ON THE ISLAND SENSITIVITY OF TOPICALIZATION IN NORWEGIAN: AN EXPERIMENTAL INVESTIGATION

DAVE KUSH | TERJE LOHNDAL | JON SPROUSE

NTNU Norwegian University of Science and Technology and Haskins Laboratories | *NTNU Norwegian University of Science and Technology and UiT The Arctic University of Norway* | *University of Connecticut*

Mainland Scandinavian languages have been reported to allow movement from embedded questions, relative clauses, and complex NPs—domains commonly considered to be islands crosslinguistically. Yet in formal acceptability studies Scandinavian participants often show 'island effects': they reject island-violating movement similarly to native speakers of 'island-sensitive' languages. To investigate this apparent mismatch between informal and formal judgments, we conducted two acceptability judgment experiments testing the acceptability of topicalization from various island domains in Norwegian. We were interested in determining whether we could (i) find evidence for island insensitivity and (ii) pin down the source of qualitatively different island effects. We asked whether such effects are best explained as reflecting violations of a uniform syntactic constraint or extrasyntactic factors. Our results suggest that embedded questions and relative clauses are not uniform syntactic islands for topicalization, but complex NPs are. Unexpectedly, we also found evidence suggesting that conditional adjunct clauses may not be islands.*

Keywords: island effects, wh-movement, Norwegian, topicalization, experimental syntax, crosslinguistic variation

1. INTRODUCTION. Mainland Scandinavian languages like Norwegian, Swedish, and Danish are often reported to be typological outliers when it comes to A′-movement, allowing dependencies to cross domains that are islands in other languages. The domains commonly discussed are EMBEDDED QUESTIONS (WH-islands), RELATIVE CLAUSES (RCs), and, to a lesser extent, COMPLEX NPs (CNPs). Attested examples of such island-violating movement dependencies are given in 1.

(1) a. WH-movement from an embedded question (Norwegian)
  Hvilke bøker spurte Jon hvem som hadde skrevet __ ?
  which books asked Jon who C had written
   'Which books did Jon ask who had written?' (Maling & Zaenen 1982:232)

  b. DP topicalization from a relative clause (Norwegian)
  Rødsprit slipper vi ingen inn [RC som har drukket __].
  Red.spirit let we no.one in that has drunk
   'Redspirit, we don't let anyone in who has drunk.' (Taraldsen 1982, ex. 11)

  c. WH-movement from a complex NP (Swedish)
  Vem levde Herodes i [CNP hoppet om att Salome skulle
  who lived Herod in hope.DEF about that Salome would
    förföra __]?
    seduce
   'Who did Herod live in the hope that Salome would seduce?'

(Allwood 1982:19)

Many researchers have argued that such examples demonstrate that movement from embedded questions, RCs, and CNPs is, as a general rule, syntactically licit in Mainland Scandinavian (Allwood 1982, Christensen 1982, Engdahl 1982, 1997, Maling & Zaenen 1982, Lindahl 2014, 2017, Nyvad, Christensen, & Vikner 2017, among others).

However, although acceptable island violations can apparently be found in naturalistic discourse (see e.g. Engdahl 1997, Lindahl 2017), it has proven challenging to verify that such dependencies are uniformly acceptable using formal experimental methods. Mainland Scandinavian participants often rate such long-distance dependencies as less acceptable than might be expected if the domains were not islands for movement. That is, judgments of these dependencies seem to display 'island effects'.

Two qualitatively and quantitatively distinct judgment patterns have been categorized as 'island effects' in previous experiments. We refer to the first response pattern as a CLASSIC ISLAND EFFECT. Classic island effects occur when movement from an island domain is not only significantly less acceptable than movement from a comparable control condition, but also sharply unacceptable on the absolute scale. The second response pattern has been called a SUBLIMINAL ISLAND EFFECT (Almeida 2014, Keshev & Meltzer-Asscher 2019). Subliminal island effects are also characterized by a significant difference between island-violating sentences and their nonisland counterparts, but they differ in that island-violating sentences are judged, on average, to be somewhere within the range of acceptability.

Experiments that have investigated the acceptability of movement from RCs and CNPs have consistently found classic island effects. Christensen and Nyvad (2014) found that movement out of RCs in Danish was rated low on average. In Kush, Lohndal, & Sprouse 2018, we investigated the sensitivity of Norwegian WH-movement to various islands using the factorial design used by Sprouse and colleagues (Sprouse 2007, Sprouse et al. 2011, Sprouse, Wagers, & Phillips 2012, and Sprouse et al. 2016). We observed that WH-movement from both RCs and CNPs was rated to be as unacceptable as uncontroversial subject island violations were. Participants rejected such dependencies outright.

Subliminal island effects have been observed with movement from WH-islands. Christensen, Kizach, and Nyvad (2013) found that Danish participants rated WH-dependencies that spanned embedded questions significantly lower than WH-dependencies that spanned embedded declaratives. However, the mean acceptability of WH-island-violating sentences was also significantly higher than other unambiguous syntactic violations. Kush et al. 2018 found that although WH-movement from an embedded question was rated as significantly less acceptable on average than movement from an embedded declarative, the resulting island effect was considerably smaller than other island effects (e.g. CNP, RC, subject, and adjunct islands) in the same experiment. Moreover, the average acceptability rating of WH-island-violating sentences consistently fell above the midpoint of the rating scale—the range typically reserved for acceptable sentences.

The source of these different island effects remains unclear, but there are, broadly speaking, three possible explanations. First, contrary to prior claims, embedded questions, RCs, and CNPs could be UNIFORM SYNTACTIC ISLANDS in Mainland Scandinavian, just as in other languages. Movement would never be allowed, and the apparent acceptability of the naturalistic examples would be 'illusory'. A second possibility is that the island effects in past experiments reflect SYNTACTIC AMBIGUITY. Under this possibility, the domains in question could have at least two underlying analyses: one an island, the other a nonisland (see e.g. Kush et al. 2018, Sichel 2018). Average island effects would emerge if participants tend to erroneously adopt the island parse when presented with a sentence in an experimental setting. Finally, the observed acceptability decrease could reflect EXTRASYNTACTIC factors. According to proponents of extrasyntactic theories, these island effects emerge when semantic, pragmatic, or processing factors conspire to make an otherwise syntactically well-formed sentence seem less ac-

ceptable. A number of authors endorse some version of an extrasyntactic explanation for the unacceptability of the dependencies in question. Engdahl (1997) argues that the apparent variable acceptability of RC island violations in Swedish is primarily determined by discourse-pragmatic and information-structural conditions (see also Allwood 1982, Andersson 1982, Christensen 1982, Lindahl 2017). Alternatively, Christensen et al. (2013) argued that the unacceptability of extraction from wh-islands in Danish should be attributed to processing effects.

Experimental arguments against the uniform syntactic explanation are typically made when subliminal island effects are observed (Kush et al. 2018, Keshev & Meltzer-Asscher 2019). In some cases arguments appeal to distributional patterns in participant judgments that are assumed to be less likely under a (categorical) syntactic approach to island effects.

Arguments made on the basis of such SECOND-ORDER ACCEPTABILITY EFFECTS have been made to explain subliminal island effects in Danish and Norwegian wh-islands. For example, Christensen et al. (2013) argued for a processing-based explanation of their Danish wh-island effects on the basis of a weak positive correlation between acceptability of long-distance wh-movement and order of presentation in their experiments: judgments of wh-movement from both embedded declaratives (nonislands) and embedded questions (wh-islands) improved over the course of their experiments. Similar amelioration was not observed for ungrammatical controls. The authors argued that such amelioration was inconsistent with the (often implicit) assumption that amelioration can occur only with grammatical structures (as in Sprouse 2007). Given the amelioration, they argued that embedded questions are not islands in Danish. In a similar vein, Kush et al. 2018 showed that apparent subliminal *whether*-island effects in Norwegian arise as an artifact of averaging over highly variable judgment distributions in the island-violating condition. Many individual participants show no island effects whatsoever, consistently accepting wh-movement from *whether*-islands, while others judge the same dependencies INCONSISTENTLY across trials, sometimes accepting the dependencies and other times rejecting them. We suggested that such variability was more easily accommodated under either ambiguity-based or extrasyntactic accounts than under an account where subliminal island effects result from a gradient syntactic constraint violation (contra Almeida 2014). This line of reasoning presupposes that a syntactic constraint violation should have a consistent negative effect on acceptability across multiple exposures.[1]

Second-order acceptability arguments are difficult to make for classic island effects where participants consistently reject movement from islands, since such a response pattern is, in principle, equally consistent with all three accounts. Thus, second-order acceptability arguments have not previously been made for the RC and CNP island effects observed in previous studies.

It is possible that classic RC and CNP island effects emerged because participants reliably chose an inappropriate island-parse of the domains even when an alternative parse was available, or due to extragrammatical factors that prior researchers failed to control for when constructing their materials. Felicitous movement from RCs and

---

[1] We do not necessarily suppose, however, that participants should report the same degree of subjective unacceptability across multiple exposures to the same violation. It is enough that participants would—all else being equal—reliably judge the sentences to be less acceptable than comparison sentences that do not violate the same constraint.

CNPs may, for instance, be subject to currently unknown semantic or pragmatic conditions that either affect the probability of selecting a nonisland parse in an experiment or constrain the resulting dependencies directly. Should this be the case, one would expect Norwegian speakers to give different judgments when these conditions are met.

Given that a precise formalization of the relevant extrasyntactic factors remains elusive, one option for indirectly investigating the source of classic island effects is to test whether second-order acceptability effects emerge when different A′-dependency types are tested. The logic behind such an experiment is as follows. A′-dependencies such as wh-question formation, topicalization, and relative clause formation are presumed to be subject to the same syntactic constraints, but different semantic and pragmatic restrictions.

By testing different dependency types we can evaluate whether the effects observed in previous studies are likely to reflect a wholesale ban on A′-movement in general (as predicted by uniform syntactic accounts) or whether they arise because of the accidental convergence of extrasyntactic factors relating to the dependency used in previous experiments.

With this goal in mind we continued the search for second-order acceptability effects by testing the same island types as we did in Kush et al. 2018, but using a different A′-dependency: topicalization. Topicalization has been claimed to be subject to the same syntactic locality constraints as wh-movement, but different semantic and discourse-pragmatic factors. Thus, we reason that if RC and CNP island violations were judged to be unacceptable due to the accidental convergence of extrasyntactic factors specific to wh-movement, then we are less likely to trigger those specific semantic or discourse-pragmatic factors with topicalization, and therefore more likely to observe a different pattern of effects.[2]

2. Experiments. We ran two acceptability judgment studies to test the island-sensitivity of topicalization in Norwegian using materials adapted from Kush et al. 2018. The experiments used the factorial design for investigating island effects originally due to Sprouse 2007, which has since been used in several crosslinguistic investigations (Sprouse et al. 2011, Sprouse et al. 2012, Almeida 2014, Sprouse et al. 2016, Kush et al. 2018, Tucker et al. 2019). Using the same design as previous experiments facilitates cross-experimental (and ultimately crosslinguistic) comparison. In the current case, it also makes it possible to compare differences between topicalization and wh-movement in Norwegian.

2.1. Experiment 1: topicalization (no context). Experiment 1 tested the acceptability of DP topicalization dependencies spanning the five island types tested in Kush et al. 2018. Crucially, the test sentences in this experiment were presented in isolation, without any context to support the topicalization.

Participants. Thirty-five native Norwegian speakers (mean age 42.6, twenty-two female) were recruited via public announcement on various social media sites. In addition to providing their age and gender, participants answered a small questionnaire on their language/dialect background. We excluded the data of two participants who reported being raised bilingual.

Materials. Materials were adapted from the test sentences used in experiment 3 of Kush et al. 2018, which were composed of a matrix clause, an embedded clause, and a

---

[2] There is some indirect, though highly suggestive, evidence that judgments of topicalization may differ from wh-movement: a survey of the literature reveals that most attested examples of extraction from RCs and CNPs involve topicalization, rather than wh-movement (Christensen 1982, Taraldsen 1982, Engdahl 1997, Lindahl 2017). Observe 1b above.

left-peripheral wh-phrase. Items followed a 2 × 2 factorial design that crossed the factors STRUCTURE and DISTANCE. Structure controlled whether the embedded clause was an ISLAND or a NONISLAND. Distance controlled whether the wh-phrase originated in the matrix clause (SHORT conditions) or in the embedded clause (LONG conditions). An example of a *whether*-island item is given below.

(2) Example *whether*-island item from Kush et al. 2018

    a. Hvilken gjest tror   at Hanne bakte kaken?
       which   guest believes that Hanne baked cake.DEF
         'Which guest believes that Hanne baked the cake?'   (Short-noIsland)

    b. Hvilken gjest lurer   på om   Hanne bakte kaken?
       which   guest wonders on whether Hanne baked cake.DEF
         'Which guest wonders whether Hanne baked the cake?'   (Short-Island)

    c. Hvilken kake tror   gjesten   at Hanne bakte?
       which   cake believes guest.DEF that Hanne baked
         'Which cake does the guest believe that Hanne baked?'   (Long-noIsland)

    d. Hvilken kake lurer   gjesten   på om   Hanne bakte?
       which   cake wonders guest.DEF on whether Hanne baked
         'Which cake does the guest wonder whether Hanne baked?'

                                                 (Long-Island)

Within this design a significant (negative) Structure × Distance interaction effect indicates that participants judged the Long-Island condition to be less acceptable, on average, than the Short-noIsland condition; this difference is more than can be explained by simply subtracting the average acceptability penalty for having a longer dependency or a more complex syntactic structure (see Sprouse 2007, Sprouse et al. 2012 for more detailed discussion). That is, there is residual unacceptability that must be accounted for: there is an island effect. We use the simple interaction logic to identify island effects. If we observe an interaction with a particular island, we have evidence for an island effect. If the interaction does not achieve significance, there is no evidence of an island effect. We point out, however, that the mere presence of a superadditive interaction effect does not automatically determine the source of that effect. Interaction effects themselves could arise equally well under the uniform syntactic, syntactic-ambiguity, or extrasyntactic analyses discussed above.

We converted the wh-movement items from Kush et al. 2018 to topicalization by replacing the left-peripheral wh-phrase with a referential DP: the wh-phrase was replaced with a pronoun in short conditions (*han* 'he' in 3 below), and with a full lexical DP in long conditions (*kaken* 'the cake' in 3). This simple conversion had one notable consequence for the structure of our factorial design. In the wh-movement items the wh-phrase in short conditions was assumed to move to the specifier of the matrix CP. The syntax of the short conditions was therefore marked in a way that was analogous to long conditions (both contained evidence of wh-movement). In the current experiment, the pronouns in the short conditions are in sentence-initial position. Because this corresponds to the default subject position in declaratives, the word order in these test sentences is unmarked.

Given that there is less obviously movement in the erstwhile short conditions, keeping 'Movement' as a factor name in our experiment might seem misleading.[3] We therefore label the second factor in our experiments TOPICALIZATION. The factor had two levels, NO TOPICALIZATION and TOPICALIZATION, corresponding to the old levels short

---

[3] Though preserving the naming convention from Kush et al. 2018 would arguably be appropriate, since it is relatively standard to assume that the matrix subject undergoes raising to a left-peripheral position in SVO

and long. Given the logic of the superadditive design, an island effect could still be defined as the superadditive interaction of the two factors: Structure × Topicalization.

We provide example items for each island type. An example *whether*-island item is provided in 3. One other change was made to the *whether*-island template from above: the Norwegian equivalent of the verb 'to wonder', *å lure*, does not take a CP complement directly. Instead, the CP is complement to the preposition *på* 'on'. Thus, in the example item above, island and no-island conditions differ not only in their embedding predicate ('believe' vs. 'wonder'), but also in the immediate syntactic environment of the embedded clause: in island conditions there is an extra level of phrasal embedding not present in the no-island condition. One might worry that such structural mismatches confound the ability to cleanly isolate an island effect. In Kush et al. 2018 we found no evidence that such a difference had a confounding effect, but we nevertheless strove to increase structural similarity across conditions here. To achieve this, embedding predicates in no-island conditions were changed to include a preposition in every item where *å lure på* was used as the embedding predicate for island conditions. Thus, *tror* 'believes' from 2 was changed to *er sikker på* 'is sure (on)'.[4]

The reader will note that in all topicalization conditions (for all island types), the topicalized DP is followed by the finite matrix verb since Norwegian is a V2 language (e.g. Holmberg & Platzack 1995). In these conditions the matrix subject comes directly after the verb.

(3) *Whether*-island
  a. Han er sikker på at   Hanne bakte kaken.
     he   is sure   on that Hanne baked cake.DEF
       'He is sure that Hanne baked the cake.'   (noTop-noIsland)
  b. Han lurer   på om   Hanne bakte kaken.
     he   wonders on whether Hanne baked cake.DEF
       'He wonders whether Hanne baked the cake.'   (noTop-Island)
  c. Kaken   er han sikker på at   Hanne bakte   __ .
     cake.DEF is he   sure   on that Hanne baked
       'The cake he is sure that Hanne baked.'   (Topic-noIsland)
  d. Kaken   lurer   han på om   Hanne bakte   __ .
     cake.DEF wonders he   on whether Hanne baked
       'The cake he wonders whether Hanne baked.'   (Topic-Island)

An example subject island item is given in 4.

(4) Subject island
  a. Hun sa   det faste avfallet   forurenser sjøen.
     she said the solid waste.DEF pollutes   sea.DEF
       'She said the solid waste is polluting the sea.'   (noTop-noIsland)
  b. Hun sa   det faste avfallet   fra fabrikken   forurenser sjøen.
     she said the solid waste.DEF from factory.DEF pollutes   sea.DEF
       'She said the solid waste from the factory is polluting the sea.'

                                                              (noTop-Island)

---

clauses in Norwegian when no other phrase occupies the left periphery (Holmberg & Platzack 1995, Åfarli & Eide 2003, though see Westergaard, Lohndal, & Alexiadou 2016).
  [4] Analogous changes were made for materials in other island subexperiments.

    c. [Det faste avfallet]$_i$ sa   hun   __ $_i$ forurenser sjøen.
       the solid waste.DEF said she   pollutes   sea.DEF
       'The solid waste she said is polluting the sea.'   (Topic-noIsland)

    d. Fabrikken$_i$ sa   hun det faste avfallet fra   __ $_i$ forurenser sjøen.
       factory.DEF said she the solid waste from   pollutes   sea.DEF
       'The factory she said the solid waste from is polluting the sea.'
                                             (Topic-Island)

An adjunct island item is given in 5. In constructing the island conditions (5b,d) we made sure that the matrix predicate was intransitive, so that the topicalized DP (*bakdøren* 'the back door') could not be misinterpreted as the object of the matrix predicate in the topicalization-island condition. This was to avoid the possibility of licensing potential parasitic-gap readings.

(5) Adjunct island (conditional 'if')

    a. Han mistenker at   de   lar   bakdøren   stå   ulåst.
       he   suspects   that they leave back.door.DEF stand unlocked
       'He suspects that they leave the back door unlocked.'   (noTop-noIsland)

    b. Han blir nervøs om de   lar   bakdøren   stå   ulåst.
       he   gets nervous if   they leave back.door.DEF stand unlocked
       'He gets nervous if they leave the back door unlocked.'   (noTop-Island)

    c. Bakdøren   mistenker han at   de   lar   stå   ulåst.
       back.door.DEF suspects   he   that they leave stand unlocked
       'The back door he suspects that they leave unlocked.'   (Topic-noIsland)

    d. Bakdøren   blir han nervøs om de   lar   stå   ulåst.
       back.door.DEF gets he nervous if   they leave stand unlocked
       'The back door he gets nervous if they leave unlocked.'   (Topic-Island)

An example RC island item is in 6. All RC items used indefinite or weak quantificational RC heads, since naturalistic examples of extraction feature such RC heads (Allwood 1982, Engdahl 1997, Lindahl 2017). We selected embedding predicates that are also attested in naturalistic examples (Engdahl 1997, Lindahl 2017).

(6) RC island

    a. De tror   at mange psykologer   ville   anbefale
       they believe that many   psychologists would recommend
          antidepressiva.
          antidepressants
       'They believe that many psychologists would recommend antidepressants.'   (noTop-noIsland)

    b. De kjenner til mange psykologer   som ville   anbefale
       they know   to many   psychologists who would recommend
          antidepressiva.
          antidepressants
       'They know (of) many psychologists who would recommend antidepressants.'   (noTop-Island)

    c. Antidepressiva tror   de at mange psykologer   ville   anbefale.
       antidepressants believe they that many   psychologists would recommend
       'Antidepressants they believe that many psychologists would recommend.'   (Topic-noIsland)

    d. Antidepressiva kjenner de   til mange psykologer   som ville
       antidepressants know   they to many   psychologists who would
          anbefale.
          recommend
      'Antidepressants they know (of) many psychologists who would rec-
      ommend.'   (Topic-Island)

An example CNP island item is given in 7.

    (7) CNP island
        a. De kunne rapportere at   han vant løpet.
          they could report   that he won race.DEF
            'They could report that he won the race.'   (noTop-noIsland)
        b. De kunne rapportere nyheten om   at   han vant løpet.
          they could report   news.DEF about that he won race.DEF
            'They could report the news that he won the race.'   (noTop-Island)
        c. Løpet   kunne de   rapportere at   han vant.
          race.DEF could they report   that he won
            'The race they could report that he won.'   (Topic-noIsland)
        d. Løpet   kunne de   rapportere nyheten om   at   han vant.
          race.DEF could they report   news.DEF about that he won
            'The race they could report the news that he won.'   (Topic-Island)

PROCEDURE. The experiments were conducted using IbexFarm (Drummond 2012). Participants completed the task on their own personal computers. Items were presented one at a time, centered on the screen. Participants were instructed to judge the acceptability of each item on a seven-point scale located below the item. The endpoints of the scale were labeled 1 *Dårlig* ('bad') and 7 *Bra* ('good'). Every survey contained forty test items. For each of the five island types participants rated eight sentences, two from each condition (2 tokens × 4 conditions × 5 island types). Test items were interspersed pseudo-randomly among forty-eight filler sentences (sixteen acceptable, thirty-two unacceptable) adapted from Kush et al. 2018. At the beginning of the experiment each participant was presented with an identical set of five filler items that varied from unacceptable to acceptable. These items were included to give participants the same opportunity to 'calibrate' their judgments before ever encountering a test item.

ANALYSIS. We *z*-score transformed raw ratings by participant before statistical analysis, which was conducted using linear mixed-effects models using the lme4 package (Bates et al. 2015) in R (R Core Team 2018). Separate models were constructed for each island type with fixed effects of Structure, Topicalization, and the Structure × Topicalization interaction. Models included random intercepts for both subject and item and by-subject random slopes for all fixed effects and their interaction. In the rare event that a model did not converge, we simplified the random-effects structure such that it included a by-subject random slope only for the Structure × Topicalization interaction. The Satterthwaite approximation was used to calculate *p*-values using the lmerTest package (Kuznetsova, Brockhoff, & Christensen 2017).

Although we fit the full models described above for all of our analyses, we report statistical summaries only of the interaction effects, rather than the full model readouts for each island. We made this choice for considerations of space and because main effects of Structure and Topicalization are ultimately immaterial to our main questions of interest—that is, whether there is an irreducible superadditive 'island effect'. In addition to reporting the significance of the interaction term from each model, we also report the

DIFFERENCES-IN-DIFFERENCES (DD) score (Maxwell & Delaney 2003) for each interaction/island effect. DD scores provide a standardized quantitative measure of island effect size, which can be useful for cross-island or crosslinguistic comparison. DD scores of zero, or close to zero, can be interpreted as an absence of an island effect. Although there is, in principle, no quantitative threshold for defining a 'true' island effect, prior studies using the factorial design have found that DD scores for island effects typically fall within the range of 0.75–1.25 and there is often little variation in average effect size across islands (Sprouse et al. 2012, Sprouse et al. 2016). Given the consistency in DD scores across islands in prior results, any intermediate-sized island effect bears closer scrutiny.

RESULTS AND DISCUSSION. Participants rated acceptable filler sentences highly (mean rating $z = 0.92$) and unacceptable filler sentences low (mean rating $z = -0.72$) on average, indicating that they understood and performed the task as expected. Figure 1 plots the mean z-scored ratings for each experimental condition by island type. A statistical summary of the island effects is given in Table 1.
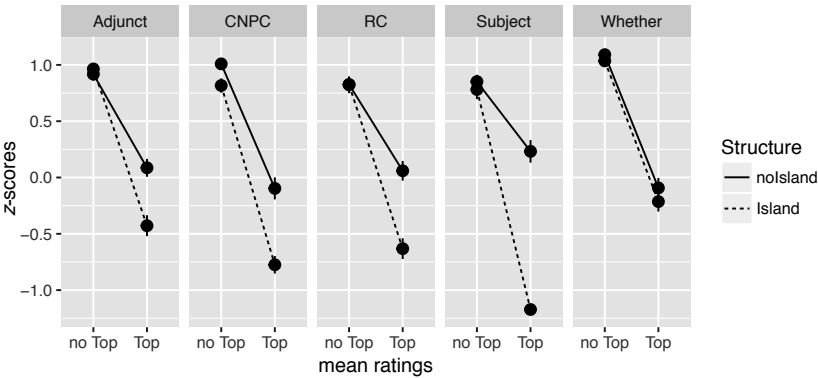

FIGURE 1. Interaction plots for all island types in experiment 1. Error bars indicate one standard error.

	t-VALUE	SIGNIFICANCE	EFFECT SIZE (DD)
ADJUNCT ISLAND	−4.55	$p < 0.001$	0.55
COMPLEX NP	−3.57	$p = 0.001$	0.51
RELATIVE CLAUSE	−4.09	$p < 0.001$	0.74
SUBJECT ISLAND	−8.85	$p < 0.001$	1.33
WHETHER-ISLAND	−0.35	$p = 0.642$	0.05

TABLE 1. Statistical summary of island effects from experiment 1.

Significant superadditive island effects were observed for four of the five islands tested: adjunct, CNP, RC, and subject islands. We failed to observe a statistically significant *whether*-island effect. The absence of an island effect is confirmed by visual inspection of Fig. 1, where there is no hint of an interaction. Both the strong subject island effect and the absence of a *whether*-island effect are largely in line with the Kush et al. 2018 WH-movement findings.

Average effect sizes for adjunct, CNP, and RC islands were smaller than the subject island effect and their corresponding effects in Kush et al. 2018 (DDs < 1.20). The diminished effect sizes can be partially attributed to the relatively low average ratings of the topicalization-noIsland conditions. The cause of the lower average ratings must be extrasyntactic, since long-distance topicalization is unquestionably grammatical in

Norwegian. Extrasyntactic unacceptability could stem from either parsing difficulty, semantic or discourse-pragmatic infelicity, or some combination thereof. We doubt that simple parsing-related factors like dependency length caused the acceptability decrement, or else the same effect should have been found for WH-movement in Kush et al. 2018, contrary to fact. Other parsing-related difficulties might have reduced average ratings. For example, it is conceivable that some of the unacceptability is the result of a temporary misanalysis: if participants occasionally misanalyzed the topicalized DP as the main subject of the topicalization-island sentences, reanalysis would have been required at the true pronominal subject after the verb. Reanalysis has been shown to have lingering effects on acceptability (e.g. Sprouse 2008). As for semantic or discourse-pragmatic effects, participants may have reacted adversely to the stimuli because they seemed 'unnatural' in the absence of accompanying context. Although the cause of these effects is of interest, it should be stressed that the lower ratings do not completely confound our ability to conclude that adjunct, CNP, and RC island violations are less acceptable than ordinary long topicalization in a way not explained by simple distance and complexity effects alone.

In Kush et al. 2018 we argued that it is insufficient to draw conclusions about the origin of island effects based on aggregate means alone, because they obscure the second-order effect of judgment variability. Inspection of rating distributions provides information about variability, which can in turn yield insight into (i) why the average scores were so low for grammatical long-distance topicalization, and (ii) whether interactions are driven by inconsistent ratings in the topicalization-island conditions.

To address these two possibilities, we looked at the distribution of ratings by condition (see Figure 2). Unusual variability can be identified through comparison to judgments in 'anchor' conditions where we expect participants to consistently either accept or reject the sentences. When participants give consistent judgments, z-scores are distributed UNIMODALLY about the population mean. Roughly unimodal distributions are seen across the no-topicalization conditions for all island types, where the mode of approximately +1 connotes unanimous acceptability. Judgments in the topicalization-island condition for subject islands are also unimodally distributed below −1, reflecting that participants reliably rejected subject island violations. These then can be seen as archetypical cases against which the other distributions can be evaluated.

We first turn to the cause of the low ratings in topicalization-noIsland conditions. Figure 2 makes clear that judgments of long-distance topicalization are characterized by significant intertrial variability. Z-scores in all five topicalization-noIsland conditions are either bimodally or near uniformly distributed across the range.[5] Thus, the average that we observe is best understood as a mixture over inconsistent judgments, many of which appear to be categorical—either full acceptance or rejection.

We turn next to ratings of island violations. First, z-scores in the *whether* topicalization-island condition are nearly indistinguishable from the corresponding topicalization-noIsland condition, further reinforcing the conclusion that participants do not perceive topicalization from embedded questions as different from declarative complement clauses. Second, though topicalization from CNPs and RCs was outright rejected on most trials, the probability of participants accepting such sentences was not zero. Topicalization from RCs was judged to be totally acceptable on a small number of trials. Such judgment patterns could indicate extrasyntactic variability, but it is difficult to

---

[5] These data rule out one possible interpretation of the reduced average acceptability discussed above: it does not reflect a consistent assessment of long-distance topicalizations as 'middling'—as might be argued by proponents of gradient syntactic constraints (e.g. Keller 2000, Featherston 2007, Bresnan & Hay 2008).
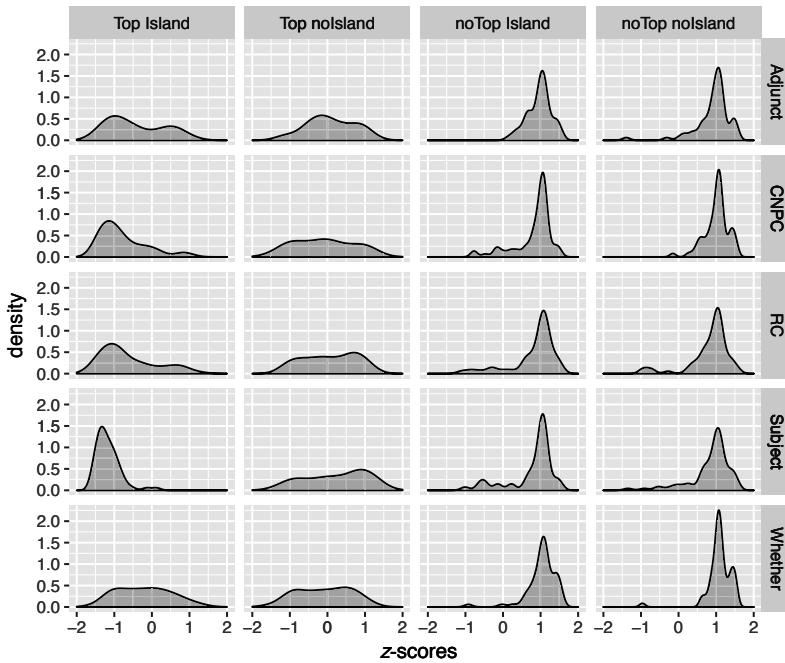
FIGURE 2. Ratings distributions from experiment 1 by island and condition.

determine whether the pattern here differs reliably from noise. Finally, and perhaps most surprisingly, Fig. 2 shows that *z*-scores in the adjunct topicalization-island condition may be bimodally distributed. Moreover, topicalization from adjuncts was judged to be acceptable on a nontrivial number of trials. This raises the possibility that conditional adjuncts are not structural islands in Norwegian.

INTERIM SUMMARY AND DISCUSSION OF RESULTS. We found significant Structure × Topicalization interaction effects for topicalization from adjunct, CNP, RC, and subject islands. There was not a significant Structure × Topicalization interaction for topicalization from *whether*-islands, indicating the absence of an island effect.

Long-distance topicalization from nonislands was judged lower on average than might have been expected for a grammatical transformation. Analysis of the judgment distributions in these conditions revealed that the lower mean ratings reflected an average over highly variable responses ranging from full acceptance to outright rejection. We suggested that this kind of variability reflects the influence of nonsyntactic factors on the judgment process, not that long-distance topicalization is ungrammatical in Norwegian.

We further speculated that if we saw similar variability (or a bimodal response pattern) in judgments of island violations, then the unacceptability should be attributed to an extrasyntactic source, rather than a syntactic constraint violation. The speculation was based on the fact that most past studies have found little to no evidence for inter- or intraindividual variability in judgments for most islands (Sprouse et al. 2012, Sprouse et al. 2016, Kush et al. 2018).

Ratings in the *whether* topicalization-island condition showed the same degree and pattern of variability as long-distance topicalization from nonislands, consistent with the conclusion that topicalization from embedded *whether*-questions is perceived as indistinguishable from standard long-distance topicalization. Therefore, we conclude that embedded *whether*-questions are not islands in Norwegian.

Experiment 1 uncovered one intriguing and unexpected result: judgments of adjunct island violations were bimodally distributed. On many trials participants appeared to judge topicalization from a conditional adjunct to be acceptable, but on others they rejected the dependencies. Such variability, particularly the number of categorical acceptances, might be taken as evidence that conditional adjuncts are not uniform syntactic islands. Under such a view, what unacceptability we observe would be rooted in extrasyntactic factors (e.g. violable semantic/discourse-pragmatic conditions that are difficult but not impossible to accommodate *in vacuo*). If variability is indeed due to such factors, we predict that distributions should shift toward unimodality if we provided supporting contexts that facilitate accommodation.

Analysis of the remaining judgment distributions revealed a small degree of variability in the judgments of RC and CNP island violations: participants occasionally judged topicalization from these domains to be acceptable. However, the pattern was not widespread enough to rule out the uniform syntactic explanation of the average unacceptability of the conditions. Although participant judgments of RC and CNP island sentences were less variable than the adjunct or *whether*-island sentences, we cannot automatically conclude that the two constructions are uniform syntactic islands. Rejection could still reflect structural ambiguity or extrasyntactic factors. The unexpected variability in the grammatical topicalization-noIsland conditions attests to the persistent influence of extrasyntactic considerations on the acceptability of long-distance topicalization. If extrasyntactic factors could cause participants to reject topicalization from simple embedded clauses, it is possible that similar factors played a role in their rejecting topicalization from more complex structures.

If participants rejected some or all topicalization from RCs or CNPs on extrasyntactic grounds, then judgments of such topicalizations might change when extrasyntactic factors are manipulated. We tested this possibility in experiment 2 by adding contextual support to each of the items in the experiment.

2.2. Experiment 2: DP topicalization with context. Experiment 2 tested the same sentences as in experiment 1, but each sentence was paired with contextual support.

Participants. Thirty-six native Norwegian speakers (mean age 33.2, twenty-five female) were recruited and screened as for experiment 1. We excluded the data of four participants who reported being raised bilingual.

Materials. Test items were adapted from the items in experiment 1. Test sentences from the previous experiment were augmented so that they were preceded by a contextual preamble intended to facilitate the use of topicalization in the test sentence. Preambles established a mini-discourse in which the potentially fronted DP in the test sentence could be interpreted contrastively. Preambles also provided an antecedent for the subject pronoun in the matrix clause of the test sentence. Example items from the CNP and RC island subexperiments are provided below. A full list of materials is available on the project's Open Science Framework site: https://osf.io/k68ew/.

(8) CNP island
　　Preamble:
　　Journalistene　visste ikke om Anders hadde vunnet hele　turneringen, …
　　journalists.DEF knew NEG if　Anders had　won　whole tournament.DEF
　　　'The journalists didn't know if Anders had won the whole tournament, … '
　　a. men de　kunne rapportere at　han vant løpet.
　　　but　they could report　　that he　won race.DEF
　　　　'but they could report that he won the race.'

  b. men de kunne rapportere nyheten om at han vant løpet.
    but they could report news.DEF about that he won race.DEF
     'but they could report the news that he won the race.'
  c. men løpet kunne de rapportere at han vant.
    but race.DEF could they report that he won
     'but the race they could report that he won.'
  d. men løpet kunne de rapportere nyheten om at han vant.
    but race.DEF could they report news.DEF about that he won
     'but the race they could report the news that he won.'

(9) RC island
  Preamble:
  Forsvareren bekreftet at alle ville anbefale   paracetamol, …
  defender.DEF confirmed that all would recommend paracetamol
   'The lawyer confirmed that everyone would recommend paracetamol, … '
  a. men de tror at mange psykologer ville anbefale
    but they believe that many psychologists would recommend
     antidepressiva.
     antidepressants
     'but they believe that many psychologists would recommend antide-
     pressants.'
  b. men de kjenner til mange psykologer som ville anbefale
    but they know to many psychologists who would recommend
     antidepressiva.
     antidepressants
     'but they know (of) many psychologists who would recommend anti-
     depressants.'
  c. men antidepressiva tror   they at mange psykologer ville
    but antidepressants believe they that many psychologists would
     anbefale.
     recommend
     'but antidepressants they believe that many psychologists would rec-
     ommend.'
  d. men antidepressiva kjenner de til mange psykologer som ville
    but antidepressants know they to many psychologists who would
     anbefale.
     recommend
     'but antidepressants they know (of) many psychologists who would
     recommend.'

PROCEDURE. The procedure for experiment 2 was identical to that for experiment 1 but with the following modifications. For each item, the context preamble and the test sentence were presented simultaneously. The preamble was presented in italics and center-aligned above the test sentence. Test sentences were preceded by an ellipsis '…'. Although participants saw both the preamble and test sentence simultaneously, they were instructed at the beginning of the experiment to judge the acceptability of the test sentence alone.

RESULTS AND DISCUSSION. Acceptable filler sentences received high ratings on average (mean z-score = 0.72) and unacceptable filler sentences low ratings (mean z-score = −0.89). Figure 3 provides the interaction plots for each island type, and Table 2 provides a statistical summary of the island effects.
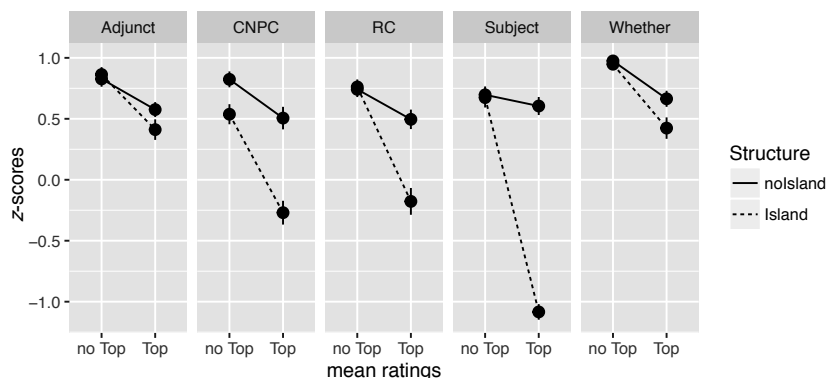
FIGURE 3. Interaction plots for all island types in experiment 2. Error bars indicate one standard error.

	t-VALUE	SIGNIFICANCE	EFFECT SIZE (DD)
ADJUNCT ISLAND	−1.40	$p = 0.168$	0.20
COMPLEX NP	−2.70	$p = 0.010$	0.49
RELATIVE CLAUSE	−4.34	$p < 0.001$	0.69
SUBJECT ISLAND	−9.90	$p < 0.001$	1.66
WHETHER-ISLAND	−1.81	$p = 0.081$	0.21

TABLE 2. Statistical summary of island effects from experiment 2.

Before discussing individual island effects, we point out the effect of the preamble. Long conditions (both island and no-island) were rated higher on average in experiment 2 than their corresponding conditions in experiment 1, except for the subject topicalization-island condition. As discussed further below, context conferred a consistent acceptability benefit on topicalization-noIsland conditions, but affected topicalization-island conditions differentially.

As in experiment 1, we observed a classic subject island effect: a significant Structure × Topicalization interaction accompanied by outright rejection of the island-violating sentences. The numerically small, marginally significant *whether*-island effect clearly fits the description of a subliminal island effect: average judgments of the island-violating condition are well above the midpoint of the scale and are roughly comparable to the average judgments of grammatical sentences across the other conditions.

We failed to observe a significant adjunct island effect (despite a slight numerical trend toward an interaction). Participants judged topicalization from conditional adjuncts to be almost as acceptable on average as topicalization from embedded declarative clauses.

CNP and RC island effects were again significant and comparable in size to those found in experiment 1, though notably smaller than CNP and RC island effects for WH-movement from Kush et al. 2018. The island effects in these two subexperiments are not easily classified as either classic or subliminal island effects. The smaller effect sizes are driven by higher average ratings of the island violations than in our earlier study: the average acceptability values of topicalization from CNP and RC islands fall close to the midpoint of the scale, whereas WH-movement from the same structures resulted in unacceptability comparable to subject island violations in Kush et al. 2018.

Once again, we analyzed the *z*-score distributions by condition to gain more information about participant judgment patterns than the mean values afford. Figure 4 plots individual rating distributions by island and condition.
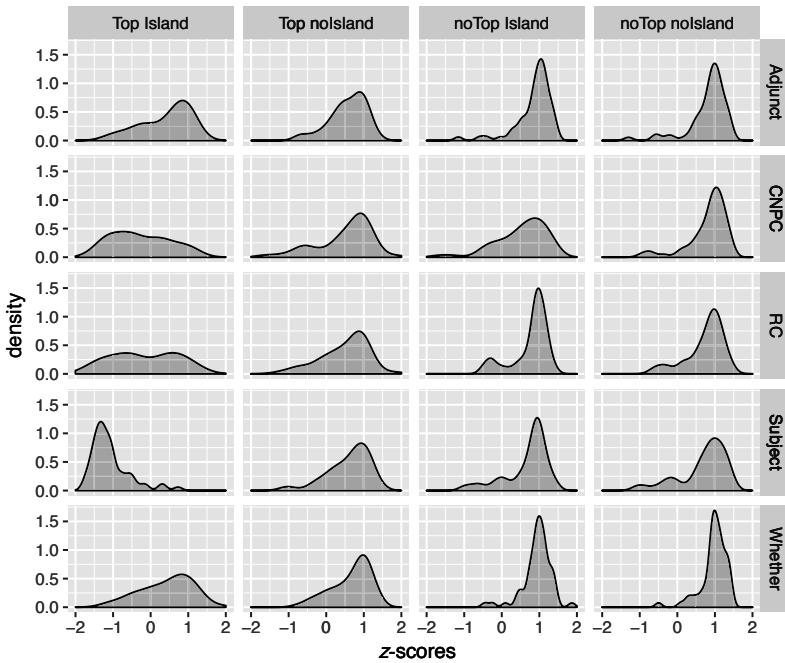
Figure 4. Ratings distributions from experiment 2 by island and condition.

As in experiment 1, judgment distributions in the no-topicalization conditions exhibit the hallmarks of consistent acceptability: unimodality centered roughly around +1. Some conditions have longer left tails, reflecting the kind of probabilistic noise common to large judgment studies. Judgments of subject island violations, once again, cluster below −1, providing an example of a condition where participants clearly reject topicalization.

The ratings of topicalization-noIsland sentences are markedly different from those in experiment 1. Z-scores across all five topicalization-noIsland conditions follow a unimodal distribution, with most judgments falling above 0. The elongated left tails of these distributions reveal that participants occasionally perceived these test sentences to be slightly marked relative to sentences without topicalization, but rarely deviant enough to reject the sentence outright. Overall, the distributions confirm that long-distance topicalization is grammatical in Norwegian, and that the lower scores for the dependencies in experiment 1 reflected the impact of nonstructural (semantic or pragmatic) factors on acceptability that were likely minimized by the addition of contextual support in experiment 2.

Judgments of topicalization from *whether*-islands are again qualitatively similar to judgments of indisputably acceptable long-distance topicalization: they have a single mode close to +1, with a longer left tail, indicating occasional marginal judgments. Participants were slightly more likely, on average, to give *whether*-island violations middling scores relative to regular long-distance topicalization, a statistical fact borne out by the numerical trend toward an interaction. However, these differences were slight, and there seemed to be few if any trials on which participants clearly rejected the movement. Thus, we can tentatively conclude that topicalization from *whether*-islands is not meaningfully different from standard long-distance topicalization from a simple declar-

ative complement clause, consistent with experiment 1. Inspection of the full distribution alone, however, permits this conclusion only at the POPULATION level. The distributions may still mask interindividual variation. For example, it is possible that the long left tail is composed of individuals who consistently rated *whether*-islands lower.

By the same logic, the distribution of judgments in the adjunct topicalization-island condition also seems to suggest that conditional adjuncts are not islands for topicalization in Norwegian. The adjunct topicalization-island distribution is nearly identical to its topicalization-noIsland counterpart.

Judgments of both CNP and RC island violations differ markedly from those of the three other islands tested. Both distributions display a high degree of variability, with many 'acceptable' (above 0) and 'unacceptable' (below 0) judgments. This is especially apparent for RC islands, where the distribution appears roughly bimodal. Bimodal rating distributions could arise because of inconsistent judgments BETWEEN participants or WITHIN participants. We explore this in our final analysis.

In order to determine whether between- or within-participant variability underlay the patterns we observed, we inspected individual participants' absolute ratings of island-violation sentences by trial (following Kush et al. 2018). For each island type we plotted individual participants' first judgment against their second (see Figure 5). Visualizing participant responses in this manner aids in identifying participant groups with specific response patterns. Those who consistently rejected island-violating sentences should cluster in quadrant 3 (bottom left), while those who judged both island violations to be acceptable should cluster in quadrant 1 (top right). Inconsistent participants—those who accepted one trial and rejected another—should occupy quadrant 2 (bottom right) or quadrant 4 (top left).
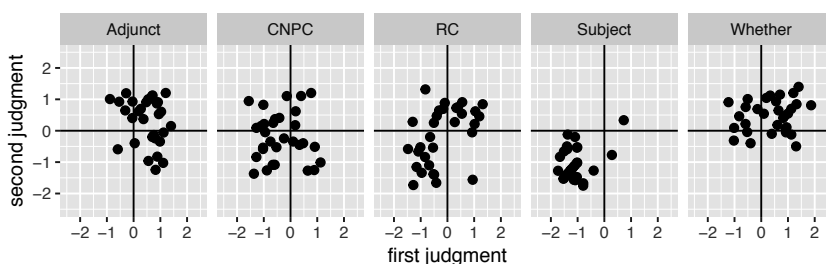

FIGURE 5. Participant ratings for each island in experiment 2. Each dot represents an individual participant, with the participant's first exposure to an island plotted against the second exposure. Participants in quadrant 1 (top right) are categorized as consistent accepters, those in quadrant 3 (bottom left) consistent rejecters, while those in quadrants 2 (bottom right) and 4 (top left) are inconsistent in their judgments.

Subject islands serve as the model of cross-participant consistency. The overwhelming majority of participants cluster in quadrant 3 because they rejected both subject island violations.

Whether-island judgments pattern differently. Almost all participants fall either into quadrant 1—that is, they were consistent accepters—or into quadrant 4, indicating that they accepted the second *whether*-island item they saw. The fact that all inconsistent participants occupied quadrant 4 suggests a degree of adaptation or recalibration: participants came to accept *whether*-island violations after rejecting their first exposure to the construction. We expect that in an experiment with a greater number of observations per participant, participants who show adaptation of this sort would continue to judge

whether-island violations to be acceptable once they had made the switch. We leave testing this prediction to future studies.[6]

Participant response patterns to adjunct islands resembled those to *whether*-islands in that there were many consistent accepters (quadrant 1) and several participants who accepted topicalization from an adjunct on the second trial. The only difference between the two island types is that some participants rejected their second exposure to adjunct island violations and thus fall into quadrant 2. It is difficult to know whether this pattern should be attributed to probabilistic noise, but we note that we observed a very similar distribution for *whether*-islands in experiments 2 and 3 of Kush et al. 2018. Nevertheless, the fact that many participants consistently accepted adjunct island violations suggests that the domains are not islands, at least for some participants.

Turning to CNPs, the sparseness of quadrant 1 shows that few participants consistently accepted topicalization out of a CNP. Slightly over one-third of participants consistently rejected topicalization out of CNPs, while the rest judged inconsistently. The near-equal number of participants in quadrants 4 and 2 indicate that presentation order did not consistently influence judgments (unlike with *whether*-islands).

Finally, three of four response patterns were observed with RC islands. Many participants consistently rejected RC island violations on both trials. However, unlike CNPs, a second group ($N = 9$) accepted both. All but one of the remaining participants fell into quadrant 4, indicating that they accepted the second RC island violation they saw. It may be that these participants display signs of adaptation, as we reasoned for response patterns in the *whether*-island subexperiment. Under this interpretation, these participants would have rejected their initial exposure to movement out of an RC, but would later 'recalibrate' their judgments. This interpretation is, of course, speculative, given the small number of observations that we collected per participant. Future research with more observations per individual should test whether such adaptation is a stable response pattern.

3. Discussion. Many researchers have argued that embedded questions, RCs, and CNPs are never syntactic islands in Mainland Scandinavian languages (Allwood 1982, Engdahl 1982, Maling & Zaenen 1982, 1997, Lindahl 2017, among others). However, the few formal acceptability judgment studies that have systematically tested movement from these domains have largely failed to support these claims. Although some Scandinavian participants appear to accept wh-movement from *whether*-islands, the same participants consistently reject wh-movement from RCs and CNPs (Kush et al. 2018).

The fact that wh-movement from RCs and CNPs has yielded strong island effects in past experiments does not necessarily entail that the domains are uniform syntactic islands. Even if movement from these domains is in principle possible in Mainland Scandinavian, participants nevertheless could have judged the dependencies to be unacceptable. We outlined two possibilities for how this could occur. First, if the constructions in question are string-ambiguous between island and nonisland parses, past participants may have incorrectly opted for the island parse when judging the sentences,

---

[6] We use the generic term *adaptation* here, rather than a more specific term like *satiation*, because we are not sure what relationship our effects have to satiation. Satiation is typically an increase in acceptability for a sentence that is considered uniformly ungrammatical. Here, participants' judgments appear to undergo a dramatic directional shift, indicating two possible parses—one of which is grammatical. See Do & Kaiser 2017 for both a review of the syntactic satiation literature and an investigation of the relationship between satiation and another potentially related phenomenon: syntactic priming.

resulting in an island effect. Alternatively, unacceptability might have arisen purely on extrasyntactic (i.e. semantic, discourse-pragmatic, or processing) grounds.

We tested whether we could find experimental support for the claim that embedded questions, RCs, and CNPs are not syntactic islands for A′-movement in Norwegian. Topicalization was our test dependency, unlike most previous studies. Experiment 1 tested sentences containing decontextualized topicalization from five different island domains: subject, *whether*-, RC, CNP, and conditional adjunct islands. Experiment 2 tested the same sentences, but provided contextual support for the dependencies. If, on the one hand, the domains in question are indeed syntactic islands, judgment patterns should be impervious to both the change in A′-dependency and our contextual manipulation: participants should uniformly reject topicalization from RCs and CNPs, just as they have rejected wh-movement in prior studies. On the other hand, we reasoned that if past participants rejected movement from RCs and CNPs because they chose the wrong structural analysis, or for extrasyntactic reasons, then judgment patterns might change with topicalization (under the assumption that the transformation is subject to different semantic and discourse-pragmatic conditions) or with contextual support.

Our experiments used the factorial design for testing island effects (Sprouse 2007, Sprouse et al. 2012, among others). In cases where we found a significant Structure × Distance interaction in the average analysis, we concluded that there was an island effect. However, the mere presence of an interaction effect was insufficient for determining whether the source of the observed average unacceptability was a syntactic constraint violation or something else.

We also inspected judgment distributions for more information about the origin of the interaction effects. Bimodal distributions where participant judgments were split between acceptance and rejection were taken to be incompatible with the uniform syntactic island view, under the relatively commonplace assumption that syntactic constraints are consistent and CATEGORICAL. We expected robust subject and adjunct island effects and relative uniformity among participants in their rejecting such dependencies. Our primary goal was to ascertain whether the three remaining domains would exhibit different response patterns.

We review the basic empirical findings for each island type below, and then briefly discuss their theoretical implications. Our intent is not to argue for a specific account of island constraints or the ability to violate them; our results are compatible with a number of different formal analyses that our experiments were not designed to adjudicate between. Instead, we discuss how our results inform theorizing about islands and crosslinguistic variation more broadly and, as always, the future research questions raised by these results.

3.1. EMPIRICAL SUMMARY.

SUBJECT ISLANDS. Norwegian participants rejected topicalization out of complex subjects in both experiments. Rejection was unequivocal in experiment 1. No participants accepted subject island violations on even a portion of trials. In experiment 2, judgments of subject island violations proved resistant to the contextual manipulation. These results confirm that complex subjects are islands for DP topicalization in Norwegian. More importantly, the persistent unacceptability of topicalization from subject islands provides a clear baseline against which other islands can be evaluated. The clear subject island effects militate against the possibility that topicalization is always 'island-insensitive' or that it is a 'nonmovement' dependency. Thus, we can be more confident that insensitivity to other islands is unlikely due to a nonmovement analysis.

Whether-islands. Judgment patterns from experiments 1 and 2 largely support previous claims that embedded (polar) questions are not syntactic islands for topicalization in Norwegian.[7] In experiment 1, we found no superadditive island effect. Judgment distributions corroborated the conclusion that participants did not perceive an acceptability difference between topicalizing from an embedded question and from an embedded declarative clause. In experiment 2 we found a marginally significant island effect. The effect is best categorized as a subliminal island effect since topicalization from *whether*-islands was judged to be as acceptable, on average, as grammatical items. Distributional analysis revealed that the majority of participants consistently accepted topicalization from *whether*-islands. Embedded questions are clearly not syntactic islands for these participants. The small interaction effect was driven by a group of participants who rated topicalization from *whether*-islands inconsistently. Importantly, these participants' judgments conformed to a clear pattern: the participants rated the first example of a *whether*-island violation they saw as degraded, but ultimately accepted later examples. We suggest, following Kush et al. 2018, that inconsistent judgments are compatible with a string-ambiguity account or an extrasyntactic explanation of the effect.

RC islands. In both experiments 1 and 2, average island effects were significant and considerably larger than interaction effects for *whether*-islands. Analysis of judgment distributions showed that almost no participants accepted decontextualized topicalization from RCs in experiment 1. Judgment patterns shifted when topicalization was presented with supporting contexts in experiment 2. *Z*-scores in the topicalization-island condition were bimodally distributed, with roughly half of all examples being accepted and the other half rejected. Further analysis showed that this bimodal distribution reflected a mixture of judgments from three groups of participants with distinct response patterns. One group consistently accepted topicalization from RCs, another rejected their first exposure to the constructions but accepted their second, and the final group rejected all dependencies spanning an RC. The first two groups display a pattern reminiscent of those in the *whether*-island subexperiment. Following the same logic as above, we conclude that RCs are not uniform syntactic islands, at least for consistent accepters.

There are two options for interpreting the behavior of the third group, the consistent rejecters. First, it is possible that movement from RCs is, in principle, possible for these participants, but that they nevertheless rejected the sentences because they selected the wrong syntactic analysis, or perceived the resulting dependencies to be semantically or pragmatically ill-formed. Under this hypothesis consistent rejecters are essentially more conservative versions of inconsistent participants. The alternative is that consistent rejecters are grammatically divergent from their peers and never allow syntactic movement from RCs. Teasing apart these two possibilities will require systematically manipulating the factors that could influence structural choice or semantic/pragmatic felicity. The prediction is that it should be possible to turn these participants into consistent accepters with the correct manipulation. Conversely, corroborating the hypothesis that RCs are syntactic islands for these participants would require identifying a group-level property that correlates with these judgments, such as geographic region, specific input during childhood, or even other syntactic properties of these participants' grammars.

Complex NP islands. Judgments of topicalization from CNPs showed little variability in experiment 1, but were slightly more variable in experiment 2. A small num-

[7] We expect that comparable results would hold in the other Mainland Scandinavian languages, Swedish and Danish, given the high degree of syntactic similarity across these languages.

ber of participants were consistent accepters, yet most were either consistent rejecters or inconsistent raters. These findings align with the literature: although it has been claimed that CNPs are not islands in Mainland Scandinavian (Allwood 1982), such claims are rarely paired with supporting examples from naturalistic discourse. This contrasts with RC islands and embedded questions, naturalistic examples of which appear to be (relatively) prevalent.

Taken at face value, the small number of consistent accepters might suggest that CNPs are not islands for some participants. However, the preponderance of rejections indicates that if such an analysis of CNPs exists, most Norwegians lack it. Alternatively, it could suggest that CNP islands are more resistant to contextual amelioration than RCs or *whether*-islands, but we consider such an analysis unlikely: what semantic or discourse-pragmatic conditions could be tailored specifically to CNPs, which are essentially synonymous with verbal complement CPs that generally allow extraction?

Although there is uncertainty about how best to interpret these results, we prefer to stake the stronger and more conservative claim that CNPs are syntactic islands in Norwegian until evidence is provided to the contrary. We believe this is the claim that is most consistent with the evidence, but welcome the possibility of new experimental evidence against this conclusion (e.g. an experimental manipulation of an extrasyntactic property that substantially lowers the proportion of rejecters).

CONDITIONAL ADJUNCT ISLANDS. Judgment of adjunct island violations did not pattern as expected. There was a significant interaction effect in experiment 1, but judgments of adjunct island violations were distributed bimodally: nearly half of all violations were judged to be acceptable. In experiment 2 we failed to observe a significant interaction effect. Participants accepted more than half of the adjunct island violations. All participants but one were either consistent accepters or inconsistent raters.

These results suggest that conditional adjuncts are not islands for A′-movement in Norwegian. We doubt that the results can be attributed to a confound: we specifically guarded, for example, against the possibility of parasitic-gap dependencies that could license a gap inside the adjunct, and, as discussed above, the subject island findings argue against the possibility that topicalization itself is inherently island-insensitive.

To our knowledge, these experiments are the first to demonstrate that conditional adjuncts are not islands for topicalization in Norwegian. This raises questions regarding how widespread insensitivity to adjunct islands is in Mainland Scandinavian languages. It has occasionally been reported that movement from some adjuncts can seem acceptable in Mainland Scandinavian (Anward 1982, Faarlund 1992, Hansen & Heltoft 2011), but little work has methodically investigated this possibility (though see Müller 2017 for preliminary exploration of movement from purpose and certain temporal adjuncts in Swedish). The next step is to explore specific theories of the extrasyntactic properties that give rise to the apparent adjunct island effects in Kush et al. 2018 and experiment 1 here.

3.2. IMPLICATIONS.

EMBEDDED QUESTIONS, RCs, AND COMPLEX NPs. Our results suggest that embedded questions and RCs are not uniform syntactic islands for topicalization (at least for a number of speakers), but CNPs are. If we assume that the underlying structural analyses of Norwegian embedded questions and RCs are identical in all relevant respects to their English counterpart structures, then our results are unexpected under theories such as SUBJACENCY (Chomsky 1977) and the BARRIERS framework (Chomsky 1986), as well as many modern phase-based approaches.

Past proposals have tried to square acceptable movement from embedded questions with cycle-based theories either by parameterizing bounding nodes or phase heads (e.g.

Rizzi 1982) or by positing the existence of additional 'escape hatches' in the left periphery of embedded clauses to enable successive cyclic movement even if another element occupies a left-peripheral specifier (e.g. Reinhart 1981). To our knowledge, no authors have tried the first strategy in Mainland Scandinavian, but a number of authors have endorsed some version of the latter (e.g. Lindahl 2017, Nyvad et al. 2017, Vikner 2017, Vikner, Christensen, & Nyvad 2017).[8]

Such analyses can be extended to other islands insofar as the underlying structure of the other island domains can be likened to embedded questions. For example, Sichel (2018) proposes that acceptable extraction from RCs in Hebrew is allowed with RAISING RCs, where the RC is not a complement/adjunct to N (see Bianchi 1999). The structure of a raising RC is exemplified in 10.

(10) I know [$_{DP}$ many [$_{CP}$ psychologists$_i$ [$_{C'}$ that [$_{TP}$ t$_i$ would recommend them

The raising RC structure resembles an embedded question in that there is no NP layer above CP, as there would be under a matching analysis (Sauerland 1998, Bianchi 1999), illustrated in 11.

(11) I know [$_{DP}$ many [$_{NP}$ psychologists$_i$ [$_{CP}$ ~~psychologists$_i$~~ [$_{C'}$ that [$_{TP}$ t$_i$ would recommend them

Sichel argues that extraction from a matching RC such as the one in 11 is blocked by the NP layer above CP.

The fact that N layers block extraction in such structural configurations predicts that CNPs should be islands, since a raising analysis is not available for the structures. Sichel shows that this prediction is borne out in Hebrew. Adapting a Sichelian analysis of Hebrew RC island extraction to Norwegian (and other Mainland Scandinavian languages more generally) is tempting. Such an account conforms with our observations that CNPs appear to be islands, but movement from embedded questions and RCs is at times acceptable.

Suppose that such a locality-compatible alternate structure is indeed available to any participant who accepts even a single example (consistent and inconsistent accepters). If there is such an analysis, why would participants ever reject movement from embedded questions or RCs? That is, why do we ever see island effects, subliminal or classic, in our experiments? We illustrate avenues of explanation using embedded questions.

A pure structural-ambiguity account holds that the parser faces a choice when reading a sentence where an A′-dependency spans an embedded question. It can either analyze the embedded question as an ordinary WH-island—that is, an embedded clause with a filled left periphery—or apply the alternate structure that allows movement. Taking the first option results in unacceptability because the resulting dependency would violate locality conditions on movement. Taking the second option results in acceptability (modulo any additional semantic well-formedness considerations). Thus, for any island effect the size of the interaction effect is proportionate to the probabilities with which the different analyses are chosen. Individual differences between participants in

---

[8] Variants of this type of account can differ, so we outline an abstract version below. Such accounts stipulate an additional head, call it Z, atop CP, whose specifier A′-movement can use as an escape hatch out of the island.

(i) [$_{ZP}$ [$_{Z'}$ Z [$_{CP}$ Op [$_{C'}$ C [$_{TP}$ …]]]]]

If Z, and not C, is the phase head (or bounding node) in these embedded clauses, then the possibility of extraction from embedded questions in Mainland Scandinavian can be reconciled with locality-based theories. Crosslinguistic differences between Mainland Scandinavian languages and languages such as English are explained by the stipulation that English does not have ZP.

acceptance rate could then follow from different participant-specific biases toward one analysis over the other. Extending the structural-ambiguity analysis to RC island effects would require relativizing the probability of adopting an alternative parse by construction type. Differences in rate of acceptance between wh- and RC island sentences entail that participants are more likely to adopt the alternate parse for embedded questions than for RCs, perhaps reflecting differences in the rates at which the different structures are encountered in the input. Further, probabilities would also have to be sensitive to differences in A′-movement type and contextual factors, as participants are more likely to accept contextually motivated topicalization from RCs than wh-movement.

It is also possible that participants freely generate the alternate analysis, allowing free A′-movement in the syntax. If that is the case, then island effects would have to be excluded on extrasyntactic grounds, such as violations of semantic constraints or pragmatic felicity conditions. There are, at present, very few clear formal proposals for what such constraints could be. Kush et al. 2018 considered semantic explanations for certain weak island effects (e.g. Szabolsci & Zwarts 1993) as an explanation for inconsistent judgments of *whether*-island structures, but it is not clear how those constraints generalize beyond wh-movement or embedded questions. Some proposals have been offered that link the ability to move out of an RC to concepts like BACKGROUNDEDNESS, PRESUPPOSITION, or other information-structural concepts (e.g. Erteschik-Shir 1973, 1982, Engdahl 1982, Ambridge & Goldberg 2008, Löwenadler 2015), but have not offered an explanation for how to best formalize those constraints. Discussing such proposals in the abstract, the explanation for variability is essentially the same. Different constructions and dependencies should exhibit different patterns of extractability because they are subject to different well-formedness conditions. Interparticipant variation should be explained by differences in how easy individuals find it to adopt or coerce a constraint-compatible interpretation of the dependencies in question.

We have considered explanations of the behavior of those who accept these extractions. What about those who consistently rejected RC island violations? In principle, such interindividual differences could reflect between-participant differences at the level of grammar. Some participants could lack the alternate structure in their mental grammars. Such a state of affairs might arise due to dialect variation[9] or through accidents of idiosyncratic exposure (some speakers might never receive the relevant input that alerts them to the need to posit such a structure). While such an outcome is of course possible, we favor the conclusion that there are not such differences. One weak argument against this kind of grammatical variation comes from the theorized relationship between wh-islands and RC islands. If the same structure that allows one to move out of an embedded question is used for movement from RCs, then all of our participants have the relevant analysis, since none of our participants consistently rejected movement out of embedded questions.

Finally, before moving on to considering the implications of our adjunct island results, we make one observation. We have cast our discussion above in terms of Sichel's (2018) raising analysis of escapable RCs for the sake of concreteness. Nevertheless, none of our experiments were designed to specifically test the predictions of this analysis. Future work should specifically determine whether RC island effects in fact are toggled on or off according to whether a raising/matching analysis is available.

---

[9] Insofar as we can tell, the interparticipant variation does not seem to be reducible to known dialectal varieties of Norwegian.

ADJUNCT ISLANDS. We have shown that conditional adjuncts are not islands for topicalization, and arguably not for A′-movement generally, but we do not know at present whether our results generalize to all adjuncts. Müller (2017) provides some initial evidence that other adjuncts, specifically purpose and some temporal adjuncts, may also be transparent for topicalization in Mainland Scandinavian, but also shows that extraction is not completely free. It is conceivable that transparent adjuncts have a peculiar structural property that allows for A′-extraction, which is not shared with those that block extraction. Yet it is unclear what that structural property could be. Extraction from adjuncts, conditional and otherwise, is traditionally ruled out by Huang's (1982) CONDITION ON EXTRACTION DOMAIN (CED).

(12) CONDITION ON EXTRACTION DOMAIN (paraphrased): A phrase A may be extracted out of a domain B only if B is a COMPLEMENT.[10]

Manipulations of the internal syntax of transparent adjuncts would not have any ameliorative effects on CED violations.[11] So far as we can tell, the only way to circumvent CED effects would be to posit that 'adjunct' clauses that allow extraction are, in fact, not adjuncts at all, but complements. We take this to be an unlikely solution.

If conditional adjuncts (and some others) are not syntactic islands in Norwegian—and no alternative structural analysis for this subset is tenable—then it would stand to reason that no adjuncts are syntactic islands in Norwegian. This has interesting typological consequences: the traditional CED approach to adjunct islands assumes that extraction is ruled out from adjuncts and subjects by the same syntactic principles. Yet insofar as we can conclude that our subject island effects are syntactic in origin, our results suggest that subject and adjunct island effects arise for different reasons in Mainland Scandinavian. This, in turn, casts doubt on the tenability of a unified analysis of subject islands and adjunct islands across languages. Stepanov (2007) argued that subject island effects were dissociable from adjunct island effects based on the observation that many languages allow extraction out of subjects. However, Stepanov showed only a one-way dissociation, which is less-strong evidence for the claim that subject and adjunct island effects are independent. Although we admittedly did not run a comprehensive study of extraction from different subject and adjunct types, our results provide preliminary evidence of a true double dissociation, thereby strengthening the argument for independent explanations for the phenomena.

Further, if adjuncts are not SYNTACTIC islands in Scandinavian, support for a universal, syntactic prohibition on movement from adjuncts is somewhat weakened. This does not mean that adjunct island EFFECTS are illusory. Such a stance is patently false. Instead, it means that we may have to consider alternative explanations for the diverse range of adjunct island effects within and across languages. For example, it is possible that these effects stem from (potentially universal) constraints on semantic well-formedness. Such semantic constraints could presumably allow 'gaps' in coverage, whereby certain movement dependencies would be allowed when various semantic principles accidentally conspired to satisfy semantic conditions. The idea that movement from adjuncts is se-

[10] Huang's CED was originally stated in terms of *proper government*: 'A phrase A may be extracted out of a domain B only if B is a properly governed'.

[11] Vikner (2017) seems to imply that his recursive-*c*P analysis provides an explanation for extraction from adjuncts, but the exact mechanism for this is unclear to us. Others have suggested that Haegeman's (2012) distinction between central and peripheral adjuncts may provide an appropriate way of at least categorizing those adjuncts that allow movement and those that do not. We suspect that the central-peripheral distinction will not work for explaining the porousness of conditional adjuncts.

mantically, but not syntactically, constrained has been tentatively advanced in the past. For example, Truswell (2007, 2011) proposes that extraction from adjuncts is allowed so long as a certain relation holds between the events associated with the verbs in the adjunct and its container VP. In order to explain our results under Truswell's account, we would have to suppose that our conditional adjunct items met these conditions. Because we did not specifically control for the interpretive relations that held between our matrix and adjunct clauses, we cannot say for certain whether this is the case. Future work should test this more systematically. However, we are somewhat skeptical that an explanation of this type can account for the entirety of the effects. Most importantly, Truswell's account is presented as a universal condition, which does not provide an explanation for why extraction appears to be freer in Norwegian than in other languages. Second, the account does not provide an explanation of why we should observe differences between WH-movement and topicalization. Whether an individual proposal is the correct analysis for our cases is less our concern; we think the general idea of semantic constraints merits serious consideration.

ACQUISITION. We concluded that (at least some) embedded questions and RCs are not syntactic islands in Norwegian. Nor are conditional adjuncts. The outstanding puzzle is now to determine how the Scandinavian child comes to learn that these domains are nonislands, while children elsewhere come to a different conclusion.

If island effects reflect innate constraints on syntactic representations or operations (Ross 1967, Chomsky 1977, 1986), then, as discussed above, acceptable movement from an apparent island entails either constraint parameterization or an alternative structural analysis of the domain that allows a constraint violation to be circumvented. We considered the second option as an explanation for the acceptability of topicalization from embedded questions and RCs.

How could a Norwegian child learn of the existence of such an alternative analysis? Broadly speaking, the child could receive DIRECT EVIDENCE OF APPARENT ISLAND VIOLATIONS in their input. Observing A′-dependencies spanning embedded questions or RCs would alert the child to the need to posit a new analysis, but the sentences alone would not guarantee the form of that analysis. Absent additional assumptions, the learner would not be forced to hypothesize, for example, that extra escape hatches were allowed in Norwegian. The child could receive INDIRECT EVIDENCE for an alternative structure by observing its COMPONENT PARTS in other constructions. For example, the child might observe sentences whose analysis requires the generation of an extended left periphery,[12] or evidence that a raising analysis of RCs is available in their language if the Sichelian approach can be extended to Norwegian. These components could then be combined and pressed into service to explain apparent island violations 'for free'.[13]

At present it is difficult to determine whether children receive direct evidence of island violations or evidence of the syntactic pieces of the alternate analysis. This is for two reasons. First, there are currently no large corpora of Norwegian child-directed speech, so we cannot estimate the frequency of any type of evidence, direct or indirect.

---

[12] Vikner et al. (2017) suggest that direct evidence for an extended left periphery may come from the possibility of 'stacked complementizers' in certain embedded clauses. For example, complementizers such as *fordi* 'because' and *hvis/om* 'if' may be followed by the declarative C^0 *at* 'that' in adjunct clauses. We further speculate that the possibility of embedded V2 word order (Holmberg & Platzack 1995, Julien 2007, among others) may alert learners to the possibility of an extended left periphery.

[13] Another alternative, which we do not pursue here, may be to investigate whether there are clusters of RC CONSTRUCTIONS that allow extraction, as in Löwenadler 2015.

Second, we do not have concrete proposals for which structures could constitute evidence for the various parts of the analysis.

It does seem as though apparent island violations occur naturally in Mainland Scandinavian (for Swedish, see Engdahl 1997 and Lindahl 2017), but it is unclear whether these attested cases are restricted to certain registers or what their base frequency is. Thus, it is unclear whether the violations appear consistently and reliably enough to ensure successful acquisition (see e.g. Ambridge et al. 2015, and Legate & Yang 2002 for discussion of the necessary frequencies required to learn a construction). Based on our own anecdotal experience, we speculate that there may be sufficient direct evidence of apparent wh-island violations in everyday speech to conclude that embedded questions are not islands,[14] though we underscore that these are only speculative remarks that remain to be evaluated rigorously and empirically.

Learning the nonisland status of conditional adjuncts in Norwegian is potentially more challenging. If conditional adjuncts are to be singled out as an exceptional class (if, for example, they require a special syntactic analysis), the input should provide either reliable direct evidence of A′-dependencies that span conditional adjuncts, or indirect evidence that the underlying analysis of conditional adjuncts is different from that of other adjuncts. As with the other islands, it is unclear whether such input exists. But if we assume that adjuncts are NOT syntactic islands, then we must account for the semantic or possibly pragmatic conditions that enable extraction. If such conditions must be learned (as opposed to falling out from the interaction of more basic principles), then we are faced with the same challenge as above: we must figure out what properties of the input actually condition such learning.

The discussion above assumes that children come to learn a language pre-equipped with innate constraints on syntactic movement. Recent work has explored the idea that specific island patterns might be learnable without having to posit specific hard-wired constraints on movement per se. Generalizations about acceptable A′-dependencies can be induced by learners equipped with weaker (though nevertheless nontrivial) and more-general innate learning biases. How might we explain the crosslinguistic differences under such models?

Pearl and Sprouse (2013a,b) propose a computational learner capable of acquiring English island judgments through exposure to a phrase-structure-annotated subset of child-directed speech from the CHILDES corpus (MacWhinney 2000). The model is trained by extracting A′-dependencies from the corpus, which it represents as sequences of phrase structure nodes. The model then tabulates the frequency of all three-node sequences (trigrams) in the input. The acceptability of a new dependency according to the model relates proportionally to the probability of the sequence of its container nodes, calculated by computing the product of its constituent trigram probabilities. Pearl and Sprouse show that the model assigns probabilities to island violations that are orders of magnitude lower than nonisland dependencies of comparable length, thereby mimicking the pattern of acceptability of such dependencies in off-line judgment experiments.

Pearl and Sprouse's model rejects island violations in English, because such dependencies are not observed in the input. In order to mimic the pattern of effects that we observed in Norwegian, such a model would need to be exposed to input that contained

---

[14] If a common explanation for the nonisland status of (certain) RCs in Norwegian follows from the analysis of embedded questions, then we could consider there to be indirect negative evidence to learn the appropriate distribution of acceptable movement from RCs.

enough direct evidence of apparent island violations. In future work, we hope to evaluate whether such a model could succeed in learning the Norwegian pattern of judgments.

4. CONCLUSION. Our experiments investigated topicalization out of five different environments: conditional adjunct islands, complex noun phrase islands, relative clause islands, subject islands, and *whether*-islands, with and without contextual support. Judgments of topicalization from embedded questions and relative clauses exhibited a degree of interindividual and intertrial variation incompatible with the conclusion that the two domains are uniform syntactic islands. Contrary to some claims, we found little evidence that complex NPs were not islands. Unexpectedly, the experiments also found evidence that conditional adjunct clauses may not be islands for movement in Norwegian.

REFERENCES

ÅFARLI, TOR A., and KRISTIN MELUM EIDE. 2003. *Norsk generativ syntaks*. Oslo: Novus.
ALLWOOD, JENS. 1982. The complex NP constraint in Swedish. In Engdahl & Ejerhed, 15–32.
ALMEIDA, DIOGO. 2014. Subliminal wh-islands in Brazilian Portuguese and the consequences for syntactic theory. *Revista da ABRALIN* 13.55–93. DOI: 10.5380/rabl.v13i2.39611.
AMBRIDGE, BEN, and ADELE GOLDBERG. 2008. The island status of clausal complements: Evidence in favor of an information structure explanation. *Cognitive Linguistics* 19.357–89. DOI: 10.1515/COGL.2008.014.
AMBRIDGE, BEN; EVAN KIDD; CAROLINE F. ROWLAND; and ANNA L. THEAKSTON. 2015. The ubiquity of frequency effects in first language acquisition. *Journal of Child Language* 42.239–73. DOI: 10.1017/S030500091400049X.
ANDERSSON, LARS-GUNNAR. 1982. What is Swedish an exception to? Extractions and island constraints. In Engdahl & Ejerhed, 33–45.
ANWARD, JAN. 1982. Basic Swedish. In Engdahl & Ejerhed, 47–75.
BATES, DOUGLAS; MARTIN MÄCHLER; BEN BOLKER; and STEVE WALKER. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1).1–48. DOI: 10.18637/jss.v067.i01.
BIANCHI, VALENTINA. 1999. *Consequences of antisymmetry: Headed relative clauses*. Berlin: Walter de Gruyter.
BRESNAN, JOAN, and JENNIFER HAY. 2008. Gradient grammar: An effect of animacy on the syntax of *give* in New Zealand and American English. *Lingua* 118.245–59. DOI: 10.1016/j.lingua.2007.02.007.
CHOMSKY, NOAM. 1977. On *wh*-movement. *Formal syntax*, ed. by Peter Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.
CHOMSKY, NOAM. 1986. *Barriers*. Cambridge, MA: MIT Press.
CHRISTENSEN, KEN RAMSHØJ; JOHANNES KIZACH; and ANNE METTE NYVAD. 2013. Escape from the island: Grammaticality and (reduced) acceptability of *wh*-island violations in Danish. *Journal of Psycholinguistic Research* 42.51–70. DOI: 10.1007/s10936-012-9210-x.
CHRISTENSEN, KEN RAMSHØJ, and ANNE METTE NYVAD. 2014. On the nature of escapable relative islands. *Nordic Journal of Linguistics* 37.29–45. DOI: 10.1017/S0332586514000055.
CHRISTENSEN, KIRSTI KOCH. 1982. On multiple filler-gap constructions in Norwegian. In Engdahl & Ejerhed, 77–98.
DO, MONICA L., and ELSI KAISER. 2017. The relationship between syntactic satiation and syntactic priming: A first look. *Frontiers in Psychology* 8:1851. DOI: 10.3389/fpsyg.2017.01851.
DRUMMOND, ALEX. 2012. IbexFarm (version 0.3.7). [Software] Online: http://spellout.net/ibexfarm.
ENGDAHL, ELISABET. 1982. Restrictions on unbounded dependencies in Swedish. In Engdahl & Ejerhed, 151–74.
ENGDAHL, ELISABET. 1997. Relative clause extractions in context. *Working Papers in Scandinavian Syntax* 60.51–79.

Engdahl, Elisabet, and Eva Ejerhed (eds.) 1982. *Readings on unbounded dependencies in Scandinavian languages*. Stockholm: Almquist & Wiksell.

Erteschik-Shir, Nomi. 1973. *On the nature of island constraints*. Cambridge, MA: MIT dissertation.

Erteschik-Shir, Nomi. 1982. Extractability in Danish. In Engdahl & Ejerhed, 175–91.

Faarlund, Jan Terje. 1992. *Norsk syntaks i funksjonelt perspektiv*. Oslo: Universitetsforlaget.

Featherston, Sam. 2007. Data in generative grammar: The stick and the carrot. *Theoretical Linguistics* 33.269–318. DOI: 10.1515/TL.2007.020.

Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and composition of the left periphery: The cartography of syntactic structures, vol. 8*. Oxford: Oxford University Press.

Hansen, Erik, and Lars Heltoft. 2011. *Grammatik over det Danske Sprog*. Roskilde: University Press of Southern Denmark.

Holmberg, Anders, and Christer Platzack. 1995. *The role of inflection in Scandinavian syntax*. Oxford: Oxford University Press.

Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Cambridge, MA: MIT dissertation.

Julien, Marit. 2007. Embedded V2 in Norwegian and Swedish. *Working Papers in Scandinavian Syntax* 80.103–61.

Keller, Frank. 2000. *Gradience in grammar: Experimental and computational aspects of degrees of grammaticality*. Edinburgh: University of Edinburgh dissertation.

Keshev, Mayaan, and Aya Meltzer-Asscher. 2019. A processing-based account of subliminal wh-island effects. *Natural Language and Linguistic Theory* 37.621–57. DOI: 10.1007/s11049-018-9416-1.

Kush, Dave; Terje Lohndal; and Jon Sprouse. 2018. Investigating variation in island effects: A case study of Norwegian wh-extraction. *Natural Language and Linguistic Theory* 36.743–79. DOI: 10.1007/s11049-017-9390-z.

Kuznetsova, Alexandra; Per B. Brockhoff; and Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82.1–26. DOI: 10.18637/jss.v082.i13.

Legate, Julie Anne, and Charles D. Yang. 2002. Empirical re-assessment of stimulus poverty arguments. *The Linguistic Review* 19.151–62. DOI: 10.1515/tlir.19.1-2.151.

Lindahl, Filippa. 2014. Relative clauses are not always strong islands. *Working Papers in Scandinavian Syntax* 93.1–25.

Lindahl, Filippa. 2017. *Extraction from relative clauses in Swedish*. Gothenburg: University of Gothenburg dissertation.

Löwenadler, John. 2015. Relative clause extraction: Pragmatic dominance, processing complexity and the nature of crosslinguistic variation. *Nordic Journal of Linguistics* 38.37–65. DOI: 10.1017/S0332586515000050.

MacWhinney, Brian. 2000. *The CHILDES project: Tools for analyzing talk*. 3rd edn. Mahwah, NJ: Lawrence Erlbaum.

Maling, Joan, and Annie Zaenen. 1982. A phrase-structure account of Scandinavian extraction phenomena. *The nature of syntactic representation*, ed. by Pauline Jacobson and Geoffrey K. Pullum, 229–82. Dordrecht: Reidel.

Maxwell, Scott E., and Harold D. Delaney. 2003. *Designing experiments and analyzing data: A model comparison perspective*. London: Routledge.

Müller, Christiane. 2017. Extraction from adjunct islands in Swedish. *Norsk Lingvistisk Tidsskrift* 35.67–85. Online: http://ojs.novus.no/index.php/NLT/article/view/1399/1384.

Nyvad, Anne Mette; Ken Ramshøj Christensen; and Sten Vikner. 2017. CP-recursion in Danish: A cP/CP-analysis. *The Linguistic Review* 34.449–77. DOI: 10.1515/tlr-2017-0008.

Pearl, Lisa, and Jon Sprouse. 2013a. Syntactic islands and learning biases: Combining experimental syntax and computational modeling to investigate the language acquisition problem. *Language Acquisition* 20.23–68. DOI: 10.1080/10489223.2012.738742.

Pearl, Lisa, and Jon Sprouse. 2013b. Computational models of acquisition for islands. *Experimental syntax and island effects*, ed. by Jon Sprouse and Norbert Hornstein, 109–31. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139035309.006.

R Core Team. 2018. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Online: https://www.r-project.org/.

Reinhart, Tanya. 1981. A second COMP position. *Theory of markedness in generative grammar: Proceedings of the 1979 GLOW conference*, ed. by Adriana Belletti, 518–57. Pisa: Scuola Normale Superiore.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT dissertation. [Published as *Infinite syntax!*, Norwood, NJ: Ablex, 1986.]

Sauerland, Uli. 1998. *The meaning of chains*. Cambridge, MA: MIT dissertation.

Sichel, Ivy. 2018. Anatomy of a counterexample: Extraction from relative clauses. *Linguistic Inquiry* 49.335–78. DOI: 10.1162/LING_a_00275.

Sprouse, Jon. 2007. *A program for experimental syntax: Finding the relationship between acceptability and grammatical knowledge*. College Park: University of Maryland dissertation.

Sprouse, Jon. 2008. The differential sensitivity of acceptability to processing effects. *Linguistic Inquiry* 39.686–94. DOI: 10.1162/ling.2008.39.4.686.

Sprouse, Jon; Ivano Caponigro; Ciro Greco; and Carlo Cecchetto. 2016. Experimental syntax and the variation of island effects in English and Italian. *Natural Language and Linguistic Theory* 34.307–44. DOI: 10.1007/s11049-015-9286-8.

Sprouse, Jon; Shin Fukuda; Hajime Ono; and Robert Kluender. 2011. Reverse island effects and the backward search for a licensor in multiple *wh*-questions. *Syntax* 14.179–203. DOI: 10.1111/j.1467-9612.2011.00153.x.

Sprouse, Jon; Matt Wagers; and Colin Phillips. 2012. A test of the relation between working-memory capacity and syntactic island effects. *Language* 88.82–123. DOI: 10.1353/lan.2012.0004.

Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10.80–126. DOI: 10.1111/j.1467-9612.2007.00094.x.

Szabolcsi, Anna, and Frans Zwarts. 1993. Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics* 1.235–84. DOI: 10.1007/BF00263545.

Taraldsen, Knut Tarald. 1982. Extraction from relative clauses in Norwegian. In Engdahl & Ejerhed, 205–21.

Truswell, Robert. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117.1355–77. DOI: 10.1016/j.lingua.2006.06.003.

Truswell, Robert. 2011. *Events, phrases, and questions*. Oxford: Oxford University Press.

Tucker, Matthew A.; Ali Idrissi; Jon Sprouse; and Diogo Almeida. 2019. Resumption ameliorates different islands differentially: Acceptability data from Modern Standard Arabic. *Perspectives on Arabic Linguistics XXX: Papers from the annual symposia on Arabic Linguistics, Stony Brook, New York, 2016 and Norman, Oklahoma, 2017*, ed. by Amel Khalfaoui and Matthew A. Tucker, 159–93. Amsterdam: John Benjamins.

Vikner, Sten. 2017. CP-recursion and the derivation of verb second in Germanic main and embedded clauses. *Two perspectives on V2: The invited talks of the DGfS 2016 Workshop 'V2 in grammar and processing: Its cause and its consequences'*, ed. by Constantin Freitag, Oliver Bott, and Fabian Schlotterbeck, 1–26. Konstanz: Universität Konstanz.

Vikner, Sten; Ken Ramshøj Christensen; and Anne Mette Nyvad. 2017. V2 and cP/CP. *Order and structure in syntax I: Word order and syntactic structure*, ed. by Laura R. Bailey and Michelle Sheehan, 313–24. Berlin: Language Science. DOI: 10.5281/zenodo.1117724.

Westergaard, Marit; Terje Lohndal; and Artemis Alexiadou. 2016. The asymmetric nature of V2: Evidence from learner languages. Paper presented at the Comparative Germanic Syntax Workshop 31, University of Stellenbosch, December 2.

Kush
NTNU Norwegian University of Science and Technology
Department of Language and Literature
NO-7491 Trondheim, Norway
[dave.kush@ntnu.no]
[terje.lohndal@ntnu.no] (Lohndal)
[jon.sprouse@uconn.edu] (Sprouse)