GLOSSA psycholinguistics

Nyvad, A. M., Müller, C., & Christensen, K. R. (2026). Moving away from the island: Extraction from adjunct clauses in Danish. *Glossa Psycholinguistics*, 23(1): 1, pp. 1–48. DOI: https://doi.org/10.5070/G6011.35249

eScholarship University of California

Moving away from the island: Extraction from adjunct clauses in Danish

Anne Mette Nyvad, Department of English, Aarhus University, Jens Chr. Skous Vej 4, 8000 Aarhus, Denmark, amn@cc.au.dk

Christiane Müller, Department of English, Aarhus University, Jens Chr. Skous Vej 4, 8000 Aarhus, Denmark, jane.mueller@web.de

Ken Ramshøj Christensen, Department of English, Aarhus University, Jens Chr. Skous Vej 4, 8000 Aarhus, Denmark, krc@cc.au.dk

This study explores island sensitivity in Danish. We present data from a series of acceptability judgment experiments on extraction from three types of finite adjunct clauses in Danish: relativization with and without supporting context, and topicalization with context. The results revealed a strikingly stable pattern: Extractions from temporal and causal clauses are significantly less acceptable than extractions from conditionals, which in turn are significantly less acceptable than extractions from complement clauses. The results indicate that Danish exhibits acceptability patterns similar to those found in Swedish and Norwegian regarding extraction from conditional and causal adjunct clauses. However, the raw scores of extraction from adjunct clauses are unexpectedly low in Danish compared to English. We argue that the acceptability ratings across constructions and languages vary as a function of extra-syntactic factors, including processing and discourse function, and that an account which is more finely grained than the ones currently offered is needed in order to capture the observed variability. This research contributes to a deeper understanding of island phenomena and highlights the need for investigation involving direct comparisons between languages.

1. Introduction

Finite adjunct clauses are typically considered to be strong islands across languages, banning all extraction (Cinque, 1990; Szabolcsi & Dikken, 2003). Example (1) demonstrates the impossibility of extracting *wh*-phrases from adjunct clauses in English:

(1) *[Who]$_i$ did Mary cry [after John hit $\_i$]?
(Huang, 1982, p. 503)

The island status of adjuncts has traditionally been captured by the Condition on Extraction Domain (CED, Huang, 1982, p. 505), universally banning extraction from domains that are not properly governed (roughly, subjects and adjuncts) (for a Minimalist account that distinguishes between subject and adjuncts, see Stepanov, 2007). However, Danish and the other Mainland Scandinavian (MSc.) languages have been argued to permit extraction from adjunct clauses, based on examples like (2) (and, indeed, such examples are common in Danish, cf. Müller & Eggers, 2022)[1]:

(2) a. [Det]$_i$ blev han smaddersur [fordi jeg sagde $\_i$].
that became he very-angry because I said
'He became very angry because I said that.'
(Jensen, 2002, p. 102)

b. [De sko]$_i$ kan jeg godt vente [mens du reparerer $\_i$].
those shoes can I well wait while you repair
'I can easily wait while you repair those shoes.'
(Hansen & Heltoft, 2011, p. 1814)

In (2), the DPs *det* 'that' (demonstrative pronoun) and *de sko* 'those shoes' have been felicitously topicalized from a finite adjunct clause introduced by *fordi* 'because' and *mens* 'while', respectively. Examples such as (2) appear to violate the CED and thus raise the question whether island constraints are subject to crosslinguistic variation after all (see also Rizzi, 1982; Stepanov, 2007). Getting a proper understanding of the unexpected behavior of MSc. regarding adjunct islands is important, as crosslinguistic variation in island effects poses a problem for all accounts of islands so far. Crosslinguistic variation suggests *possibility*, whereas at least traditional syntactic accounts (with inviolable constraints) would predict universal uniform *blocking* across languages.[2] If Universal Grammar (e.g. Chomsky, 2000;

[1] Likewise, in English, numerous counterexamples to the CED have been observed in the literature, e.g. the following example from Chaves (2013, p. 293) (adapted from Grosu, 1981):

(i) Th[is] is the house that Mary died [before she sold _].

[2] This is also reminiscent of Rizzi's (1982) original parametric account of crosslinguistic variation in subjacency and bounding nodes in English and Italian. However, subjacency cannot account for the differences between the complementizer that we observe in Danish here and in English (Nyvad et al., 2022), at least not without stipulating additional (potentially non-overt) bounding nodes.

Roberts, 2016) blocks island extraction, then no language should allow it. The data, however, says otherwise. If languages differ with respect to some putative universal, then that universal cannot be an absolute principle, but may be a parameter which can be set in a number of ways, and which may be sensitive to processing factors. This matters for theories about what *Language* (capital L) is and about how *languages* (plural) are and can be different (constraints on variation).

Recent studies of Norwegian (Bondevik et al., 2020) and Swedish (Müller, 2017) suggest a more nuanced picture, indicating that adjunct clause extraction in MSc. is not unconstrained and that extraction possibilities in these languages appear to vary across different types of adjunct clauses. Specifically, while topicalization from conditional adjunct clauses headed by *om* 'if' and from some temporal adjuncts was judged on the upper half of the scale by most participants in the studies of Norwegian (Bondevik et al., 2020) and Swedish (Müller, 2017), topicalization from causal adjuncts (clauses headed by *fordi* 'because' in Norwegian and *eftersom* 'because' in Swedish) received ratings on the low end of the scale, while also exhibiting an interaction (or superadditive) effect in comparison to the non-extracted version. Similarly, Kobzeva et al. (2022) found an island effect for *wh*-movement out of *if*-clauses, but not for relativization. The question is whether Danish also patterns with Swedish and Norwegian in showing variable sensitivity to island constraints across different adjunct clause types. Preliminary evidence for variation in extractability from different types of Danish adjuncts comes from a recent corpus study (Müller & Eggers, 2022). While the corpus search in this study returned a range of examples displaying extraction from adjunct clauses headed by *hvis* ('if') and *når* ('when'), no cases of extraction from clauses headed by *fordi* ('because') were found. This result matches the previous observations from Swedish and Norwegian that conditional and (certain) temporal adjuncts appear to allow extraction more freely than causal adjuncts. However, additional data from acceptability studies in Danish are needed to confirm the island sensitivity of different adjunct clause types in Danish.

The finding that adjunct clause extraction in the MSc. languages is not unrestricted raises the question of how different these languages really are from English in this regard. Recent studies suggest that English might be more similar to MSc. than previously assumed, in that different adjunct clause types display a non-uniform behavior when it comes to extraction. Nyvad et al. (2022) compared contextually facilitated relativization from *if*-, *when*-, and *because*-clauses in English and found that the three adjunct clause types showed a rather heterogeneous pattern. First, extraction from *if*-clauses resulted in significantly higher ratings than extraction from *when*-clauses, which, in turn, resulted in higher ratings than extraction from *because*-clauses (a similar gradience was found by Chaves & Putnam, 2020, pp. 235–239, for English). Second, extraction from all three clause types showed somewhat different acceptability distributions. The finding that extraction from *if*-clauses yielded ratings on the upper half of the scale and showed no significant difference from non-island *that*-clause extraction suggests that at least

for relativization, *if*-adjuncts are not invariably strong islands in English. This picture is further corroborated by data from Sprouse et al. (2016) showing that relativization from *if*-adjunct clauses did not result in an island effect (under their factorial definition of an island effect and contra to their findings for *wh*-questions), although the ratings for relativization from *if*-clauses remained relatively low in Sprouse et al.'s study.

However, a more direct comparison of the island constraints in MSc. languages and English has so far been difficult, because previous studies did not use matching conditions. For instance, while many recent acceptability studies of islands in MSc. tend to focus on topicalization (Bondevik et al., 2020; Kush et al., 2019; Müller, 2017), most existing studies of islands in English solely investigate extraction by *wh*-movement (Michel & Goodall, 2013; Müller, 2019; Sprouse, 2007; Sprouse et al., 2012). Moreover, many studies of island extractions in the MSc. languages make an explicit effort to use stimulus sentences that are as natural-sounding as possible, by modeling the sentences after authentic examples reported in the literature (Kush et al., 2018, p. 755, 2019, p. 399). By contrast, many of the existing studies of islands in English suffer from using stimuli that are pragmatically odd or unnatural even without the extraction, because they describe very implausible scenarios (see Chaves & Putnam, 2020, for the same observation and examples). These differences prevent a fair and direct comparison of the relative sensitivity to island constraints between English and the MSc. languages.

An acceptability judgment study testing extraction in the form of relativization from three different types of finite adjunct clauses in English (*if*-, *when-,* and *because*-clauses) in the presence of supporting context has previously been conducted (Nyvad et al., 2022). Here, we present a follow-up study that investigates Danish adjunct clauses under comparable conditions. The goal of this study is twofold:

I. The first goal is to investigate whether the acceptability of adjunct clause extraction in Danish varies across different types of adjunct clauses. This would be evidence that Danish patterns with Swedish and Norwegian in showing variable acceptability patterns across different adjunct clauses.
II. The second goal is to compare the island sensitivity of different adjunct clauses between Danish and English. A comparison is possible between the current study and the results from Nyvad et al.'s (2022) experiment, because both studies investigate extraction using the same experimental design and comparable materials.

In what follows, we present the results of three experiments testing the acceptability of extraction from three different types of adjunct clauses. In the first experiment, the filler-gap dependency is constructed by way of relativization and facilitated by a context, whereas the context was removed in second experiment. The third experiment changed the dependency type from relativization to topicalization.

2. Methods

2.1 Participants

A total of 335 native speakers of Danish (aged 17 to 72 years, mean = 29.5, SD = 7.5) volunteered to participate in the experiment. The number of participants per list ranged from 31 to 52: list 1 = 41, list 2 = 45, list 3 = 41, list 4 = 52, list 5 = 43, list 6 = 32, list 7 = 50, list 8 = 31).

2.2 Materials

To facilitate extraction as much as possible, we used relativization in the presence of a context in both studies. The presence of contextual cues has been suggested to be necessary for the felicity of at least some island extractions (Bondevik et al., 2020; Kush et al., 2018; Müller, 2019; Tutunjian et al., 2017; Wiklund et al., 2017). Furthermore, Kush et al. (2019) provide experimental evidence for the beneficial effect of context on acceptability ratings for at least some adjunct clause extractions.

We chose to test extraction in the form of relativization rather than *wh*-extraction, since *wh*-extraction from at least some islands is regularly reported to be less acceptable or less common than other forms of extraction (i.e. topicalization or relativization) in both English and the MSc. languages (Abeillé et al., 2020; Bondevik et al., 2020; Kush et al., 2018, 2019; Lindahl, 2017; Müller & Eggers, 2022; Sprouse et al., 2016). *Wh*-extraction would therefore not create the optimal conditions for felicitous extraction. In turn, extraction in the form of topicalization would have been problematic for the English experiment presented in Nyvad et al. (2022), as topicalization in English is less frequently used and is generally considered to be fairly marked compared to the MSc. languages (e.g. Engdahl, 1997; Poole, 2017, p. 15). Relativization from adjunct clauses, on the other hand, allows for a fair comparison of English and Danish because it appears to be easier to accept (and to occur more frequently) than *wh*-extraction, both in English (Abeillé et al., 2020; Sprouse et al., 2016) and in Danish (Müller & Eggers, 2022).

Following Nyvad et al. (2022), we used a 2×4 factorial design with the two factors, namely, a binary Extraction factor and a Complementizer factor with four different levels: *at* 'that' (introducing declarative complement clauses, which are typically non-island clauses), *hvis* 'if', *da* 'when', and *fordi* 'because' (all introducing adjunct clauses). This yielded the eight target types listed in **Table 1**.

Table 1: Experimental design.

	Complementizer	**[–Extraction]**	**[+Extraction]**
Non-island	*At* 'that'	Type 1	Type 5
Island	*Hvis* 'if'	Type 2	Type 6
	Da 'when'	Type 3	Type 7
	Fordi 'because'	Type 4	Type 8

The stimulus set was a Danish translation of the English stimulus used in Nyvad et al. (2022), slightly modified whenever direct translation was not possible. The Danish [+Extraction] conditions are *it*-clefts (or rather, *det*-clefts), not relative clauses as in the English study; we will come back to this in 5.2.4. A sample set is shown in (3).

(3) TARGET. CONTEXT:
I det sidste træningsprogram jeg udarbejdede for Emma, ville jeg gerne gøre det så godt som umuligt for hende og inkluderede derfor endnu et sæt virkelig brutale pull-ups.
'In the latest workout routine I designed for Emma, I really wanted to make it impossible for her and included another set of particularly brutal pull-ups.'

NON-ISLAND STRUCTURE, [–EXTRACTION]:

a. Det er åbenlyst at jeg blev overrasket over, **at** hun faktisk gennemførte
it is obvious that I became surprised over **that** she actually completed
dét program.
that exercise
'It's obvious that I was surprised **that** she actually completed this exercise.'

ISLAND STRUCTURE, [–EXTRACTION]:

b. Det er åbenlyst at jeg ville blive overrasket, **hvis** hun faktisk
it is obvious that I would become surprised **if** she actually
gennemførte dét program.
completed that exercise
'It's obvious that I would be surprised **if** she actually completed this exercise.'

c. Det er åbenlyst at jeg blev overrasket, **da** hun faktisk gennemførte dét program.
'It's obvious that I was surprised **when** she actually completed that exercise.'

d. Det er åbenlyst at jeg blev overrasket, **fordi** hun faktisk gennemførte dét program.
'It's obvious that I was surprised **because** she actually completed that exercise.'

NON-ISLAND STRUCTURE, [+EXTRACTION]:

e. Det er <u>dét program</u> som jeg blev overrasket over, **at** hun faktisk
it is that exercise that I became surprised over **that** she actually
gennemførte _.
completed
'This is the exercise that I was surprised **that** she actually completed.'

ISLAND STRUCTURE, [+EXTRACTION]:

f. Det er <u>dét program</u> som jeg ville blive overrasket, **hvis** hun faktisk
it is that exercise that I would become surprised **if** she actually
gennemførte _.
completed
'This is the exercise that I would be surprised **if** she actually completed.'

g. Det er dét program som jeg blev overrasket, **da** hun faktisk gennemførte _.
'This is the exercise that I was surprised **when** she actually completed.'

h. Det er dét program som jeg blev overrasket, **fordi** hun faktisk gennemførte _.
'This is the exercise that I was surprised **because** she actually completed.'

As shown in (3), each target sentence was preceded by a facilitating context. Participants received the instruction to base their judgment only on the sentence that followed the context. For the reasons mentioned in Section 1, we chose to test extraction in the form of relativization. Since relativization does not have an unextracted counterpart which can be used for comparison between the [+Extraction] and [–Extraction] sentences, we constructed our [–Extraction] baseline sentences such that they were all embedded in a *Det er åbenlyst at…* 'It's obvious that' or *Det er klart at…* 'It's clear that' construction, parallel to the English baseline sentences in Nyvad et al. (2022). Our stimuli thus deviated slightly from a completely strict factorial design, and this has consequences for the cross-condition comparisons. Even though the main matrix clauses of our [+Extraction] and [–Extraction] were not minimally different, the *It's obvious / clear that…* constructions used in our baseline sentences were of comparable length and complexity as the relativization structure used in the [+Extraction] sentences, and the target sentences were matched on the parameters that we have reason to suspect might play a role for extractability.[3]

The non-island *that*-clauses and the three different types of adjuncts were embedded under adjectival psych-predicates, such as *overrasket* 'surprised' or *glad* 'happy', which can be felicitously combined with *that*-clauses as well as with all three types of adjunct clauses tested. In the conditions involving the complementizer *hvis* 'if' (3b,f), the matrix structure had to be in the future past tense, instead of the simple past tense, to be felicitous. To facilitate extraction, we constructed all the stimuli with adjectival psych-predicates to trigger causal, and thus, semantically coherent readings of the matrix and adjunct event (Müller, 2019; Tanaka, 2015; Truswell, 2011).

To ensure the optimal conditions for extraction, following Nyvad et al. (2022), we made sure that all the adjunct clauses tested in our experiment would be classified as central (rather than

[3] Erteschik-Shir and Lappin argue that any matrix clause which can be interpreted as serving the function of introducing into the sentence the head of the relative clause should allow extraction to the front of the clause more easily. Erteschik-Shir and Lappin (1979, p. 57) furthermore argue that, in addition, "the head of the relative clause must be indefinite (as with the existential operator) and the subject of the matrix must be first person." Otherwise, "[h]aving anything but a first-person pronoun as the subject of the matrix also blocks the dominant interpretation of the relative clause and hence extraction" (footnote 17). However, we do not agree with their intuition. For example, we do not think that there is any difference in acceptability between *Den historie kender **jeg** ingen der tror på* 'That story I don't know anyone who believes" and *Den historie kender **hun** ingen der tror på* "That story, she doesn't know anyone who believes".

peripheral) according to Haegeman's (2012, pp. 161–164) definition, given that they provide information about the condition for ('if'), the time ('when'), or the cause ('because') of the state or event in the matrix clause (e.g. being or becoming surprised in (3)). All things being equal, extraction from central adjunct clauses should be more easily acceptable than from peripheral ones.

Twenty-four items of the type shown in (3) were created and distributed across eight lists in a Latin Square design, such that each participant saw three instances of each type, but only one sentence from each item set (scenario). The order in which the sentences were presented was randomized on each list. As in Nyvad et al.'s (2022) study of English adjunct clauses, two types of controls were added. Control 1 involved extraction from NP subjects and Control 2 involved extraction from coordinate structures. These structures served additionally as points of comparison, as both subjects and coordinate structures are also assumed to be strong islands in both English and Danish. Eight sets of control type 1 and eight sets of control type 2 were constructed, once again by translating the English sentence items used in Nyvad et al. (2022) into Danish. As demonstrated in the sample sets in (4) (Control 1, involving subject islands with preposition stranding) and (5) (Control 2, involving coordinate structure islands), the NP subjects and the coordinate structures were embedded in the same four clause types used in the target items (non-island *at*-clauses, and adjunct clauses introduced by *hvis* 'if', *da* 'when', and *fordi* 'because'), each of which had an extracted and unextracted version. This yielded eight control conditions for each control type that could be used for comparison with our target items. We chose to keep the number of controls to an absolute minimum to avoid experimental fatigue. Subject islands and, in particular, coordinate structure violations were used as the benchmark for unacceptability. The controls were distributed across the eight lists in a Latin Square design, such that each participant saw two instances of each type. Each list thus contained 40 sentences (24 targets and 16 controls).

(4) CONTROL 1. CONTEXT:
Mit team har udviklet en COVID-19-vaccine på rekordtid, og jeg synes, at vi fortjener en vis anerkendelse.
'My team has developed a COVID-19 vaccine in record time, and I think we deserve some recognition.'

SUBJECT ISLAND + NON-ISLAND THAT-CLAUSE, [–EXTRACTION]:

a. Det er klart at vi blev glade for, **at** vores vaccine mod dén virus
it is clear that we became glad for **that** our vaccine against this virus
endelig fik Nobelprisen.
finally got Nobel Prize.the
'It's clear that we were pleased **that** our vaccine against this virus finally got the Nobel Prize.'

SUBJECT ISLAND + ADJUNCT ISLAND, [–EXTRACTION]:

b. Det er klart at vi ville blive glade, **hvis** vores vaccine mod dén virus
it is clear that we would become glad **if** our vaccine against this virus
endelig fik Nobelprisen.
finally got Nobel Prize.the
'It's clear that we would be pleased **if** our vaccine against this virus finally got the Nobel Prize.'

c. Det er klart at vi blev glade, **da** vores vaccine mod dén virus endelig fik Nobelprisen.
'It's clear that we were pleased **when** our vaccine against this virus finally got the Nobel Prize.'

d. Det er klart at vi blev glade, **fordi** vores vaccine mod dén virus endelig fik Nobelprisen.
'It's clear that we were pleased **because** our vaccine against this virus finally got the Nobel Prize.'

SUBJECT ISLAND + NON-ISLAND THAT-CLAUSE, [+EXTRACTION]:

e. Det er dén virus som vi blev glade for, **at** vores vaccine mod _ endelig
It is that virus that we became glad for **that** our vaccine against finally
fik Nobelprisen.
got Nobel Prize.the
'This is the virus that we were pleased **that** our vaccine against finally got the Nobel Prize.'

SUBJECT ISLAND + ADJUNCT ISLAND, [+EXTRACTION]:

f. Det er dén virus som vi ville blive glade, **hvis** vores vaccine mod _
It is that virus that we would become glad **if** our vaccine against
endelig fik Nobelprisen.
finally got Nobel Prize.the
'This is the virus that we would be pleased **if** our vaccine against finally got the Nobel Prize.'

g. Det er dén virus som vi blev glade, **da** vores vaccine mod _ endelig fik Nobelprisen.
'This is the virus that we were pleased **when** our vaccine against finally got the Nobel Prize.'

h. Det er dén virus som vi blev glade, **fordi** vores vaccine mod _ endelig fik Nobelprisen.
'This is the virus that we were pleased **because** our vaccine against finally got the Nobel Prize.'

(5) CONTROL 2. CONTEXT:

Jeg havde lovet en ven at passe hans kæledyr i en uge. Ved en fejl lod jeg uheldigvis hoveddøren stå åben et sekund for længe og både hans kat og hans prisvindende udstillingshund løb ud.

'I had promised a friend to watch his pets for a week. Unfortunately, I accidentally left the front door open just a bit too long and both his cat and his prize-winning show dog ran out.'

Coordinate structure island + non-island that-clause, [–Extraction]:

a. Det er åbenlyst at jeg blev flov over, **at** jeg faktisk mistede både
it is obvious that I became ashamed over **that** I actually lost both
katten og hunden samme dag.
cat.the and dog.the same day
'It's obvious that I was ashamed **that** I actually lost both the cat and this dog in one day.'

Coordinate structure island + adjunct island, [–Extraction]:

b. Det er åbenlyst at jeg ville blive flov, **hvis** jeg faktisk mistede både
it is obvious that I would become ashamed **if** I actually lost both
katten og hunden samme dag.
cat.the and dog.the same day
'It's obvious that I would be ashamed **if** I actually lost both the cat and this dog in one day.'

c. Det er åbenlyst at jeg blev flov, **da** jeg faktisk mistede både katten og hunden samme dag.
'It's obvious that I was ashamed **when** I actually lost both the cat and this dog in one day.'

d. Det er åbenlyst at jeg blev flov, **fordi** jeg faktisk mistede både katten og hunden samme dag.
'It's obvious that I was ashamed **because** I actually lost both the cat and this dog in one day.'

Coordinate structure island + non-island that-clause, [+Extraction]:

e. Det er <u>hunden</u> som jeg blev flov over, **at** jeg faktisk mistede både
it is dog.the that I became ashamed over **that** I actually lost both
katten og _ samme dag.
cat.the and same day
'This is the dog that I was ashamed **that** I actually lost both the cat and in one day.'

Coordinate structure island + adjunct island, [+Extraction]:

f. Det er <u>hunden</u> som jeg ville blive flov, **hvis** jeg faktisk mistede både
it is dog.the that I would become ashamed **if** I actually lost both
katten og _ samme dag.
cat.the and same day
'This is the dog that I would be ashamed **if** I actually lost both the cat and in one day.'

g. Det er <u>hunden</u> som jeg blev flov, **da** jeg faktisk mistede både katten og _ samme dag.
'This is the dog that I was ashamed **when** I actually lost both the cat and in one day.'

h. Det er <u>hunden</u> som jeg blev flov, **fordi** jeg faktisk mistede både katten og _ samme dag.
'This is the dog that I was ashamed **because** I actually lost both the cat and in one day.'

2.3 Procedure

The acceptability survey was disseminated in the form of an online questionnaire using Google Forms. Participants were recruited through social media platforms and were assigned in a pseudo-random fashion (based on their birth month) to one of the eight presentation lists. They rated the sentences on a seven-point Likert scale ranging from 1 = *helt uacceptabel* 'completely unacceptable' to 7 = *helt acceptabel* 'completely acceptable'.

3. Results

Before we analyze the z-transformed (standardized) data, we first analyze the raw data (7-point Likert scale), which shows significant effects in the experiment using the actual scale in the acceptability judgments. The mean acceptability scores for both target and control conditions are presented in **Table 2**. The z-transformed data allows for comparisons across experiments and other studies.

Table 2: Mean acceptability scores. [±Ex] = ±Extraction.

		Target sentences		**Control 1**		**Control 2**	
		Acceptability	**SD**	**Acceptability**	**SD**	**Acceptability**	**SD**
at ('that')	[–Ex]	6.53	0.97	6.32	1.15	6.59	0.90
hvis ('if')	[–Ex]	6.56	0.93	6.36	1.05	6.48	1.00
da ('when')	[–Ex]	6.46	1.08	6.27	1.30	6.61	0.83
fordi ('because')	[–Ex]	6.05	1.51	6.17	1.22	6.35	1.14
at ('that')	[+Ex]	6.19	1.25	2.62	1.68	1.55	1.13
hvis ('if')	[+Ex]	3.45	1.86	2.20	1.33	1.45	0.94
da ('when')	[+Ex]	2.77	1.65	1.75	1.07	1.36	0.77
fordi ('because')	[+Ex]	2.79	1.66	1.71	1.05	1.41	0.86

The data was analyzed using *R* version 4.1.2 (R Core Team, 2021). The target conditions (extraction from adjunct clauses), Control 1 conditions (extraction from embedded subjects), and Control 2 conditions (coordinate structure violation) were analyzed using separate mixed-effects models. (We had no a priori hypotheses about statistical differences between the three; furthermore, as the three [target, Control 1, Control 2] were not minimally different, direct comparison across types would be inappropriate.) The models in the analyses are the largest models that converged without errors. All plots were made in R using *ggplot2* (Wickham, 2016) and *gridExtra* (Augie, 2016).

3.1 Target: Adjunct islands

To test for significant positive interactions between extraction and complementizer, which would indicate significant island effects, we applied a cumulative link mixed-effects model (clmm), using the

ordinal package for *R* (Christensen, 2015), with acceptability as outcome variable and the interaction between extraction and complementizer as predictor, and random intercepts for participant and item, and random slopes for extraction by participant. As the summary in **Table 3** shows, the interactions were significant ($p < 0.001$), as was the main effect of Extraction ($p < 0.001$).

Table 3: Summary of the cumulative link mixed-effects model for interaction between Extraction and Complementizer on the raw ratings of the **target sentences (adjunct islands)**: clmm(Acceptability ~ Complementizer * Extraction + (1 + Extraction | Participant) + (1 | Item). Significance indicator: *** $p < 0.001$.

	Estimate	**SE**	**z**	**p**	
if	0.110	0.203	0.544	0.587	
when	–0.188	0.201	–0.933	0.351	
because	–0.990	0.199	–4.973	0.000	***
Extraction	–0.845	0.199	–4.248	0.000	***
if × Extraction	–3.871	0.283	–13.662	0.000	***
when × Extraction	–4.367	0.283	–15.411	0.000	***
because × Extraction	–3.588	0.281	–12.795	0.000	***

For pairwise comparisons of the eight target types, we used the *emmeans* package for *R* (Lenth, 2022) on a cumulative mixed-effects model with acceptability as outcome variable and type as predictor and random intercepts for participant and item. The significance levels of the contrasts, Bonferroni corrected for multiple comparisons, are shown in the contrast matrix in **Table 4**.

Table 4: Pairwise contrasts between the eight types for **targets (adjunct islands)**. clmm(Acceptability ~ Type + (1 | Participant) + (1 | Item), Bonferroni corrected for 28 tests. Significance indicators: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, n.s. = not significant.

2. *if* [–Ex]	**3. *when* [–Ex]**	**4. *because* [–Ex]**	**5. *that* [+Ex]**	**6. *if* [+Ex]**	**7. *when* [+Ex]**	**8. *because* [+Ex]**	
n.s.	n.s.	***	***	***	***	***	1. *that* [–Ex]
	n.s.	***	***	***	***	***	2. *if* [–Ex]
		**	*	***	***	***	3. *when* [–Ex]
			n.s.	***	***	***	4. *because* [–Ex]
				***	***	***	5. *that* [+Ex]
					***	***	6. *if* [+Ex]
						n.s.	7. *when* [+Ex]

The results and adjacent pairwise comparisons for the eight target sentence types are also shown in **Figure 1 (A)**. The acceptability ratings for the four baseline conditions [–Ex] all had a rating above 6 on the 7-point scale. For the three adjunct types (*that*, *if*, and *because*), there was

a significant drop in acceptability between the [–Ex] and [+Ex] types, indicating a significant negative main effect of extraction (cf. also the model in **Table 3**), [+Ex] > [–Ex] ($p < 0.001$, see also **Figure 1**), and extraction from *if*-clauses was more acceptable than extraction from *when*-clauses ($p < 0.001$). The effect of extraction from complement clauses, *that* [+Ex], was not significantly different from baseline *because*-clauses. However, though the difference between *that* [–Ex] and *that* [+Ex] was small (0.34 points on the 7-point scale), it was indeed significant ($p < 0.001$), showing that the cost of extraction is not confined to adjunct clauses. This model also showed a significant difference between *that* [–Ex] and *because* [–Ex] ($p < 0.001$), while the differences between *that* [–Ex] and *if/when* [–Ex] was not significant.

3.2 Control 1: Subject islands

For the subject islands, we used the same approach as with the target condition, and used *emmeans* on a model with acceptability as outcome variable and type as predictor and random intercepts for participant and item. The pairwise contrasts are summarized in **Table 5**.

Table 5: Pairwise contrasts for **Control 1 (subject islands)**. clmm(Acceptability ~ Type + (1 | Participant) + (1 | Item), Bonferroni corrected for 28 tests. Significance indicators: *** $p < 0.001$, * $p < 0.05$, n.s. = not significant.

2. *if* [–Ex]	**3. *when* [–Ex]**	**4. *because* [–Ex]**	**5. *that* [+Ex]**	**6. *if* [+Ex]**	**7. *when* [+Ex]**	**8. *because* [+Ex]**	
n.s.	n.s.	n.s	***	***	***	***	1. *that* [–Ex]
	n.s.	n.s.	***	***	***	***	2. *if* [–Ex]
		n.s.	***	***	***	***	3. *when* [–Ex]
			***	***	***	***	4. *because* [–Ex]
				n.s.	***	***	5. *that* [+Ex]
					n.s.	*	6. *if* [+Ex]
						n.s.	7. *when* [+Ex]

The results and adjacent pairwise comparisons for the eight sentence types are also shown in **Figure 1 (C)**. There were no significant differences between the four baseline [–Ex] types, which were all rated above 6 on the 7-point scale. In contrast, the four [+Ex] sentence types were all rated very low, and they were all significantly different from the [–Ex] types. In other words, there was a clear and significant negative main effect of extraction. There was also a significant difference between *that* [+Ex] and *when* and *because* [+Ex], while *if* [+Ex] did not differ significantly from *that* [+Ex] and *when* [+Ex].

3.3 Control 2: Coordinate structures

The converging model was identical to the one used for Control 1, with acceptability as outcome variable, type as predictor and random intercepts for participant and item. The pairwise contrasts are summarized in **Table 6**.

Table 6: Pairwise contrasts for **Control 2 (coordinate structures)**. clmm(Acceptability ~ Type + (1 | Participant) + (1 | Item), Bonferroni corrected for 28 tests. Significance indicators: *** $p < 0.001$, n.s. = not significant.

2. *if* [–Ex]	**3. *when* [–Ex]**	**4. *because* [–Ex]**	**5. *that* [+Ex]**	**6. *if* [+Ex]**	**7. *when* [+Ex]**	**8. *because* [+Ex]**	
n.s.	n.s.	n.s	***	***	***	***	1. *that* [–Ex]
	n.s.	n.s.	***	***	***	***	2. *if* [–Ex]
		n.s.	***	***	***	***	3. *when* [–Ex]
			***	***	***	***	4. *because* [–Ex]
				n.s.	n.s.	n.s.	5. *that* [+Ex]
					n.s.	n.s.	6. *if* [+Ex]
						n.s.	7. *when* [+Ex]

See also **Figure 1 (C)** for the results and adjacent pairwise comparisons for the eight sentence types. As with Control 1, the acceptability ratings for the four baseline conditions [–Ex] were all in the top range (all > 6), whereas the four [+Ex] sentence types were very low (<1.6), indeed, lower than all the other control and target conditions. The differences between the baseline conditions and the [+Ex] conditions were all significant ($p < 0.001$).

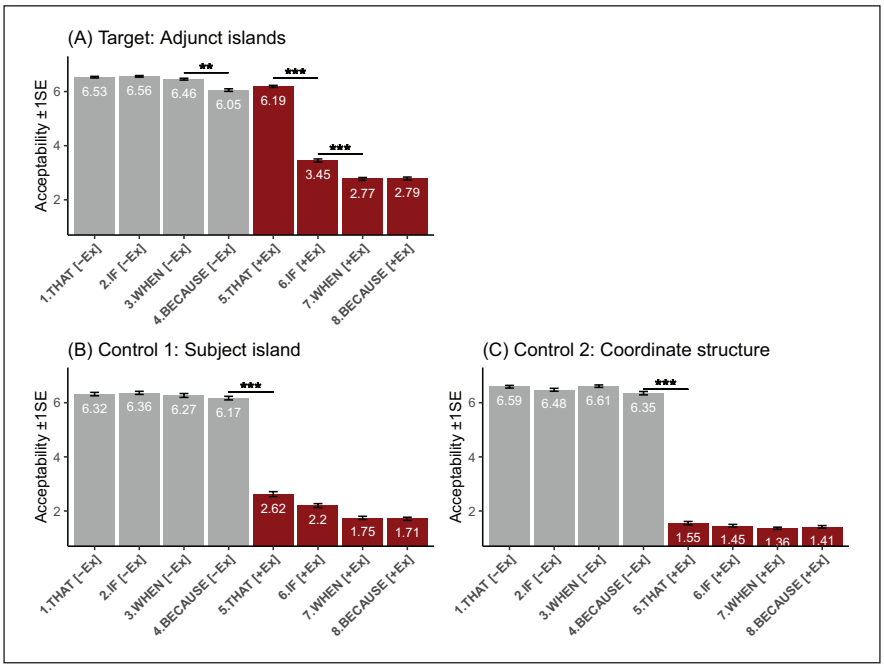

Figure 1: Mean (raw) acceptability ratings. *** $p < 0.001$, * $p < 0.05$.

To check that the general acceptability pattern in the target condition (**Figure 1 (A)**) was consistent across all the 24 sets/contexts, we plotted acceptability by type by participant for each set, as shown in **Figure 2**. As predicted, the overall patterns (i.e. the fitted curves) look quite similar.

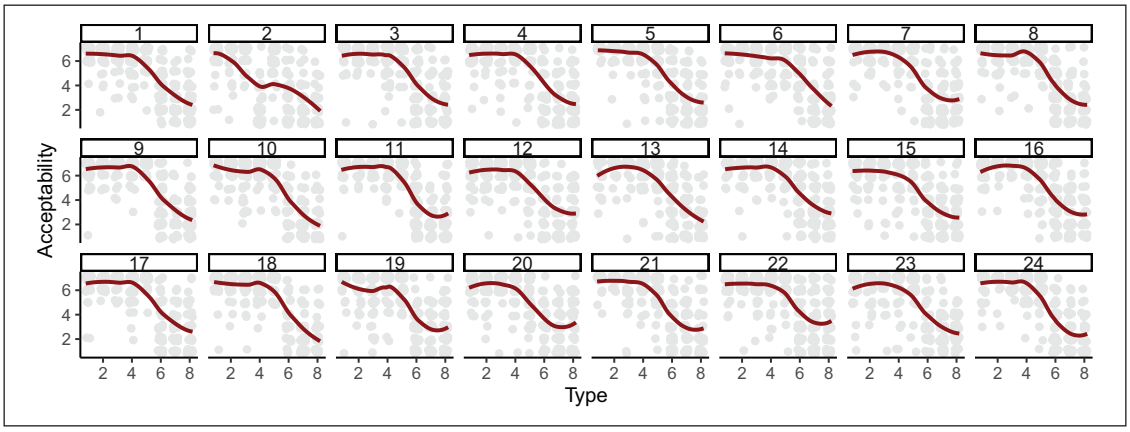

Figure 2: Mean (raw) acceptability ratings by type (see **Figure 1**) for each of the 24 target sets.

3.4 Standardized (z-transformed) acceptability ratings

Following Sprouse et al. (2012, 2016), we z-score transformed the ratings by participant (a) to control for potential individual scale bias, such as using only one end of the scale or a larger or smaller range (Bondevik et al., 2020; see also Kush et al., 2018, 2019), and (b) to make direct comparisons with a number of other studies easier, in particular, Nyvad et al. (2022). A participant's z-transformed rating represents the number of standard deviations the raw (non-transformed) rating is from that participant's mean rating. The z-transformed results were analyzed using a linear mixed-effects model with the *lmerTest Package* (Kuznetsova et al., 2017); the model has type as predictor and random intercepts for participant and item.[4] The model converged, although with a boundary (singular) fit. Again, we used the *emmeans* package (Lenth, 2022) to calculate the paired contrasts, Bonferroni corrected for multiple comparisons. The results are summarized in **Table 7** and plotted in **Figure 3 (A)**. Note that the overall acceptability pattern and significant contrasts are the same as with the non-transformed ratings (compare **Figure 1 (A)** and **Figure 3 (A)**).

[4] Here, none of the models with different combinations of random slopes converged without errors or warnings. The simplest model converged with a boundary fit warning.

Table 7: Pairwise contrasts between the eight types for for the **z-transformed ratings of the target sentences (adjunct islands)**: lmer(Acceptability_z ~ Type + (1 | Participant) + (1 | Item), Bonferroni corrected for 28 tests. Significance indicators: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, n.s. = not significant.

2. *if* [–Ex]	**3. *when* [–Ex]**	**4. *because* [–Ex]**	**5. *that* [+Ex]**	**6. *if* [+Ex]**	**7. *when* [+Ex]**	**8. *because* [+Ex]**	
n.s.	n.s.	**	n.s.	***	***	***	1. *that* [–Ex]
	n.s.	***	*	***	***	***	2. *if* [–Ex]
		*	n.s.	***	***	***	3. *when* [–Ex]
			n.s.	***	***	***	4. *because* [–Ex]
				***	***	***	5. *that* [+Ex]
					***	***	6. *if* [+Ex]
						n.s.	7. *when* [+Ex]

According to Sprouse, Wagers, and Phillips (2012), an island violation should lead to a superadditive effect of complementizer and extraction (corresponding to a positive interaction effect), as reflected in the differences-in-differences (DD) score (Bondevik et al., 2020; Kush et al., 2018, 2019; Nyvad et al., 2022; see also Sprouse et al., 2012). The DD score is calculated for each [+Ex] condition by subtracting the D score for *that* (the difference in rating between *that* [–Ex] and *that* [+Ex], for the non-island type) from the D scores for the three adjunct island types: *if*, *when*, and *because*, see **Figure 3 (C)** and **(D)**. Because the D scores are based on z-transformed ratings, the DD score provides a standardized measure of the island effect size, which enables comparisons between different island types across experiments and languages. According to Kush et al. (2019, p. 401), a DD score above 0.75 is indicative of an island violation, whereas a lower DD score "bears closer scrutiny". The DD scores for extraction from adjunct clauses headed by *if*, *when*, and *because* were 1.15, 1.39, and 1.21, respectively, all well above the assumed threshold for islandhood (**Figure 3 (D)**).

To test for significance, we analyzed the z-transformed data with a linear mixed-effects model with fixed effects of Complementizer and Extraction and their interaction, and random intercepts for participant and item, and random slopes for Complementizer and Extraction by participant. The model converged, albeit with a boundary (singular) fit. We first used *that* as intercept (summarized in **Table 8**), then we used *when* as intercept (summarized in **Table 9**). Note that the results (the significant contrasts) are the same as in **Table 4**: a significant main effect of Extraction and significant interactions between Extraction and Complementizer (island effects).

Table 8: Summary of the mixed-effects model of Complementizer and Extraction (**z-transformed ratings**, **target sentences (adjunct islands)**), using *that* as intercept: lmer(Acceptability_z ~ Complementizer * Extraction + (1 + Complementizer + Extraction | Participant) + (1 | Item). *** p < 0.001, ** p < 0.01.

	estimate	SE	df	t	p	
(Intercept: *that* [–Ex])	0.781	0.038	225.334	20.474	0.000	***
if [–Ex]	0.018	0.053	206.193	0.337	0.736	
when [–Ex]	–0.034	0.052	197.140	–0.658	0.511	
because [–Ex]	–0.196	0.052	200.984	–3.754	0.000	***
Extraction	–0.144	0.053	220.597	–2.695	0.008	**
if × Extraction	–1.155	0.072	186.010	–15.981	0.000	***
when × Extraction	–1.393	0.072	186.010	–19.275	0.000	***
because × Extraction	–1.237	0.072	186.010	–17.113	0.000	***

Table 9: Summary of the mixed-effects model of Complementizer and Extraction (**z-transformed ratings**, **target sentences (adjunct islands)**), using *when* as intercept: lmer(Acceptability_z ~ Complementizer * Extraction + (1 + Complementizer + Extraction | Participant) + (1 | Item). *** p < 0.001, ** p < 0.01, * p < 0.05.

	estimate	SE	df	t	p	
(Intercept: *when* [–Ex])	0.747	0.036	186.057	20.666	0.000	***
if [–Ex]	0.052	0.051	186.599	1.010	0.314	
because [–Ex]	–0.162	0.051	190.553	–3.151	0.002	**
that [–Ex]	0.034	0.052	204.625	0.648	0.518	
Extraction	–1.537	0.053	215.847	–28.937	0.000	***
if × Extraction	0.238	0.072	186.057	3.294	0.001	**
because × Extraction	0.156	0.072	186.057	2.163	0.032	*
that × Extraction	1.393	0.072	186.057	19.278	0.000	***

Figure 3 shows the results of the analysis of the z-transformed acceptability ratings for the target condition: (A) shows the mean acceptability ratings (compare **Figure 1**), (B) the extraction effects for the four complementizers, (C) the extraction effect sizes (D scores, all significant, p < 001, and all significantly different from each other, p < 0.001, except for *if* versus *because*), and (D) shows the DD scores indicating island effects, significant (p < 0.001). Running the model a third time but with *if* as intercept showed that the DD scores for *if* and *because* are not

significantly different from each other ($p > 0.261$). However, as is shown in **Figure 3(A)** and **(B)**, *because*-clauses without extraction are rated significantly lower than the other baselines, which explains why the DD-score is lower and why there is no significant difference between the DD scores for *if* and *because*.

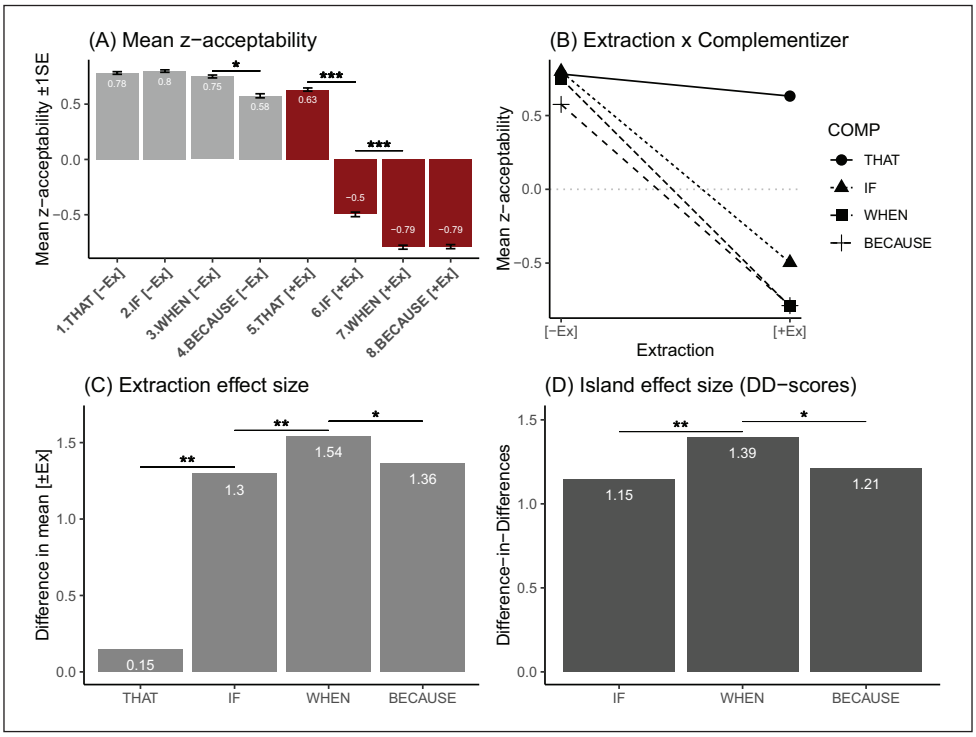

Figure 3: (A) Mean z-transformed acceptability ratings. (B) Extraction × Complementizer interaction. (C) Extraction effect size (difference between [+Ex] and [–Ex]). (D) Island effect size measured in DD scores (Complementizer x Extraction interaction) between *that* and the other three complementizers). *** $p < 0.001$, * $p < 0.05$.

To examine the inter-participant variation in the judgment pattern for each of the mean scores shown in **Figure 1**, we plotted the distribution of acceptability ratings by condition; see **Figure 4**. Uniform syntactic islands should show a unimodal distribution narrowly centered around $z = -1$, reflecting a very low consistent rating across participants, whereas unambiguously well-formed types should show a unimodal distribution centered narrowly around $z = +1$, reflecting a very high consistent rating across participants (Bondevik et al., 2020; Kush et al., 2019; Nyvad et al., 2022). As shown in **Figure 4**, all the [–Ex] types (dark blue) were indeed very clearly centered unimodally and narrowly around +1, and all the [+Ex] in the Control 1 condition (B) and Control 2 condition (C) were equally clearly centered around –1. Extraction from complement clauses headed by *that* also showed a unimodal distribution around +1, reflecting the fact that such extraction is indeed fully acceptable. Extractions from *if*-, *when*-, and *because*-clauses, on

the other hand, all showed unimodal distributions centered around –1, suggestive of islandhood. However, for *if* [+Ex], the distribution is broader and less pointy, indicating more variation.

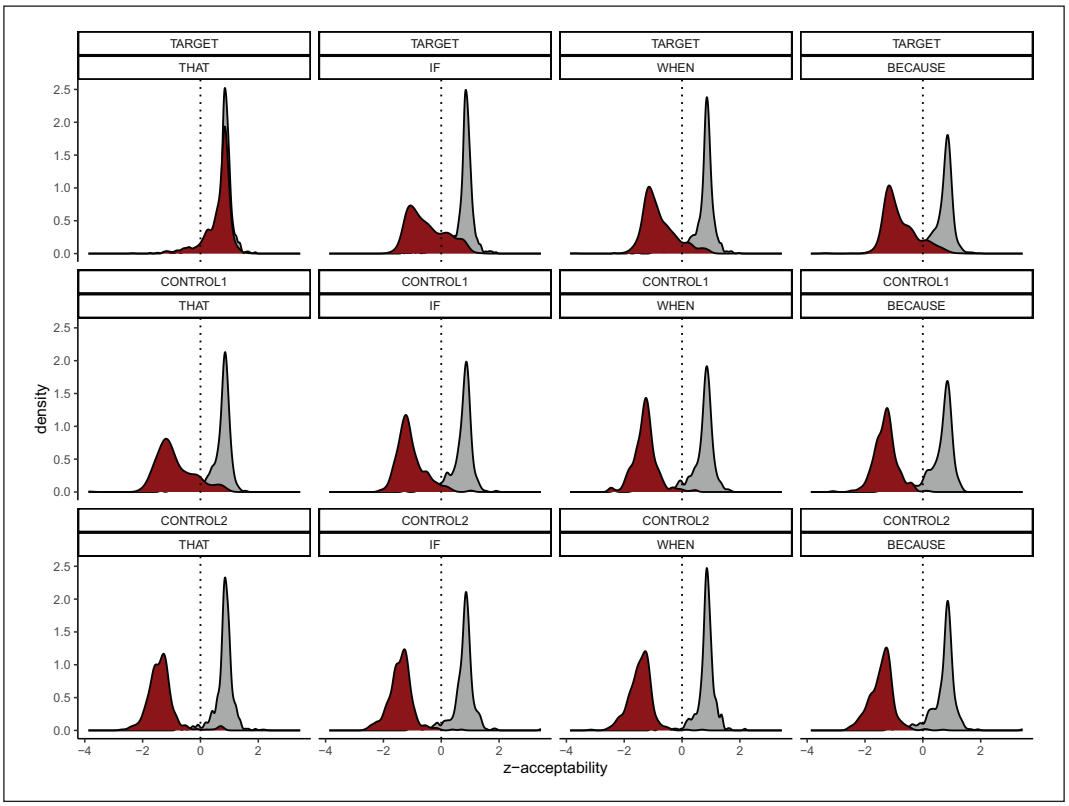

Figure 4: Density plots for each condition. Grey color: [–Ex], dark red: [+Ex].

3.5 Trial effects

To test for positive correlations between repeated exposure of type and acceptability ratings, we plotted the mean acceptability (across participants on each list) of each occurrence of the four extraction types in the target category (we leave the controls aside); see **Figure 5 (A)**. The effects were all very small and none of them were significant ($p > 0.1$) (though there was a positive trend for *that*, and a negative one for *if*). However, it should be noted that each participant was only exposed to 9 adjunct clause violations (3 of each type, *if, when* and *because*), which may not be a sufficient number of exposures to elicit a satiation effect for adjunct islands (Chaves & Putnam, 2020).

Because it is standardly assumed that adjunct clauses headed by *if*, *when*, and *because* have the same syntactic structure, we combined the three in **Figure 5 (B)** to see if their combined number of trials (i.e. 9) would show a significant effect. This was not the case ($p > 0.6$).

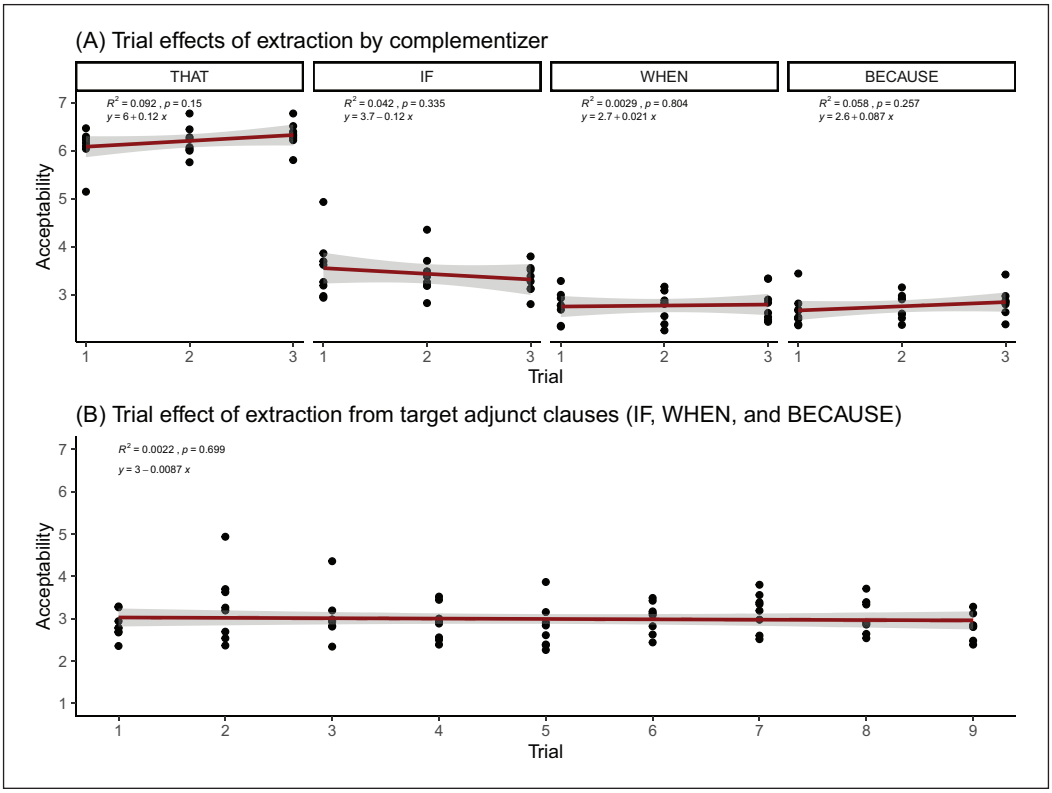

Figure 5: (A) Acceptability as a function of repetition (trial) of extractions from complement clauses headed by *that* and from adjunct clauses headed by *if*, *when*, and *because* in the target category. **(B)** Trial effect of extraction from target adjunct clauses headed by *if*, *when*, or *because*. Grey shading: 95% confidence interval.

4. Follow-up studies

The results of our study showed that the ratings for extraction from *if*-, *when*-, and *because*-clauses in Danish all showed unimodal distributions centered around –1, which is suggestive of islandhood. This is unexpected, given that Danish, just like the other MSc. languages, has been argued to permit extraction from finite adjunct clauses (Hansen & Heltoft, 2011; Jensen, 2002; Müller & Eggers, 2022) To test for the potential effects of context and extraction type, we ran the experiment two additional times. We also compared our findings with the results from Nyvad et al. (2022) to see if there was a language effect.

4.1 The effect of context

The presence of a preceding context has been shown to increase acceptability of extraction from adjunct clauses (Kush et al., 2019), and our stimulus set was carefully constructed to facilitate such extraction. However, since the present study was translated from English into Danish, it is

at least conceivable that this translation might have made the context less than natural, although we carefully screened the entire set several times. One reason why the results for extraction from adjunct clauses with *hvis* ('if'), *da* ('when'), and *fordi* ('because') are much less acceptable than predicted might be that the contexts were accidentally biased against them, favoring extraction from complement clauses with *at* ('that'). To test whether that was the case, we ran the experiment again with the same set of stimulus sentences (both target and control) just presented without context. The procedure was otherwise the same as in the experiment with context. A total of 49 people participated in the experiment (age 19–46, mean = 26 years, SD = 4.8). The number of participants per list ranged from 4 to 10. To test the effect of context, we ran a linear mixed-effects model for the targets, with z-transformed acceptability as outcome variable and type (see **Table 1**) and context and their interactions as predictors and random intercepts by item, and random slopes for context by item. The results are summarized in **Table 10** and **Figure 6** (second column from the left).

Table 10: Statistical model for the [±context] (Experiment 1 vs. Experiment 2) × type interaction effect on z-transformed acceptability: lmer(Acceptability_z ~ Type * Context + (1 + Context | Item).

	estimate	SE	df	t	p	
[+Context] vs. [–Context]	–0.081	0.084	172.540	–0.958	0.339	
hvis ('if') [–Ex] × [–Context]	0.042	0.119	172.540	0.351	0.726	
da ('when') [–Ex] × [–Context]	0.020	0.119	172.540	0.167	0.868	
fordi ('because') [–Ex] × [–Context]	–0.020	0.119	172.540	–0.170	0.865	
at ('that') [+Ex] × [–Context]	0.031	0.119	172.540	0.262	0.793	
hvis ('if') [+Ex] × [–Context]	0.205	0.119	172.540	1.717	0.088	.
da ('when') [+Ex] × [–Context]	0.178	0.119	172.540	1.488	0.139	
fordi ('because') [+Ex] × [–Context]	0.083	0.119	172.540	0.699	0.486	

The overall result is basically the same with and without context; compare the two leftmost columns in **Figure 6**. There were no significant interactions between context and type for any of the target types ($p > 0.1$; though it was marginal for extraction from *hvis* 'if', $p < 0.062$). The fixed (main) effect for context was also not significant ($p > 0.34$).

4.2 Extraction type (relativization vs. topicalization)

Another potential cause for the low ratings for extraction from adjunct clauses could be dependency type. Previous studies have shown that relativization (Sprouse et al., 2016) and

topicalization from an island is better than *wh*-movement in English and Norwegian, respectively (Kush et al., 2019), and there might be language-specific differences that make relativization from adjunct islands more marked in Danish than in English. Danish is a V2 language, and according to Mikkelsen (2015, p. 637), approximately 35% of main clauses have XVS word order (where X is a an overt non-*wh*-constituent), suggesting that topicalization (or at least, fronting) is fairly common.[5] In a study by Kush et al. (2019), it is reported that the acceptability of topicalizing out of a conditional clause in Norwegian, on average, is almost on a par with that of extracting from an embedded declarative clause. To investigate for the potential effect of extraction type, we ran the original experiment (with context) again.

We removed the main matrix part of the sentences ("It is obvious that…") and inserted a clause-initial adverb or discourse particle (corresponding to "so"); compare (6) and (3) above.

(6) NON-ISLAND STRUCTURE, [–EXTRACTION]:

a. Så jeg blev overrasket over, **at** hun faktisk gennemførte dét program.
so I became surprised over **that** she actually completed that exercise
'So I was surprised **that** she actually completed this exercise.'

ISLAND STRUCTURE, [–EXTRACTION]:

b. Så jeg ville blive overrasket, **hvis** hun faktisk gennemførte dét program.
so I would become surprised **if** she actually completed that exercise
'So I would be surprised **if** she actually completed this exercise.'

c. Så jeg blev overrasket, **da** hun faktisk gennemførte dét program.
'So I was surprised **when** she actually completed this exercise.'

d. Så jeg blev overrasket, **fordi** hun faktisk gennemførte dét program.
'So I was surprised **because** she actually completed this exercise.'

NON-ISLAND STRUCTURE, [+EXTRACTION]:

e. Så <u>dét program</u> blev jeg overrasket over, **at** hun faktisk gennemførte _.
So this exercise became I surprised over **that** she actually completed
'So, this exercise, I was surprised **that** she actually completed.'

ISLAND STRUCTURE, [+EXTRACTION]:

f. Så <u>dét program</u> ville jeg blive overrasket, **hvis** hun faktisk gennemførte _.
so that program would I become surprised **if** she actually completed
'So, this exercise, I would be surprised **if** she actually completed.'

[5] Similarly, in a Danish corpus study, K. R. Christensen and Nyvad (2024) found 120 monoclausal examples (approx. 9%) with a fronted accusative object pronoun vs. 1,165 examples with an initial nominative subject pronoun, suggesting that object fronting is not rare. Furthermore, though subject fronting was almost 10 times more frequent than object fronting, they were equally acceptable, approx. 6 on a 7-point scale.

g. Så dét program blev jeg overrasket, **da** hun faktisk gennemførte _.
'So, this exercise, I was surprised **when** she actually completed.'

h. Så dét program blev jeg overrasket, **fordi** hun faktisk gennemførte _.
'So, this exercise, I was surprised **because** she actually completed.'

A total of 85 people participated in the experiment (age 22–78, mean = 38 years, SD = 12.2). The number of participants per list ranged from 7 to 15. We analyzed the results with a general linear model for the targets, with z-transformed acceptability as outcome variable and type (see **Table 1**) and extraction type (relativization vs. topicalization) and their interactions as predictors.[6] The results are summarized in **Table 11** and **Figure 6** (third column from the left).

Again, the overall result is basically the same; compare the leftmost column and the third column form the left in **Figure 6**. There was a significant interaction with extraction type for complement clauses headed by *that* ('at') ($p < 0.001$), and also a significant interaction with extraction type for adjunct clauses headed by *hvis* ('if') ($p < 0.05$). In other words, topicalization from a complement clause is slightly, but significantly, worse than relativization, whereas from an adjunct clause headed by *hvis* ('if'), topicalization is slightly, but significantly, better than relativization. The fixed (main) effect for extraction type was not significant ($p > 0.65$).

Table 11: Statistical model for the extraction type (relativization [Experiment 1] vs. topicalization [Experiment 3]) × type interaction effect on z-transformed acceptability: lm(Acceptability_z ~ Type * ExType).

	estimate	SE	t	p	
Extraction type (Topic vs. Rel.)	0.016	0.036	0.449	0.653	
hvis ('if') [–Ex] × Topic	–0.024	0.051	–0.476	0.634	
da ('when') [–Ex] × Topic	0.024	0.051	0.477	0.634	
fordi ('because') [–Ex] × Topic	0.104	0.051	2.059	0.040	*
at ('that') [+Ex] × Topic	–0.215	0.051	–4.256	0.000	***
hvis ('if') [+Ex] × Topic	0.105	0.051	2.068	0.039	*
da ('when') [+Ex] × Topic	–0.055	0.051	–1.092	0.275	
fordi ('because') [+Ex] × Topic	–0.057	0.051	–1.124	0.261	

[6] Since neither the participants nor the items were the same across the two experiments, we did not include participant and item as random factors.

4.3 Language differences (Danish vs. English)

To test for potential cross-language effects, we fitted a linear mixed-effects model to the target sentences (combining the data from Experiment 1 (see 3.1) and the data from Nyvad et al. (2022)), with z-transformed acceptability as outcome variable and type (see **Table 1**) and language (Danish vs. English) and their interactions as predictors and random intercepts for item, and random slopes for language by item.[7] The results are summarized in **Table 12** and **Figure 6** (rightmost column).

Table 12: Statistical model for the extraction language (Danish [Experiment 1] vs. English) × type interaction effect on z-transformed acceptability: lmer(Acceptability_z ~ Type * Language + (1 + Language | Item).

	estimate	SE	df	t	p	
Language (Danish vs. English)	–0.184	0.055	183.327	–3.373	0.001	***
hvis ('if') [–Ex] × English	0.174	0.077	184.038	2.250	0.026	*
da ('when') [–Ex] × English	0.152	0.077	183.505	1.965	0.051	.
fordi ('because') [–Ex] × English	0.149	0.077	183.605	1.934	0.055	.
at ('that') [+Ex] × English	–0.291	0.077	183.877	–3.762	0.000	***
hvis ('if') [+Ex] × English	0.875	0.077	183.480	11.333	0.000	***
da ('when') [+Ex] × English	0.630	0.077	183.704	8.163	0.000	***
fordi ('because') [+Ex] × English	0.370	0.077	183.696	4.796	0.000	***

As shown in **Table 12**, there were significant interactions with language for all clause types. All types were rated significantly higher in the English experiment, except extraction from a complement clause (*at* 'that'), which was rated significantly lower in English than in Danish. This pattern is also clear in **Figure 6**. Mere visual inspection of the top row shows (near) identical results for the three Danish experiments (Exp. 1–3), while the ratings are higher in the English experiment, except for *that* [+Ex] (top-right panel) (Exp. 4). This is also shown in **Figure 7**. Note also that the results for coordinate structures (Control 2) are the same: uniformly acceptable without extraction and, likewise, uniformly unacceptable with extraction. The crosslinguistic variation in our study is restricted to extraction from adjunct islands, not across the board (even though we did not run any statistical models to test for this).

[7] Since the participants were not the same across the two experiments, we did not include participant as a random factor.

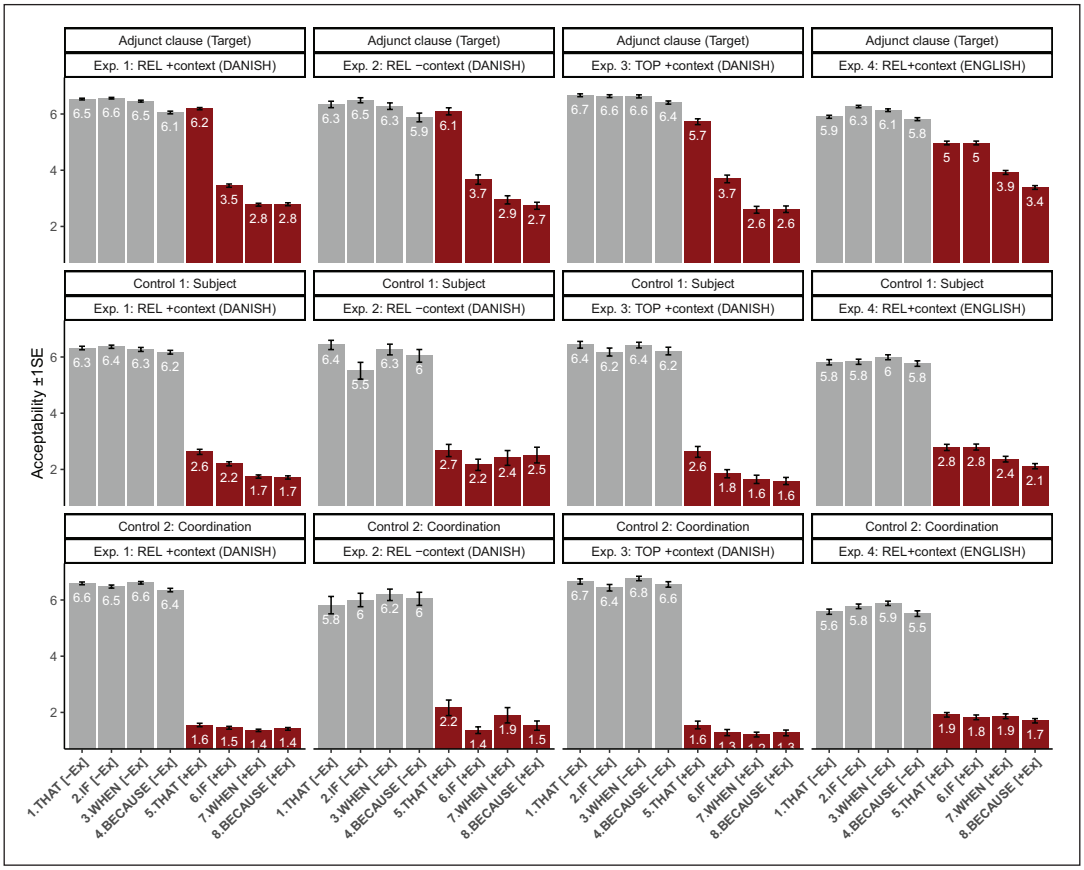

Figure 6: Comparison of the four experiments. From left: Exp. 1 = Danish relativization with context, Exp. 2 = Danish relativization without context, and Exp. 3 = Danish topicalization with context. The results in rightmost column, Exp. 4 = English relativization with context, are from Nyvad et al. (2022). (Significance markers are left out.)

In short, we have repeated the experiment twice, with slight modifications (coming close to replications of the experiment): First, we repeated the experiment but without context and got the same results, and then, we repeated it with topicalization instead of relativization, again with a very similar result. This is remarkable, given that the chance of getting the same result randomly is exceedingly unlikely, and also because replication studies are rare (Baker, 2016), perhaps especially in psycholinguistics (Kobrock & Roettger, 2023). To some extent, we also got the same the results as Nyvad et al. (2022), at least in terms of the relative average ratings in **Figure 6**, though with different effect sizes. Nyvad et al. (2022) report higher island effect size for *because* (DD = 0.74), followed by *when* and *if* (0.63 and 0.16, respectively), which, in part, can be explained by the relatively lower ranking of *that* [+Ex] in the English data.

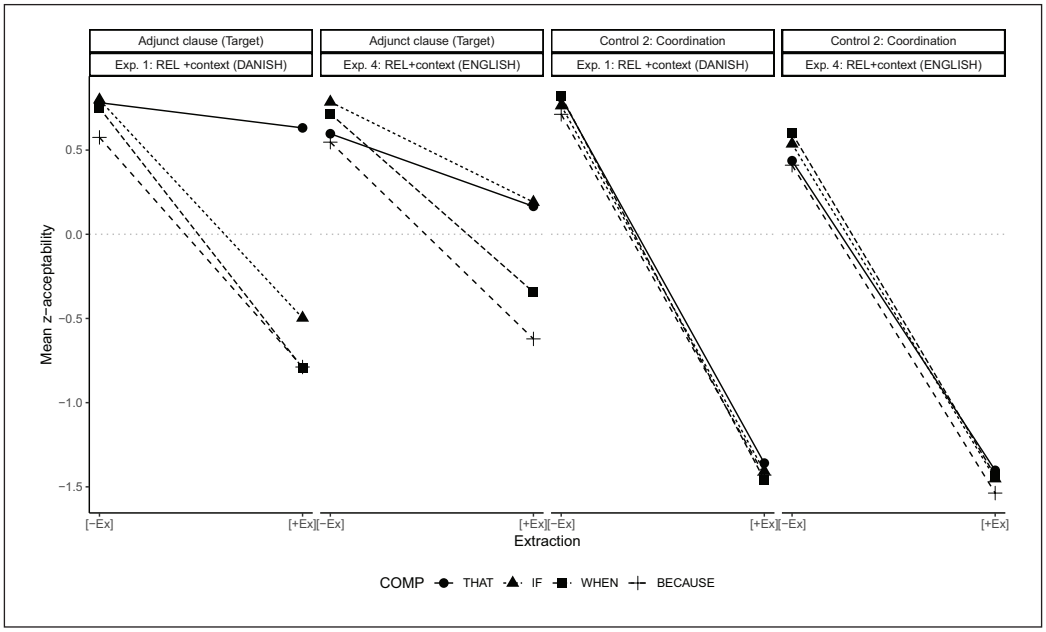

Figure 7: The effects of extraction (relativization) with context from adjunct clauses in Danish and English (left panels) and from coordinate structures in Danish and English (right panels).

5. Discussion

The purpose of the present study was twofold: First, to find out whether the acceptability of extraction out of different types of adjunct clauses varies in Danish, such as has been found in Swedish and Norwegian. Second, to compare adjunct island sensitivity in Danish and English, based on the same basic (translated) set of stimuli, differing minimally across conditions.

In our study, the participants systematically ranked the baseline [–Extraction] types (targets as well as Control 1 and 2) very high and the [+ Extraction] types in the controls (in particular Control 2) very low, as also shown in the very limited variance (cf. the tight, unimodal distribution in the density plots). We take this systematic, uniform behavior to indicate that the participants indeed did the task as instructed. Had they not done so, we would expect more variance and random answers in the control conditions, as well as data overall.

The most significant finding of this study is the variability attested both across different adjunct clause types and, perhaps even more surprisingly, across English and Danish. This significant level of variation in extraction patterns suggests that the theoretical framework accounting for the possibility of extraction from adjunct clauses needs to be more fine-grained, given that established accounts of adjunct islands (based on UG and/or the ECP) *prima facie* predict more or less uniform unacceptability across constructions as well as across languages. While the cross-construction variation is difficult for structural accounts to explain, it might

be expected on a functional account, given that different constructions arguably have different functions (see Liu et al., 2022, p. 518; Winckel et al., 2025). The crosslinguistic variation, on the other hand, might be accounted for under a structural approach that relies on e.g. parametric variation in what counts as a bounding node across languages.

On the one hand, we find that Danish is similar to Swedish (Müller, 2017), Norwegian (Bondevik et al., 2020), and English (Nyvad et al., 2022) in showing variability in the acceptability pattern of extraction from conditional, temporal, and causal adjunct clauses. On the other hand, the acceptability ratings for extraction from adjunct clauses are surprisingly low in Danish compared to the corresponding results from English – though perhaps not in isolation (as it is common that ratings are lower than expected in formal experiments, which may, in part, be due to the lack of contextual cues, Bondevik et al., 2020; Christensen & Nyvad, 2014; Kush et al., 2018; Müller, 2019, p. 209; Tutunjian et al., 2017; Wiklund et al., 2017).

5.1 Variation between the different adjunct clause types

Speakers of Mainland Scandinavian (MSc.) languages have been reported to "accept and produce filler-gap dependencies into domains that were considered islands in many other languages" (Kobzeva et al., 2022, p. 5). The present study joins a growing body of experimental work (e.g. Bondevik et al., 2020; Kush et al., 2019; Nyvad et al., 2022) suggesting that adjunct clauses are not uniform when it comes to extractability. Rather, a more fine-grained account is needed in order to capture the variation that we see across different adjunct clause types, regardless of whether the gradience is located at the level of syntax, pragmatics/information structure, processing, or all of the above.

Interestingly, we see the same overall pattern in the adjunct islands and (to a lesser extent) the subject islands, in the sense that extraction from a complement clause is better than extraction from a conditional clause (introduced by *hvis* 'if') which in turn is better than extraction from both the temporal (*da* 'when') and causal (*fordi* 'because') adjunct clauses. This parallelism in the patterns attested across the two types of island configurations suggests that the relative acceptability is dependent on the complementizers.

5.1.1 A quick glance at frequency

It is theoretically possible that the difference in acceptability between the complementizers could simply be the result of collocation, such that participants found extraction from *hvis* ('if')-clauses better than that from *fordi* ('because')-clauses due to e.g. *overrasket* ('surprised') occurring more frequently with *hvis* than *because.* We looked up the collocations between the predicative adjectives in the study and the four different complementizers that they co-occur with in the stimuli in the Danish written language corpus KorpusDK; see **Table 13**. The total number of

occurrences of the set of predicative adjectives was 5,547, which means that, overall, they occurred with one of the four complementizers 18.3% of the times.

Table 13: Collocations of the 24 predicative adjectives used in our study and the four types of complementizers in KorpusDK. Prep = *over / for* 'over / for' (e.g. *overrasket* 'surprised *over*, but *glad* 'happy' *for*). The search string allowed for 0–3 words between the lemma *blive* 'become' and the adjective.

Complementizer	Count	Percent
Adj + Prep + *at* 'that'	445	43.8%
Adj + *hvis* 'if'	199	19.6%
Adj + *da* 'when'	329	32.4%
Adj + *fordi* 'because'	44	4.3%
Total	1,017	100,0%

The numbers in **Table 13**, though significantly different ($\chi^2 = 350.96$, df = 3, $p < 0.0001$), do not explain the attested acceptability differences across complementizers. Overall, *at* 'that' is most frequent and most acceptable, and *fordi* 'because' is least frequent as well as least acceptable. However, *da* 'when' is more frequent but less acceptable than *hvis* 'if'. (For a detailed discussion of frequency, acceptability, and complexity, see Christensen & Nyvad, 2024.)

We also searched the corpus to see if there were examples with any matrix predicative adjective (not just those in our study) followed by any preposition followed by *hvis/da/fordi*, i.e. examples where the adjunct clause was realized as a complement clause under a preposition, parallel to clauses with *at*. This search, however, resulted in only a single example with *hvis* (zero hits for *da* and *fordi*), which is in line the standard analysis of *hvis*, *da*, and *fordi* as adjunct complementizers.[8]

5.1.2 No effect of context

One might speculate that the variation in acceptability of extraction across the different complementizers may be dependent on context. Keller (2000) argues that grammatical constraints can be divided into those that are affected by the presence of a context (and which should be in effect crosslinguistically) and those that are not. As pointed out by Sprouse (2007, p. 48), many

[8] The single example of a *hvis*-clause as the complement of a preposition:

(i) *Jeg har to spørgsmål, som jeg ville blive utrolig glad for, hvis I kunne svare på*
'I have two questions that I would become incredibly happy about if you could answer.'

Anecdotally, we also find such examples with *hvis* 'if' much more acceptable than corresponding examples with *da* 'when' or *fordi* 'fordi'.

frameworks have analyzed the context-dependent constraints as being essentially semantic or pragmatic in nature, whereas context-independent constraints can be interpreted as relating to syntax (or grammar) proper. This line of reasoning leads Erteschik-Shir (2006), among many others, to argue that context-independent grammatical constraints should yield a binary output (i.e. either grammatical or ungrammatical). While Sprouse (2007) found context not to interact with island structures involving *wh*-movement, several other studies have indeed found a positive effect of supplying a facilitating context on extractability (Kristensen et al., 2013; Kush et al., 2019).

In the present study, however, there was no significant effect of context (cf. **Table 10**). Though it is possible that we might have seen an effect with a pronominal filler instead of a full DP (e.g. *den* 'that' instead of *den prøve* 'that test'), and though it could be slightly more unnatural to repeat a full DP in the target sentences when it has already been mentioned in the context, the lack of significant context effects is remarkable and suggests that the variation found between complementizers is fairly independent of this pragmatic factor.

5.1.3 No effect of dependency type

Another unexpected result was the lack of effect on the overall acceptability patterns when changing the dependency type from relativization to topicalization. Importantly, the dependency lengths in the relativizations, (3), are longer than in the topicalization structures, (6), and hence there is a difference in complexity between the two, and the consequent difference in processing load is not reflected in the acceptability levels. (This may, in part, be explained by the fact that the corresponding baselines for the topicalization study are structurally simpler.) Sprouse et al. (2016) found an effect of dependency type for *if*-clauses in English, such that relativization was more acceptable than *wh*-movement, and they argue that dependency type is key to the understanding of the variation in the acceptability of extraction from islands. However, we found no difference between topicalization and relativization. As shown in **Figure 6**, *if*, *when*, and *because* are equally unaffected by the absence of context.

On the one hand, island effects have traditionally been thought to be due to universal constraints on A-bar-movement (Chomsky, 1973, 1977, 2000; Cinque, 1990; Huang, 1982), and as such, we should, in principle, expect topicalization to be as unacceptable as *wh*-movement from islands. On the other hand, in the MSc. languages, it is more common to front topical DPs than *wh*-phrases (Lindahl, 2022). In addition, topicalization generally receives a higher acceptability rating. Furthermore, both relativization and topicalization often pattern together, compared to question formation (see e.g. Williams, 2011), a connection that has also been highlighted by Kuno (1976) and Abels (2012). Douglas (2016, p. 147) even argues that relativization and topicalization share a discourse-linking feature absent in *wh*-questions. Indeed, Müller and Eggers (2022) showed that relative clause extraction is more prevalent than adjunct clause extraction

in both Danish and English. While they found more examples of topicalization out of relative clauses than relativization, the prevalence of relativization and topicalization out of adjunct clauses was at the same level in Danish. This picture fits our experimental results, which do not suggest a distinction between the two dependency types when it comes to extractability from adjunct clauses.[9]

5.1.4 The lexico-semantics of the complementizers

Bondevik et al. (2020) replicated the results of Kush et al. (2019) by finding island effects in connection with topicalization from temporal and causal adjunct clauses, but not conditional clauses. However, in our study, even though we appear to find the same overall pattern, the temporal adjunct clauses introduced by *når* in the Norwegian study differed in their island effect from the causal adjuncts (*when* > *because*), whereas our temporal and causal adjunct clauses were on a par in the raw scores. The DD scores, however, show that the island effects for *da* and *fordi* are actually different (see **Figure 3 (D)**). This may be related to the fact that our temporal adjuncts were introduced by *da*, not *når* (both translatable to 'when'). Danish, like Norwegian, is a "two-'when' language" (Vikner, 2004, p. 133). Whereas *da* covers episodic clauses describing a unique identifiable eventuality in the past, *når* can be used in both past and present habitual clauses, as well as for future clauses. Simply put, the temporal *da*-clause presupposes that there is only one situation which would correspond to the description it expresses, which is not the case with *når* in the past (see the detailed discussion in Vikner, 2004). However, like *when* in English, *da* also has an explicit causal reading, and in the stimuli of the present study, it is not possible to completely distinguish between the temporal (informal) and the causal (formal) readings.[10] But then again, the density plots in **Figure 4** do not indicate ambiguity, because the distributions are not bimodal or spread-out. It may indeed be possible that all participants got a causal interpretation on top of the temporal one.

According to Truswell (2011, p. 44), causation between matrix and embedded clause may be a crucial factor in extractability from adjunct clauses. Similarly, Chaves and Putnam (2020) argue that the degree of extractability from adjunct clauses is correlated with Van Valin's (2005, pp. 208–209) "inter-clausal relations hierarchy", which states that causal relations express more cohesion or a tighter semantic connection than purely temporal ones. However, this causative

[9] For a discussion of the possibility of accounting for why relativization and topicalization may be easier than *wh*-movement from adjunct islands using a featural Relativized Minimality account, see Nyvad et al. (2022). In short, the moved element in both topicalization and relativization can be argued to have a richer featural specification than the potential *wh*-operator in the adjunct clauses.

[10] It is also interesting to note that, unlike *hvis* ('if') and *fordi* ('because'), *da* ('when') is not able to occur in complementizer stacking: *Jeg blev sur, fordi (at)/da (*at) han glemte kaffen* ('I got upset because that/when that he forgot the coffee') (see the discussion in Nyvad et al., 2017; Vikner et al., 2017).

connection was a potential in all our adjunct clause types, both in the English study (Nyvad et al., 2022) and the present Danish one, given that they were embedded under emotive factive predicates which elicit a causal reading, and, as such, extraction from all three adjunct clause types should be equally licit, contrary to our results. It could be argued, however, that for the *da*-clauses, a causative reading was not forced, and that the between-participant and between-item variation may be partially explained by this. However, it does not appear as though this factor is determinant in the variation that we have found overall, although a causative link between the adjunct clause and the matrix may be one of the preconditions for extraction (Bondevik et al., 2020).

Our data support recent proposals that the notion of *gradience* should be properly operationalized and integrated into accounts of island phenomena in a falsifiable way. Our data suggest that variation attested across adjunct clause types may be due to the lexical semantics of the complementizers interacting with factors relating to information structure and discourse function.

5.1.5 The effect of information structure/discourse function

Information-structural and discourse-functional accounts of island phenomena have long held that the gap in filler-gap dependencies generally must occur within the potential focus domain of the sentence or in the part of the utterance that is "asserted", "focused" or "dominant" (Deane, 1991; Erteschik-Shir, 1973; Van Valin, 1996). As a result, adjunct clauses are not necessarily expected to exhibit uniform island effects. The degree to which an adjunct clause can be argued to have this special status may interact in intricate ways with a range of lexical and semantic factors. In addition, we would argue, it is possible that this "privileged" status that facilitates extraction is a product of an interplay between the semantics of the complementizer and the dependency type.

Fillers in long-distance dependencies typically occur in "discourse-prominent slots" (Goldberg, 2005, p. 135), as extraction, in and of itself, places a discourse element in a prominent position at the left edge of a clause. Adjunct clauses typically provide circumstantial information about the main event which may cause them to be considered non-essential in relation to the situational knowledge described by the matrix predication, which may, in turn, affect extractability.

According to Goldberg (2005, p. 135, (14)), it is pragmatically anomalous for an element to be both discourse-prominent and backgrounded at the same time, and she formulates the following discourse generalization: Backgrounded constituents may not serve as gaps in filler-gap constructions (Backgrounded Constructions are Islands, henceforth BCI) (see also Cuneo & Goldberg, 2023). She argues that such backgrounded constituents are characterized by being neither "primary topic" nor part of the "focus domain" of the sentence, but rather by being presupposed. Presuppositions, she asserts, can be revealed through negation tests, as they should

be implied by both the positive and negative form of the sentence (Goldberg, 2005, p. 130) (for a critical review of the negation test, see Pérez-Leroux & Kahnemuyipour, 2014, pp. e121–e122). The BCI, however, does not predict extractability to vary as a function of the dependency type involved, unlike Abeillé et al.'s (2020) Focus-Background Conflict Constraint (FBCC), according to which a backgrounded constituent cannot function as a gap for a filler-gap dependency when the filler is focalized, because that leads to a discourse-functional "clash" between filler and gap. This means that variation between dependency types is expected under the FBCC: *Wh*-extraction from a backgrounded constituent should be unacceptable, while relativization should be acceptable, since a relative clause typically just adds information about a "given" NP and does not focalize the filler (i.e. the correlate), and any type of fronting not involving focalizing should be equally unaffected by the FBCC. As also pointed out by Lindahl (2017, pp. 160–161), the information-structural/discourse-functional approaches appear to face a problem when it comes to cleft structures, given that the clefted constituent is focused and the negation test identifies the embedded clause as presupposed (hence, as an island), even though this type of construction generally seems to allow extraction in the MSc. languages.

The three adjunct clause types in our stimuli have distinct semantico-pragmatic profiles, and they are not equivalent in terms of discourse function: The temporal and causal clauses are presupposed, while the conditional clause is not. Hence, what our study appears to demonstrate is that the BCI (or FBCC) would have to be gradient in order to account for the variability in the acceptability patterns attested, such that there is a spectrum of backgroundedness (as indeed suggested by Abeillé et al., 2020; Goldberg, 2013; Winckel et al., 2025) and this continuum interacts with the possibility of extraction from e.g. an adjunct clause. As pointed out by Schütze et al. (2015, p. 35), however, the degree of backgroundedness of a particular constituent is "not fixed for a given sentence" and may vary significantly as a function of the containing discourse. Hence, we need a way to operationalize the concept of backgroundedness, which to our knowledge does not yet exist. For example, though Abeillé et al. (2020, p. 21) argue that the FBCC is gradient, they say "a gradient discourse model is outside the scope of this paper". (But see Cuneo & Goldberg, 2023.)

The adjunct clauses in our stimuli were constructed such that they express a cause or condition for the psychological state expressed by the matrix clause, and, as such, the adjunct clauses express at-issue information for the main action (Chaves & Putnam, 2020, p. 91). Following Simons (2007), Kush et al. (2021) suggest that extractability may instead depend on the clause containing the gap contributing "wholly or partially new information" (Kush et al., 2021, p. 39), or being the *main point of utterance* (MPU) of the sentence. An embedded clause, according to Simons (2007), can be MPU, especially if the predicate in the matrix clause expresses how the speaker feels about the information in the embedded clause. This is reminiscent of Kluender (2004, pp. 121–122), who argues that the filler's relationship with the predicate in the main clause is as important as the relationship that it has to the gap position.

5.2 Crosslinguistic differences

While the results for English indicated that relativization out of conditional, temporal, and causal clauses does not constitute a strong island, the present data suggest that all three types of adjunct clause extraction are island violations in Danish, for two reasons: One, the DD scores (the differential cost of extraction) are all above the >0.75 threshold suggested by Kush et al. (2019, p. 401), and two, the interactions between complementizer type and extraction (see **Table 8**) were highly significant, which is a clear indication of islandhood. Overall, compared to English, the acceptability of extraction from adjunct islands in Danish turned out to be worse than expected. The crosslinguistic variation in island sensitivity observed between Danish and English raises intriguing questions about the underlying mechanisms that govern extractability from adjunct clauses. In what follows, we will propose several possible explanations of the observed crosslinguistic differences.

5.2.1 Processing differences?

Island phenomena have been explained by means of processing-based accounts, dating back (a least) to Pritchett (1991) and Kluender (1991) (see also Fodor, 1983, p. 310; Ross, 1987, p. 190). Kluender and Kutas (1993) argued that island effects may be the result of memory load exceeding capacity in connection with the processing of complex structures in e.g. adjunct island configurations. More recent processing-based accounts have been proposed, which draw on the importance of factors such as "accessibility" (Arnold, 2010), "predictability" (Hale, 2001; Levy, 2008) and "surprisal"' (Culicover et al., 2022). Overall, from a processing perspective, the relatively low frequency of island structures (Lindahl, 2017; Müller & Eggers, 2022) would predict low levels of acceptability for these configurations, and given that islands involve higher syntactic complexity, they should incur a higher processing cost (Christensen & Nyvad, 2024).

These processing accounts seem to struggle, however, in the face of the overall crosslinguistic variation found in English and Danish, assuming that the island configurations, which are presumably relatively uniform across the two languages, also incur more or less the same processing load. Nevertheless, one could imagine that the three different types of adjunct clause may vary in processing difficulty due to other (semantic or pragmatic) factors, not just within but also across the languages. However, to the best of our knowledge, no processing account has been proposed which is able to capture and operationalize variation across constructions and across languages.

In a broader sense, it could be argued that island extraction might be relatively common overall in the MSc. Languages, because of the frequency of topicalization structures in these languages, cf. 4.2 above. (See also Engdahl, 1997; Erteschik-Shir, 1973, pp. 126–127.) Though frequency in itself is not a predictor (it is what needs to be explained), the fact that topicalization is so widespread and common in Danish but marked in English might suggest that it is easier to extract from at least some of the configurations in Danish that are more island-like in English. In

this light, it is even more surprising that the acceptability ratings for island extractions are lower in Danish than in English.

From the perspective of a processing account, the lack of a difference between topicalization and relativization out of adjunct clauses in our Danish study is surprising. For example, Gibson (1998) and Lewis & Vasishth (2005) would not make a distinction between relativization and topicalization as such, since it is the length of the dependency (distance and/or number of discourse referents between filler and gap) that is important in their frameworks. However, in our study, the topicalization structures are shorter and the dependency lengths are longer in the relativizations. In other words, the lack of a difference goes against processing accounts based on dependency length.

5.2.2 Effect of translation

The stimuli employed in the Danish study were translated as closely as possible from the English stimulus set in Nyvad et al. (2022). However, in trying to keep the stimulus sets in the two languages as minimally different as possible, structurally as well as lexically, the formality level may have been inadvertently affected, potentially leading to differences in acceptability, such that the Danish version of the original English stimulus set is more stilted or formal, at least to some participants.

We attempted to address this question, at least in part, by running the experiment again without the context to see if the crosslinguistic difference might be an effect of the context somehow being worse in Danish than in English. The context was intended to simulate a turn in a conversation where someone is talking about a certain event and we thus aimed for a low formality level, which is why we modelled the actual target sentences on naturally occurring examples found in corpora.

If it is indeed the case that the translation introduced an unintended stiltedness in the Danish stimuli, then it is not surprising that the unextracted versions of the stimulus sentences in Danish appear to be on a par with the corresponding results in English, while the extracted version is worse in Danish: Extraction from islands is typically considered a spoken language phenomenon and hence a higher level of formality in the Danish stimulus items would clash more with the phenomenon under investigation, which is most often found in a colloquial context.[11]

[11] One could speculate that native speakers of English are used to more variation in English than native speakers of Danish are in Danish, because English is spoken by L2 speakers from across the globe (whereas L2 Danish is a much less frequent phenomenon) and that this difference could potentially lead to more tolerance among the native English speakers. However, the fact that the acceptability of the unextracted versions and extraction from *that*-clauses is not significantly higher in English than in Danish speaks against this. (It should be noted, however, that there may be a ceiling effect in the unextracted versions which renders crosslinguistic contrasts in these conditions opaque.)

It may also be the case that unintended variation in formality level between the two languages played a role which differed across the conditions: For instance, the complementizer *da* ('when') can introduce both temporal and causal adjunct clauses, but its use as a causal complementizer is associated with an elevated level of formality, and in our stimuli, it is actually not possible to tell the temporal and causal interpretations apart. In other words, the lower-than-expected raw scores for extraction from clauses introduced by *da* ('when'), may be related to a clash in formality level.

However, though we cannot completely rule out that it may have contributed to the attested difference between Danish and English, the fact that there was no significant effect of context (4.1) and no significant effect of dependency type (4.2) speaks against the translation as such being problematic. In addition, the sentences were constructed and carefully evaluated by native speakers, who did not find them problematic or stilted.

5.2.3 Syntactic difference between English and Danish

The lack of effect of removing the supporting context and of changing dependency type from relativization to topicalization could speak to a syntactic account of the results. However, adjunct islands are typically treated as a uniform class by syntactic theories (e.g. Chomsky, 1986; Huang, 1982; Rizzi, 1990). The variation attested between English and Danish in the two studies is thus difficult to explain with reference to syntactic structure alone or to standard generative analysis: The sentences containing extraction from adjunct clauses are generally thought to be structurally the same, both crosslinguistically in English and Danish, as well as across the three types of adjunct clause, *if/hvis*, *when/da* and *because/fordi*.

Any explanatory account of island phenomena needs to explain the variation attested across languages. It could be argued, as is done in Schütze et al. (2015, p. 35), that a parameterization of what counts as a bounding node may be able to capture crosslinguistic facts. Potential parametric variation in the structure of the left periphery in the MSc. languages has also been discussed in connection with island phenomena; Nyvad et al. (2017) suggest a *c*P/CP-distinction, which may help capture the extraction patterns found in Danish, as it affords a division between lexical, "root clause" CP, conceivably a bounding node for extraction, and the functional, potentially recursive "embedded clause" *c*P, which may not be (see also Vikner et al., 2017). This framework is reminiscent of Wiklund et al. (2007), who argue that the matrix predicate may select a clause with a more articulated left periphery (ForceP), thus carrying illocutionary force (corresponding to CP). Even so, this theoretical distinction cannot account for the low acceptability scores in Danish compared to English. In addition, in our data, the matrix predicate is invariably an emotive factive predicate, and the same predicate is kept constant across extractions. Moreover, all the adjunct clauses in our experiment were central (rather than peripheral) according to Haegeman's (2012, pp. 161–164) definition, meaning that their structural position (and hence,

semantic scope) should be the same, and due to this low adjunction site, extraction should be more easily acceptable (compared to peripheral clauses). (See also the discussion in Nyvad et al., 2022.) In other words, it would seem that the crucial factor determining the differences in the possibility of extracting must lie elsewhere.

It is, however, possible that a syntactic factor plays a role in the generally lower acceptability levels in Danish found in the present study, as compared to those found in English by Nyvad et al. (2022): In English, finite complement clauses headed by *that* normally cannot occur immediately after a preposition selected by an adjective (Quirk et al., 1985, pp. 658–661), a phenomenon known as the Case-Resistance Principle (Stowell, 1981), while the corresponding complement clauses in Danish require the presence of the preposition:

(7) He was surprised (*at) [that she noticed him].
(Quirk et al., 1985, p. 659)

(8) Han blev overrasket *(over) [at hun lagde mærke til ham].
he became surprised (over) that she put notice to him
'He was surprised that she noticed him.'

This basic syntactic asymmetry between English and Danish can potentially have led to the differences in the acceptability patterns attested, as this crosslinguistic difference might suggest a potentially stronger tendency towards a complement reading of the adverbial clause in the English stimuli than in the Danish, given that the complement reading in Danish would require the presence of a preposition in front of the adverbial clause; compare (7) and (8). In fact, a few of the comments left by the Danish-speaking participants revealed that some felt that a preposition was "missing" in the sentences containing adjunct island configurations. This could suggest that some of the participants might indeed have had a complement reading in mind at some point during the sentence processing. It is possible that a greater propensity towards an intermediate complement reading in the English stimulus sentences could have led to overall greater acceptability in the English data, but this is an empirical question for future research. Note again, however, that the density plots in **Figure 4** for all three adjunct islands show narrow and unimodal distributions, which does not suggest speaker variation in parsing strategy. Furthermore, the corpus frequencies in 5.1.1 also show that the complement reading (Adj + Prep + Comp) is not likely for *hvis*, *da*, and *fordi*, and that examples with Adj + Prep make up less than half of the examples.

5.2.4 Demonstrative relative clause reading or cleft reading?

The crosslinguistic differences attested between English and Danish in the two studies could also be the result of a difference in the status of the filler (or correlate) that is being interpreted in the

gap position of the adjunct clause. It may be the case that the English participants constructed a demonstrative relative clause, while the Danish participants constructed a cleft.

Proponents of the FBCC (e.g. Winckel et al., 2025) could argue that the discrepancy in acceptability levels in the English and Danish studies could be due to the English participants interpreting the fronting as a demonstrative, existential relative clause while a cleft reading is the only possibility in the Danish stimuli. Existential relative clauses are typically non-presuppositional and, hence, not backgrounded, which could have positive consequences for felicity. In clefts, on the other hand, the clefted constituent is focused and the negation test identifies the embedded clause as presupposed (see Lindahl, 2017, p. 160 ff for a discussion).

In the English stimuli (Nyvad et al., 2022), the extraction sentences are most plausibly interpreted as demonstrative relative clauses (e.g. *This is the exercise that I was surprised because she completed*). According to Kobzeva et al. (2022, p. 7), demonstrative relative clauses are characterized by the head of the relative clause being definite and preceded by a demonstrative pronoun and a finite form of *to be*.[12] In the present Danish study, however, the target sentences were unambiguously cleft sentences, due to the use of the acute accent (*dét program* 'THAT exercise') and, as a result, the clefted constituent receives emphatic stress and the NP that is co-referent with the gap is focalized (cf. the examples in (3)).

Winckel et al. (2025) compared extraction from relative clauses and clefts (*it*-clefts in English and *c'est*-clefts in French) and found significant island effects in both languages, but only for the clefts. According to Abeillé et al. (2020), the extraction in the demonstrative relative clause reading leads to less of a discourse clash than the extraction in the cleft structure. In a similar vein, Abeillé et al.'s FBCC would predict that, all else being equal, the acceptability ratings of the unambiguous cleft structures in Danish should be worse than what was found for the parallel English sentences: The latter could potentially be read as a demonstrative relative clause, making associating the filler with the gap inside the adjunct clause more felicitous; conversely, in the former, associating a clefted, hence, focused filler with a gap inside a backgrounded adjunct clause is less felicitous and less acceptable. This means that the information-structural match between the discourse function of the dependency type and the status of the adjunct constituent containing the gap is better in English than in Danish. If so, the crosslinguistic pattern for relativization could potentially be accounted for with reference to Abeillé et al.'s (2020) FBCC.

---

[12] Interestingly, the stimuli in Nyvad et al. (2022) (containing *demonstrative, existential* relative clauses) can be considered fundamentally different from the extractions from the English *presentational* relative clauses found in corpora by Müller and Eggers (2022), given that the latter tend to have an indefinite/non-specific correlate, as their primary function is to introduce a new referent into the discourse. Naturally occurring examples of extraction from island environments often appear with existential relative clauses and clefts (Engdahl, 1997; Erteschik-Shir & Lappin, 1979; Lindahl, 2017).

This line of reasoning, however, begs the question of why we see similar acceptability levels when the dependency type is topicalization instead of relativization. In principle, the FBCC can be extended to account for the frequently attested topicalization out of island configurations in the MSc. languages, since the fronting typically involves non-focal elements, and can, thus, pattern with relativization in terms of the discourse-functional role of the dependency type (Müller & Eggers, 2022). In other words, in topicalization structures, the topical filler is typically backgrounded information, and, in general, this dependency type is less foregrounding than question formation, which is focus movement. However, in our Danish stimulus sentences, the topicalization involves a contrastive focus, in that the topic is also focalized (or contrastive): The context introduces a discourse set and the topicalization picks out a particular referent in that set, as shown in (9), repeated from (6f) above:

(9) Context: *I det sidste træningsprogram jeg udarbejdede for Emma, ville jeg gøre det så godt som umuligt for hende og inkluderede derfor endnu et sæt virkelig brutale pull-ups.*
'In the latest workout routine I designed for Emma, I really wanted to make it impossible for her and included another set of particularly brutal pull-ups.'

Så det program ville jeg blive overrasket, [hvis hun faktisk gennemførte _].
that exercise would I become surprised if she actually completed
'I would be surprised if she actually completed that exercise.'

In this sense, the filler in our Danish stimuli could be argued to be focalized in both the relativization structures (interpreted as clefts) and in the topicalization structures (the filler is a referent picked from a discourse set in the context). Given that the adjunct clauses are indeed backgrounded, this would incur a discourse-functional clash not found in the English stimuli: Extraction by way of relativization is not predicted by the account in Abeillé et al. (2020) to incur a strong island effect, and the English stimuli involve a demonstrative, existential relative clause, which may not be at odds with the adjunct constituent containing the gap.

It thus appears that the discourse-functional role of the filler and the embedded clause containing the gap may indeed play a role in acceptability. The problem with this kind of explanatory model is, however, that the concept of focus is slippery, open to interpretation, and not easily operationalizable, at least not in a way that can capture the subtle differences in acceptability and make the FBCC falsifiable. The lack of reliable tests for focus and backgroundedness makes it tricky to evaluate the functional accounts in question in connection with the empirical data collected in our studies here (cf. the critique of the negation test, Lindahl, 2017, p. 160 ff; Pérez-Leroux & Kahnemuyipour, 2014, pp. e121–e122; but see Cuneo & Goldberg, 2023). In addition, discourse-functional alignment between filler and gap position may just be a precondition for felicitous extraction, interacting with a host of other factors. In other words, the way in which the grammatical system interacts with discourse-pragmatic factors, such as information structure and topic-focus articulation, could play a decisive role in island sensitivity. It is also possible that

specific constraints or language-internal factors in Danish influence island sensitivity differently than in English. These constraints might further interact with structural and processing factors, contributing to the overall acceptability levels. In other words, a substantial amount of research is still needed to extricate the component parts of island effects, across constructions and languages.

6. Conclusions

A major focus of linguistic theory is to try to uncover the nature of human language by identifying basic underlying principles, which can account for both universal patterns across languages and the variation in language-specific properties. Given that we do not have direct access to the language system or tacit grammatical knowledge in our brains, evidence in linguistics stems from probing its output (i.e. real-world performance or usage) and the behavioral measure of grammaticality is sentence acceptability (i.e. the degree to which native speakers find a given sentence "good" or "natural"). In the field of linguistics, and particularly generative theory, the tradition has been to make a more or less direct link between the two, even though grammaticality is often considered to be binary (i.e. a sentence is either grammatical or ungrammatical) and acceptability, gradient. Island effects are an excellent case study of the complex relationship between acceptability and grammaticality and may be the key to a range of important questions in linguistics and cognitive science.

The present study contributes to our understanding of islands by contributing data about island sensitivity in Danish and English, and it highlights the complexity of the phenomenon. In acceptability studies such as ours, performance data can be used as a prism or lens to study competence, and the rainbow-colored reflections shining back at us suggest that the language organ allows for a great deal of variation in island structures. While various theoretical accounts have been considered, none provide a comprehensive explanation for the language-internal and crosslinguistic variability attested. In the syntax literature, claims of crosslinguistic variation in relation to island phenomena have been based on comparisons of very different structures in e.g. English and the Mainland Scandinavian languages. For instance, Schütze et al. (2015) argue that English exhibits island effects for both complex NPs and *wh*-islands, while Swedish exhibits neither. However, the examples that they employ are not parallel, so a range of extra-grammatical factors could play into the differences in acceptability. Once the structures under empirical investigation are minimally different, the impact of different dependencies (e.g. the adverse effect of *wh*-movement out of context) or the presence/absence of context can be controlled for, and the emerging picture appears to be that the crosslinguistic variation between e.g. English and the Mainland Scandinavian languages may have been exaggerated.

The parallels in the patterns attested in Norwegian and Danish in relation to the acceptability of extraction from conditional, temporal, and causal adjunct clauses, as well as the attested discrepancies between the results in the English and Danish studies, may potentially highlight

the explanatory value of the functional account over a purely structural one. Adjunct clauses are traditionally analyzed as structurally uniform, within as well as across languages, at least for languages as closely related as Danish and English. It is conceivable that different complementizers have subtly different structural properties (e.g. different types of operators, a different number of projections, differences in bounding nodes, etc.), but current syntactic theories do not offer a comprehensive account that captures the island data. Furthermore, we found no difference between topicalization and relativization, i.e. no effect of dependency type, which also suggests that the explanation does not lie (exclusively) in structural or semantic properties of the complementizer or adjunct structure as such. In short, the crosslinguistic variation adds weight to the argument that the acceptability of extraction from adjunct clauses is not solely dependent on syntactic constraints, but is also influenced by pragmatics, information structure, and processing constraints. Note, however, that the lack of difference between topicalization and relativization is not captured by pragmatics or information structure, either. Neither can it be explained by processing alone: If it is assumed that the structure for clausal adjunction is parallel in Danish and English, we would not predict a difference between the two languages, contrary to fact. Likewise, if it is assumed that different adjunct clauses have the same syntactic structure, they should be equally easy or difficult to process, again, contrary to fact.

Overall, the patterns attested in English and Danish suggest that the two languages are quite similar when it comes to adjunct clauses, a syntactic environment that has previously been thought to be a strong island. The studies show that different adjunct clause types have a non-uniform behavior when it comes to extractability and, in fact, some of them (*if/hvis* clauses) do not behave like categorical islands. However, the findings also presented a remarkable crosslinguistic variation between Danish and English which may highlight potential pitfalls when translating from one language to another. Although significant efforts were made to ensure minimal differences between the original English sentences and their Danish translations, language-specific nuances, and formality levels might have inadvertently influenced acceptability judgements, underscoring the need to carefully assess the impact of language-specific lexical, structural, and pragmatic differences and seemingly innocuous variations in formality that might influence extraction patterns. On the other hand, we found no significant effect of context or dependency type, which might otherwise have been indicative of translation issues.

From a methodological point of view, one of the main take-aways of this current study is that direct crosslinguistic comparison in acceptability studies is tricky: Subtle language-specific traits may trigger unexpected and unwanted effects. Though it is difficult to compare languages directly, it is important to do so: We call for more experimental work probing the gradience in the acceptability/grammaticality of island configurations that is becoming increasingly evident, both language-internally and crosslinguistically. However, comprehension may condition extractability in ways that we do not yet fully understand and which play into a multi-factor

explanation of the results of the present study: On the one hand, the raw scores of the adjunct island extractions in Danish are both too high, in the sense that they are significantly higher than the acceptability scores for the strong islands found in the coordinate structure violations, and they are too low, in the sense that that the scores are significantly lower than the scores for corresponding sentences in English, contrary to prediction. On the other hand, we might ask whether the Danish ratings really are surprising, after all. The ratings are in the intermediate-to-lower range, which is often the case for acceptability studies. Indeed, the question might reasonably be why the English ratings are so surprisingly high.

In sum, this study offers a steppingstone toward a more refined understanding of the factors affecting extraction potential. New avenues of inquiry into the factors influencing extractability emerge, as we slowly move away from the rigid view that adjunct clauses are strong islands.

Data accessibility statement

The stimulus set used in this study, the dataset, and the R script used to analyze the data are openly available at the Open Science Framework: https://osf.io/etnrj/.

Ethics and consent

The study was granted exemption of approval by the Research Ethics Committee. According to §2.1 and §14.2 of the Danish Act on Research Ethics Review of Health Research Projects (the Committee Act. https://danskelove.dk/komitéloven/), surveys that use questionnaires and do not involve human biological material are not subject to notification of and approval from the Research Ethics Committee (see also http://en.nvk.dk). The study was completely anonymous and voluntary, it did not involve sensitive personal data or offensive material, and it did not involve any physical or mental discomfort for the participants.

Acknowledgements

Thanks to the audiences at *Linguistic Evidence* (Université Paris Cité, Oct. 6–8, 2022), *MUDS19* (Aarhus University, Oct. 13–14, 2022), *Grammatiknetværksmødet* (Dalum, Denmark, Oct. 27–29, 2022), *Dansk Funktionel Lingvistik* (Lingvistkredsen, Copenhagen University, Nov. 27, 2024), and *Syntax at the Edge* (Aarhus University, Aug. 8, 2025). In particular, we would like to extend our gratitude to Sten Vikner, Kasper Boye, Kyle Johnson, Doug Saddy, Robert Truswell, Sandra Villata, Elodie Winckel, Anne Abeillé, Liliane Haegeman, and Nomi Erteschik-Shir. We would also like to thank three anonymous reviewers for their constructive criticism and helpful comments. Finally, we want to thank Clara Ulrich Eggers for working on the stimulus sentences.

Funding information

This study was funded by the Independent Research Fund Denmark (grant ID: DFF-9062-00047B).

Competing interests

The authors have no competing interests to declare.

Author contributions

AMN supervised the project and was responsible for its administration. AMN and KRC conceptualized the project and decided on the methodology. CM, AMN, KRC constructed the stimuli. KRC took the lead on data analysis, data curation, and visualization. AMN wrote the original draft. AMN, CM, and KRC reviewed and edited the text.

ORCiD IDs

Anne Mette Nyvad: https://orcid.org/0000-0001-9554-6716

Christiane Müller: https://orcid.org/0000-0002-5385-905X

Ken Ramshøj Christensen: https://orcid.org/0000-0001-8758-240X

References

Abeillé, A., Hemforth, B., Winckel, E., & Gibson, E. (2020). Extraction from subjects: Differences in acceptability depend on the discourse function of the construction. *Cognition*, *204*, 104293. https://doi.org/10.1016/j.cognition.2020.104293

Abels, K. (2012). *Phases. An essay on cyclicity in syntax*. De Gruyter. https://doi.org/10.1515/9783110284225

Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass*, *4*(4), 187–203. https://doi.org/10.1111/j.1749-818X.2010.00193.x

Augie, B. (2016). *Package 'gridExtra'* [Computer software]. https://github.com/baptiste/gridextra

Baker, M. (2016). Is there a reproducibility crisis? *Nature*, *533*(7604), 452–454. https://doi.org/10.1038/533452a

Bondevik, I., Kush, D., & Lohndal, T. (2020). Variation in adjunct islands: The case of Norwegian. *Nordic Journal of Linguistics*, *44*(3), 1–32. https://doi.org/10.1017/S0332586520000207

Chaves, R. P. (2013). An expectation-based account of subject islands and parasitism. *Journal of Linguistics*, *49*(2), 285–327. https://doi.org/10.1017/S0022226712000357

Chaves, R. P., & Putnam, M. T. (2020). *Unbounded dependency constructions: Theoretical and experimental perspectives*. Oxford University Press. https://doi.org/10.1093/oso/9780198784999.001.0001

Chomsky, N. (1973). Conditions on transformations. In S. R. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle* (pp. 232–286). Holt, Rinehart and Winston.

Chomsky, N. (1977). On wh-movement. In P. Culicover, T. Wasow, & A. Akmajian (Eds.), *Formal syntax* (pp. 71–132). Academic Press.

Chomsky, N. (1986). *Barriers*. MIT Press.

Chomsky, N. (2000). *New horizons in the study of language and mind*. Cambridge University Press. https://doi.org/10.1017/CBO9780511811937

Christensen, K. R., & Nyvad, A. M. (2014). On the nature of escapable relative islands. *Nordic Journal of Linguistics*, *37*(1), 29–45. https://doi.org/10.1017/S0332586514000055

Christensen, K. R., & Nyvad, A. M. (2024). Complexity, frequency, and acceptability. *Glossa: A Journal of General Linguistics*, *9*(1), 1–44. https://doi.org/10.16995/glossa.10618

Christensen, R. H. B. (2015). *ordinal—regression models for ordinal data* [Computer software]. http://www.cran.r-project.org/package=ordinal/

Cinque, G. (1990). *Types of A'-dependencies*. MIT Press.

Culicover, P. W., Varaschin, G., & Winkler, S. (2022). The radical unacceptability hypothesis: Accounting for unacceptability without universal constraints. *Languages*, *7*(2), 96. https://doi.org/10.3390/languages7020096

Cuneo, N., & Goldberg, A. E. (2023). The discourse functions of grammatical constructions explain an enduring syntactic puzzle. *Cognition*, *240*, 105563. https://doi.org/10.1016/j.cognition.2023.105563

Deane, P. (1991). Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics*, *2*(1), 1–64. https://doi.org/10.1515/cogl.1991.2.1.1

Douglas, J. A. (2016). *The syntactic structures of relativisation*. [Doctoral dissertation]. University of Cambridge. https://www.repository.cam.ac.uk/handle/1810/353739

Engdahl, E. (1997). Relative clause extraction in context. *Working Papers in Scandinavian Syntax*, *60*, 59–86.

Erteschik-Shir, N. (1973). *On the nature of island constraints*. [Doctoral dissertation]. Massachusetts Institute of Technology.

Erteschik-Shir, N. (2006). What's what? In G. Fanselow, C. Fery, M. Schlesewsky, & R. Vogel (Eds.), *Gradience in grammar* (pp. 317–335). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199274796.003.0016

Erteschik-Shir, N., & Lappin, S. (1979). Dominance and the functional explanation of island phenomena. *Theoretical Linguistics* *6*(1–3), 41–86. https://doi.org/10.1515/thli.1979.6.1-3.41

Fodor, J. D. (1983). Phrase structure parsing and the island constraints. *Linguistics and Philosophy*, *6*(2), 163–223. https://doi.org/10.1007/BF00635643

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76. https://doi.org/10.1016/S0010-0277(98)00034-1

Goldberg, A. E. (2005). Island constraints and scope. In *Constructions at work: The nature of generalization in language* (pp. 129–165). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199268511.003.0007

Goldberg, A. E. (2013). Backgrounded constituents cannot be "extracted." In J. Sprouse & N. Hornstein (Eds.), *Experimental syntax and island effects* (pp. 221–238). Cambridge University Press. https://doi.org/10.1017/CBO9781139035309.012

Grosu, A. (1981). *Approaches to island phenomena*. North-Holland.

Haegeman, L. (2012). *Adverbial clauses, main clause phenomena, and the composition of the left periphery: The cartography of syntactic structures, Volume 8*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199858774.001.0001

Hale, J. (2001). A probabilistic earley parser as a psycholinguistic model. *Proceedings of NAACL*, *2*, 1–8. https://doi.org/10.3115/1073336.1073357

Hansen, E., & Heltoft, L. (2011). *Grammatik over det Danske Sprog. Bind I–III*. Det Danske Sprog- og Litteraturselskab.

Huang, C.-T. J. (1982). *Logical relations in Chinese and the theory of grammar*. [Doctoral dissertation]. Massachusetts Institute of Technology. http://pubman.mpdl.mpg.de/pubman/item/escidoc:403106:7/component/escidoc:403105/chinese_huang1982_o.pdf

Jensen, A. (2002). Sætningsknuder i dansk. *NyS, Nydanske Sprogstudier*, *29*, Article 29. https://doi.org/10.7146/nys.v29i29.13427

Keller, F. (2000). *Gradience in grammar. Experimental and computational aspects of degrees of grammaticality*. [Doctoral dissertation]. University of Edinburgh.

Kluender, R. (2004). Are subject islands subject to a processing account? *Proceedings of WCCFL 23*, 101–125. https://doi.org/10.1016/S0304-3991(04)00135-4

Kluender, R. E. (1991). *Cognitive constraints on variables in syntax*. [Doctoral dissertation]. University of California, San Diego.

Kluender, R., & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, *8*(4), 573–633. https://doi.org/10.1080/01690969308407588

Kobrock, K., & Roettger, T. B. (2023). Assessing the replication landscape in experimental linguistics. *Glossa Psycholinguistics*, *2*(1). https://doi.org/10.5070/G6011135

Kobzeva, A., Sant, C., Robbins, P. T., Vos, M., Lohndal, T., & Kush, D. (2022). Comparing island effects for different dependency types in Norwegian. *Languages*, *7*(3), Article 3. https://doi.org/10.3390/languages7030197

Kristensen, L. B., Engberg-Pedersen, E., Højlund Nielsen, A., & Wallentin, M. (2013). The influence of context on word order processing – An fMRI study. *Journal of Neurolinguistics*, *26*(1), 73–88. https://doi.org/10.1016/j.jneuroling.2012.05.001

Kuno, S. (1976). Subject, theme, and the speaker's empathy. A reexamination of relativization phenomena in subject and topic. In C. N. Li (Ed.), *Subject and topic* (pp. 417–444). Academic Press.

Kush, D., Lohndal, T., & Sprouse, J. (2018). Investigating variation in island effects: A case study of Norwegian wh-extraction. *Natural Language & Linguistic Theory*, *36*(3), 743–779. https://doi.org/10.1007/s11049-017-9390-z

Kush, D., Lohndal, T., & Sprouse, J. (2019). On the island sensitivity of topicalization in Norwegian: An experimental investigation. *Language*, *95*(3), 393–420. https://doi.org/10.1353/lan.2019.0051

Kush, D., Sant, C., & Strætkvern, S. B. (2021). Learning island-insensitivity from the input: A corpus analysis of child- and youth-directed text in Norwegian. *Glossa: a journal of general linguistics*, *6*(1), Article 1. https://doi.org/10.16995/glossa.5774

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13). https://doi.org/10.18637/jss.v082.i13

Lenth, R. V. (2022). *emmeans: Estimated marginal means, aka least-squares means*. https://CRAN.R-project.org/package=emmeans

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 1126–1177. https://doi.org/10.1016/j.cognition.2007.05.006

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419. https://doi.org/10.1207/s15516709cog0000_25

Lindahl, F. (2017). *Extraction from relative clauses in Swedish*. [Doctoral dissertation]. University of Gothenburg. https://gupea.ub.gu.se/handle/2077/51985

Lindahl, F. (2022). Extraction from relative clauses in Icelandic and Swedish: A parallel investigation. *Languages*, *7*(3), Article 3. https://doi.org/10.3390/languages7030163

Liu, Y., Winckel, E., Abeillé, A., Hemforth, B., & Gibson, E. (2022). Structural, functional, and processing perspectives on linguistic island effects. *Annual Review of Linguistics*, *8*(1), 495–525. https://doi.org/10.1146/annurev-linguistics-011619-030319

Michel, D., & Goodall, G. (2013). Finiteness and the nature of island constraints. In N. Goto, K. Otaki, A. Sato, & K. Takita (Eds.), *Proceedings of GLOW in Asia IX 2012: The main session* (pp. 187–197). https://faculty.human.mie-u.ac.jp/~glow_mie/IX_Proceedings_Oral/12Michel&Goodall.pdf

Mikkelsen, L. (2015). VP anaphora and verb-second order in Danish. *Journal of Linguistics*, *51*(3), 595–643. https://doi.org/10.1017/S0022226715000055

Müller, C. (2017). Extraction from adjunct islands in Swedish. *Norsk Lingvistisk Tidsskrift*, *35*(1), 67–85. http://ojs.novus.no/index.php/NLT/article/view/1399

Müller, C. (2019). *Permeable islands. A contrastive study of Swedish and English adjunct clause extractions*. [Doctoral dissertation]. Lund University. https://lucris.lub.lu.se/ws/portalfiles/portal/63305372/Permeable_islands_Christiane_M_ller.pdf

Müller, C., & Eggers, C. U. (2022). Island extractions in the wild: A corpus study of adjunct and relative clause islands in Danish and English. *Languages*, *7*(2), 125. https://doi.org/10.3390/languages7020125

Nyvad, A. M., Christensen, K. R., & Vikner, S. (2017). CP-recursion in Danish: A cP/CP-analysis. *The Linguistic Review*, *34*(3), 449–477. https://doi.org/10.1515/tlr-2017-0008

Nyvad, A. M., Müller, C., & Christensen, K. R. (2022). Too true to be good? The non-uniformity of extraction from adjunct clauses in English. *Languages*, *7*(4), 244. https://doi.org/10.3390/languages7040244

Pérez-Leroux, A. T., & Kahnemuyipour, A. (2014). News, somewhat exaggerated: Commentary on Ambridge, Pine, and Lieven. *Language*, *90*(3), e115–e125. https://doi.org/10.1353/lan.2014.0049

Poole, E. (2017). *Movement and the semantic type of traces*. [Doctoral dissertation]. University of Massachusetts. https://doi.org/10.7275/10687243.0

Pritchett, B. L. (1991). Head position and parsing ambiguity. *Journal of Psycholinguistic Research*, *20*(3), 251–270. https://doi.org/10.1007/BF01067218

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rizzi, L. (1982). Violations of the Wh island constraint and the subjacency condition. In *Issues in Italian syntax* (pp. 49–76). De Gruyter Mouton. https://doi.org/10.1515/9783110883718.49

Rizzi, L. (1990). *Relativized minimality*. MIT Press.

Roberts, I. (Ed.). (2016). *The Oxford handbook of universal grammar*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199573776.001.0001

Ross, H. (1987). Islands and prototypes. In A. Bosch, B. Need, & E. Schiller (Eds.), *Papers from the twenty-third regional meeting of the Chicago Linguistic Society* (pp. 309–320). Chicago Linguistic Society.

Schütze, C. T., Sprouse, J., & Caponigro, I. (2015). Challenges for a theory of islands: A broader perspective on Ambridge, Pine, and Lieven. *Language*, *91*(2), e31–e39. https://doi.org/10.1353/lan.2015.0014

Simons, M. (2007). Observations on embedding verbs, evidentiality, and presupposition. *Lingua*, *117*(6), 1034–1056. https://doi.org/10.1016/j.lingua.2006.05.006

Sprouse, J. (2007). Continuous acceptability, categorical grammaticality, and experimental syntax. *Biolinguistics*, *1*, 123–134. https://doi.org/10.5964/bioling.8597

Sprouse, J., Caponigro, I., Greco, C., & Cecchetto, C. (2016). Experimental syntax and the variation of island effects in English and Italian. *Natural Language & Linguistic Theory*, *34*(1), 307–344. https://doi.org/10.1007/s11049-015-9286-8

Sprouse, J., Wagers, M., & Phillips, C. (2012). A test of the relation between working-memory capacity and syntactic island effects. *Language*, *88*(1), 82–123. https://doi.org/10.1353/lan.2012.0004

Stepanov, A. (2007). The End of CED? Minimalism and Extraction Domains. *Syntax*, *10*(1), 80–126. https://doi.org/10.1111/j.1467-9612.2007.00094.x

Stowell, T. A. (1981). *Origins of phrase structure.* [Doctoral dissertation]. Massachusetts Institute of Technology. https://dspace.mit.edu/handle/1721.1/15626

Szabolcsi, A., & Dikken, M. den. (2003). Islands. In *Islands* (pp. 213–240). De Gruyter Mouton. https://doi.org/10.1515/9783110890952.213

Tanaka, M. (2015). *Scoping out of adjuncts: Evidence for the parallelism between QR and WH-movement.* [Doctoral dissertation]. University College London. https://discovery.ucl.ac.uk/id/eprint/1463471/

Truswell, R. (2011). *Events, phrases, and questions*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199577774.001.0001

Tutunjian, D., Heinat, F., Klingvall, E., & Wiklund, A.-L. (2017). Processing relative clause extractions in Swedish. *Frontiers in Psychology*, *8*, 2118. https://doi.org/10.3389/fpsyg.2017.02118

Van Valin, R. D. J. (1996). Toward a functionalist account of so-called extraction constraints. In B. Devriendt, L. Goosens, & J. Van der Auwera (Eds.), *Complex structures. A functionalist perspective* (pp. 29–60). De Gruyter Mouton. https://doi.org/10.1515/9783110815894.29

Van Valin, R. D. J. (2005). *Exploring the syntax-semantics interface*. Cambridge University Press. https://doi.org/10.1017/CBO9780511610578

Vikner, C. (2004). The semantics of Scandinavian 'when'-clauses. *Nordic Journal of Linguistics*, *27*(2), 133–167. https://doi.org/10.1017/S0332586504001209

Vikner, S., Christensen, K. R., & Nyvad, A. M. (2017). V2 and cP/CP. In L. Bailey & M. Sheehan (Eds.), *Order and structure in syntax I: Word order and syntactic structure* (pp. 313–324). Language Science Press. https://doi.org/10.5281/zenodo.1117724

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* [Computer software]. Springer. https://ggplot2.tidyverse.org

Wiklund, A.-L., Heinat, F., Klingvall, E., & Tutunjian, D. (2017). An acceptability study of long-distance extractions in Swedish. In L. Escobar, V. Torrens, & T. Parodi (Eds.), *Language processing and disorders* (pp. 103–120). Cambridge Scholars Publishing.

Wiklund, A.-L., Hrafnbjargarson, G. H., Bentzen, K., & Hróarsdóttir, Þ. (2007). Rethinking Scandinavian verb movement. *The Journal of Comparative Germanic Linguistics*, *10*(3), 203–233. https://doi.org/10.1007/s10828-007-9014-9

Williams, E. (2011). *Regimes of derivation in syntax and morphology*. Routledge. https://doi.org/10.4324/9780203830796

Winckel, E., Abeillé, A., Hemforth, B., & Gibson, E. (2025). Discourse-based constraints on long-distance dependencies generalize across constructions in English and French. *Cognition*, *254*, 105950. https://doi.org/10.1016/j.cognition.2024.105950