unmarked order

DO < IO : *the marked configuration*

Collins & Thráinsson (1996), Falk (1990), Ottósson (1991): the availability of DO < IO is dependent on information structure (focus interpretation), stress patterns, definiteness, idiomaticity.



The inverted object [i.e., the DO in DO < IO structures] must preferably be unfocussed, it seems, or at least not more focused than the unmoved object. Thus, the inverted object is often definite, but can be indefinite if the other object is also indefinite.

(Ottóson 1991: 94)

Double-Object Constructions

Dehé (2004) ran an acceptability-judgement task on the IO > DO/DO > IO alternation in Icelandic. Her findings:

- The DO > IO order was very infrequent and was always rated poorly (either as *rather odd* or as *plainly wrong*).
- The **IO > DO is preferred even when the IO is in focus.**
- Dehé (2004) argues that Icelandic uses prosody rather than movement to mark focus.
- If IO is in focus: **stress shift**. The DO undergoes de-stressing and main stress is relocated to the rightmost stressable unit of the IO.

Our Research Question

What happens when we move from (1)...

1. *Ég* *gaf* *Elínu* *bókina*
I *gave* *Eileen(DAT)* *book.the(ACC)*

... To (7-8)?

7. *Elínu* *gaf* *ég* *bókina* **IO Top**
Eileen(DAT) *gave* *I* *book.the(ACC)*

8. *Bókina* *gaf* *ég* *Elínu* **DO Top**
book.the(ACC) *gave* *I* *Eileen(DAT)*

Our Research Questions

- Is the preference for $IO < DO$ maintained even when the two objects no longer appear in the same local domain?

I.e., is the preference for $IO < DO$ a linear type of constraint (e.g. it applies regardless of the relative distance and structure between DO and IO)?

Local / Non-Local Asymmetry

Relative order of focus and “if” Italian (Callegari, 2018):

Locally = *FOC < if

9. Mi domando **se** A GIANNI hai parlato.
Refl I-wonder if TO GIANNI you-have spoken
'I wonder if you have spoken to GIANNI (not to JOHN)

10. *Mi domando A GIANNI **se** hai parlato
Refl I-wonder TO GIANNI if you-have spoken

Non-locally = FOC < if

11. A GIANNI mi domando **se** hai parlato
TO GIANNI refl I-wonder **if** you-have spoken

2.

The study

IcePaHC

IcePaHC

- *Icelandic Parsed Historical Corpus* (currently, 1,002,390 words).
- Diachronic corpus of Icelandic written texts, from the 12th to the 21st century.
- The IcePaHC contains different genres: scientific, legal, narrative, religious....

IcePaHC

We searched the IcePaHC using PaCQL (*Parsed Corpus Query Language*, Ingason 2016) through the online platform treebankstudio.org

Treebank Studio (PREVIEW)

Search Documentation

- Treebank Studio is an online tool for searching parsed corpora using the PaCQL query language.
- The current preview version is configured to search IcePaHC (The Icelandic Parsed Historical Corpus).
- Look at the documentation page for advice on how to use PaCQL.

Search IcePaHC 0.9 (PaCQL): [Anchor: IP-(MAT|SUB)] Anchors only Web output ▾

```
1 IP-(MAT|SUB) idoms MDPI
2
3 ov:1
4 MDPI sprec NP-OB[12]
5 NP-OB[12] sprec VB
6
7 ov:0
8 MDPI sprec VB
9 VB sprec NP-OB[12]
10
11 meta:
12 text century
13 node nodewords NP-OB[12]
```

Clear Submit

IcePaHC

- We searched for all instances of **matrix-clause double-object constructions**, and coded these depending on whether
 - (i) IO topicalization had occurred,
 - (ii) DO topicalization had occurred,
 - (iii) no topicalization had occurred.
- What is the frequency of DO topicalization?
- What is the frequency of IO topicalization?

3.

The Results

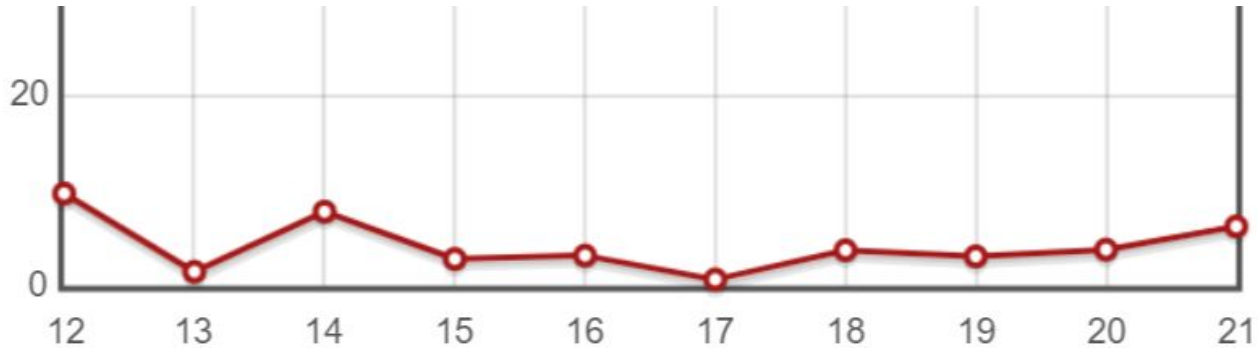
IcePaHC

- A total of 1100 instances of matrix-clause double-object constructions.
 - Of these, 89 were instances of DO topicalization (incidence: 8%).
 - 39 were instances of IO topicalization (incidence: 3,5%).
-
- Rate of topicalization with double-object constructions: 11%.
 - **DO topicalization twice as frequent** as IO topicalization.

IO vs DO topicalization



DO



IO

DO topicalization

12. Reykelsi færur vér honum ...
Incense(ACC) bring we he(DAT) ...
'We bring him incense'

(From Íslensk Hómilíubók,
late 12th/early 13th century)

- Average length of topicalized DOs: $\mu = 1,9$

IO topicalization

13. og [öllum ríkismönnum þeim sem þar voru]
and [all powerful-men those that there were](DAT)
gaf hann nokkura góða gjöf og sæmilega
gave he some good gift and respectable

'and he gave all the powerful men that were present some good and respectable gift'.

(From Finnboga Saga Ramma, 1330-1370)

- Only 6 instances of pronominal topics.
- Average length of topicalized IOs was 2.6 words, pointing to an effect of heaviness on fronting (Indriðadóttir & Ingason 2019).

4.

Analysis

Local/ Non-Local Asymmetry

- Asymmetry between topicalized and non-topicalized structures.
- While IO>DO is preferred when both objects appear VP-internally, the order DO>IO has a higher incidence the moment the two objects no longer appear in the same local domain.

Why is DO topicalization preferred?

- There is a general, cross-linguistic preference towards having topical constituents precede focal ones (Prince 1981).
- Given an **underlying**, base **IO<DO order** and a **penalty for** overt syntactic **movement**, the preference for **Topic < Focus** does predict the higher incidence of DO topicalization we see in the data.
- Optimality Theory (Prince & Smolensky 1991) : competing candidates are alternative realisations of the same input.

Topic first...

- **IO topic, DO focus**

Penalty for movement, as movement is costly (Chomsky 1995).

	Order	Topic < Focus	Stay
1	V < <u>IO</u> < DO		
2	V < DO < <u>IO</u>	*	*
3	<u>IO</u> < V < S < DO		*
4	DO < V < S < <u>IO</u>	*	*

Topic first...

- DO topic, IO focus

	Order	Topic < Focus	Stay
1	V < IO < <u>DO</u>	*	
2	V < <u>DO</u> < IO		*
3	IO < V < S < <u>DO</u>	*	*
4	<u>DO</u> < V < S < IO		*

- IO topicalization is the least preferred strategy.
- All other candidates violate some constraint, generating optionality (e.g. the $\frac{3}{4}$ signature (Bobaljik & Wurmbrand 2008), a Rank-Ordering model of EVAL (Coetzee 2004)).

The Heaviness Factor

Why does IO topicalization occur at all, then?

- Recall that the average length of topicalized IOs was 2.6 μ .
- Indriðadóttir & Ingason (2019): heaviness draws phrases to both edges of a clause, not simply the right edge.
- I & I (2009) searched the IcePaHC for:
 - i) Left-Dislocated Subjects, DOs & IOs,
 - ii) Topicalized DOs & IOs.

- *Icelandic Left Dislocation*

[Presturinn], María sá **[hann]** í bænum ígær.
the.priest Mary saw him downtown yesterday
'The priest, Mary saw him downtown yesterday.'

- They compared the average length of moved constituents vs. length of in-situ constituents.

The Heaviness Factor

- **Subjects**

In situ: $\mu = 2,1$. Left-Dislocated: $\mu = 9,6$.

- **DOs**

In situ: $\mu = 2,57$. Left-dislocated: $\mu = 8$. Topicalized: $\mu = 1,9$

- **IOs**

In situ: $\mu = 1,5$. Left-dislocated: $\mu = 8$. Topicalized: 2,6

- Left-dislocated constituents are much heavier than their in-situ counterparts.
- Topicalized IOs are longer than in-situ IOs, while topicalized DOs are shorter than in-situ DOs.

The Heaviness factor

S < V < IO < DO

- If heaviness does draw phrases to the edges of a clause, IOs are prime candidates for dislocation: their in-situ position is not an edge position.
- Heaviness: one possible trigger for the otherwise dispreferred IO topicalization structure.
- Other triggers: overt realization of contrastivity.

Conclusions

- **Asymmetry between local and non-local orders:** while the order IO < DO is preferred VP-internally, DO topicalization in ditransitive structures is twice as frequent as IO topicalization.
- We argue that the higher incidence of DO topicalization follows from the **cross-linguistic preference for having topical information first**, together with a penalty for overt movement operations (Stay).
- **Heaviness** one of the possible triggers for **IO topicalization**.

Thanks!

Elena: ecallegari@hi.is

Anton: antoni@hi.is

References



- Callegari, E. (2018). *Understanding Word Order in the Left Periphery*. PhD Thesis, University of Oslo.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Collins, C., & Thráinsson, H. (1996). VP-internal structure and object shift in Icelandic. *Linguistic inquiry*, 391-444.
- Dehé, N. (2004). On the order of objects in Icelandic double object constructions. *UCL working papers in linguistics*, 16, 85-108.
- Falk, C. (1990). On double object constructions. *Working Papers in Scandinavian Syntax*, 46, 53-100.
- Ingason, A. K. (2016). PaCQL: A new type of treebank search for the digital humanities. *Italian Journal of Computational Linguistics*, 2(2), 51-66.
- Indriðadóttir, I. H., & Ingason, A. K. (2019). Linguistic end-weight is really edge-weight. Observing heaviness in a parsed corpus. *DHN*, 240-249.
- Ottóson, K.G. (1991). Icelandic double objects as Small Clauses. *Working Papers in Scandinavian Syntax*, 48, 77-97.
- Thráinsson, H. (1994). Icelandic. In König, E. & van der Auwera, J. (eds.), *The Germanic Languages*. London/New York: Routledge. 142-189.
- Wallenberg, J. C., Ingason, A. K., Sigurðsson, E. F., & Rögnvaldsson, E. (2011). *Icelandic Parsed Historical Corpus (IcePaHC)*. Version 0.9. Available at http://www.linguist.is/icelandic_treebank