No escape from the island: On extraction from complement *wh*-clauses in English[1]

Ken Ramshøj Christensen & Anne Mette Nyvad
Aarhus University

Abstract

In theoretical syntax, English complement *wh*-clause are considered syntactic islands which block extraction in an asymmetric way: Argument extraction is more acceptable than adjunct extraction. Though this pattern is often assumed to be universal, studies have shown that Danish (and other Mainland Scandinavian languages) may be exceptions. It has also been argued that the patterns of (un)acceptability are biased by expert intuitions. We present data from 100 native speakers of English which confirms (i) that English complement *wh*-clauses are islands, (ii) that there is a (subtle) argument-adjunct asymmetry, and (iii) that this acceptability pattern is not due to participant bias. Together with earlier findings on Danish, these results are compatible with an island account that relies on parametric variation in the possibility of CP-recursion.

1. Introduction: The standard pattern

It has been reported numerous times that extracting an argument (e.g. *what* or *which*) from a complement *wh*-clause is more acceptable than extracting an adjunct (such as *how* or *where*), though neither is considered completely acceptable in English, as illustrated in (1), taken from Rizzi (1990: 4):

---

(1) a. ??<u>Which problem</u>$_1$ do you wonder [<u>how</u>$_2$ John could solve __$_1$ __$_2$]?
    b. *<u>How</u>$_2$ do you wonder [<u>which problem</u>$_1$ John could solve __$_1$ ___$_2$]?

Extracting an argument, as in (1)a, is traditionally explained as a Subjacency violation (Haegeman 1994: 402), because *what* crosses two IPs. What makes adjunct extraction, as in (1)b, worse is that in addition to the Subjacency violation, it also violates the Empty Category Principle (Haegeman 1994: 442), because the trace of *how* is not lexically governed.

Both violate the general principle of locality (cf. the *Minimal Link Condition* (Chomsky 1995: 311) and *Relativized Minimality* (Rizzi 1990)) because movement of the *wh*-element to the matrix spec-CP 'skips' the intermediate spec-CP in the embedded clause. Movement must take place in successive cyclic (i.e. local) steps, cf. the *Successive Cyclic Hypothesis* (Poole 2011: 160), a principle which is independently supported with data from many cross-linguistic studies (Kayne & Pollock 1978; Torrego 1984; Chung & McCloskey 1987; Henry 1995), language acquisition studies (Felser 2004; Crain & Thornton 1998), and psycholinguistic studies (Gibson & Warren 2004; Marinis et al. 2005).

The argument/adjunct asymmetry in *wh*-island extraction, as in (1), is assumed to be universal. However, as discussed below, it has been argued that at least some languages allow both types of extraction without asymmetry (Christensen, Kizach & Nyvad 2013a). It has also been argued that the expert intuitions on which syntactic theory is based are flawed due to confirmation bias (syntacticians presumably want the data to support their theory) (Dąbrowska 2010). It could also be that such intuitions are affected by knowledge of other languages (due to mere exposure to foreign languages or outright bilingualism) (Bohnacker 2006; Booth, Clenton & Van Herwegen 2018). In short, the question is whether we can be confident that there is an underlying universal constraint that results in the grammaticality judgements in (1), and if not, what about locality?

Studies on extraction from complement *wh*-clauses in Danish (Christensen, Kizach & Nyvad 2013a; Christensen, Kizach & Nyvad 2013b) suggest that such structures are not islands but may simply be very difficult to process. The results showed that sentences involving movement out of an embedded *wh*-clause which is uncontroversially grammatical, as in (2), are less than fully acceptable, and that it is more acceptable to extract an argument than an adjunct.

(2) a.   <u>Which problem</u>₁ do you think [that John could solve ___₁]?
   b.   <u>How</u>₁ do you think [that John could solve the problem ___₁]?

These effects naturally follow from processing considerations, since movement as such increases working memory load which reduces acceptability, and because some of the fronted *wh*-elements could (temporarily) be misconstrued as complements of the matrix verb (cf. also Fodor & Inoue's (1998) *Attach Anyway* heuristic, Frazier & Clifton's (1989) *Active Filler Hypothesis*, and lingering garden-path interpretations (Ferreira, Christianson & Hollingworth 2001)). The fronted argument *wh*-elements were Matrix Verb Compatible [+MVC], i.e. compatible as arguments/adjuncts of the matrix verb, and the temporary interpretation (attachment) at the matrix verb is well-formed (*What did she know?*). The adjunct *wh*-elements, on the other hand were matrix verb incompatible [–MVC], i.e. not compatible as arguments/adjuncts of the matrix verb, and the temporary interpretation at the matrix verb is anomalous (*Where did she know?*). The temporary anomaly induced by matrix verb incompatibility further decreases acceptability. The results in Christensen, Kizach & Nyvad (2013a) also showed effects of trial such that the participants found the island violations slightly more acceptable as a function of exposure (i.e. an amelioration effect). This was also the case for uncontroversially grammatical long movement of arguments and adjuncts, as in (2), but crucially, not for clearly ungrammatical sentences.

It has been argued that Danish allows extractions from a range of structures that are normally considered islands, possibly due to a syntactic parameter that allows recursive CPs in Danish, but not in English (Nyvad, Christensen & Vikner 2017; Vikner, Christensen & Nyvad 2017). In a nutshell, the argument is that Danish (and potentially the other Mainland Scandinavian languages) have the option of a recursive functional *c*P-layer ('little *c*P'), which allows extraction by providing extra specifier positions and complementizer stacking; all subordinate clause types (embedded clauses headed by an overt or non-overt complementizer, embedded *wh*-questions, clauses that are complements of nouns, and relative clauses) are *c*Ps ('little *c*Ps), whereas 'big' CP is only found in (embedded as well as main) V2 clauses. Modern English does not allow multiple complementizers in the same minimal clause (such as, *because that*, *if that*, *which that*), whereas Middle English did (Vikner 1995: 121–122). In Danish, it is ubiquitous: *fordi at* (because that), *hvis at* (if that), *som at*

der (which/that that that) (Nyvad 2016). However, it is also conceivable that extractions from *wh*-clauses in English are just difficult to process but not ungrammatical, as is arguably the case for Danish. In this paper, we present the results from a study on extraction from complement *wh*-clauses in English using the same experimental setup and design as in our studies on *wh*-islands in Danish.

There is also an ongoing debate about the nature of data that has traditionally been used in generative syntax. For example, Branigan & Pickering (2017: 4–5) argue that the "standard" approach to data collection, which they claim is "to ask a single informant about the acceptability of a few sentences", is fundamentally flawed. It is open to various sorts of bias from the informant, who might be influenced by what they know about linguistic theory or what they think about the information-seeking linguist; see also Gibson & Fedorenko (2010). However, all of these objections to the (caricature of the) "standard" approach have been answered in detail, and there seems to be no real reason to suspect that generative syntactic theory is based on false assumptions and flawed acceptability judgments (Featherston 2009; Phillips 2009; Sprouse & Almeida 2017; Christensen 2019). Some of the concerns should still be taken seriously, however. One concern is that there is good reason to carry out experiments with many examples and many participants when examining subtle contrasts in acceptability in order to avoid participant or expert bias (Gibson, Piantadosi & Fedorenko 2013); another concern is that expert intuitions may also be biased (Dąbrowska 2010).

Following the generally accepted assumption that complement *wh*-clauses are weak islands in English, i.e. they exhibit a selective, non-uniform extraction pattern (Szabolcsi 2006), and the uncontroversial assumption that movement in itself increases processing load, we made the following set of predictions:

Prediction 1: There are processing effects: Movement per se increases processing load which decreases acceptability (which is not an effect attributable to the specifics of the grammar of English).

Prediction 2: Complement *wh*-clauses are islands, and extraction leads to consistent ungrammaticality or (at least) severely reduced acceptability. Therefore, (non-local) movement across a *wh*-element in the embedded spec-CP is significantly less acceptable than long (local, successive-cyclic) *wh*-movement. We assume that English *wh*-islands are 'real' islands (i.e.,

what blocks extraction is structural and absolute, not a matter of processing load) and as such immune to lexical effects, and therefore, the level of acceptability of extraction from a *wh*-island does not correlate with the frequency of occurrence of the matrix verb.

Prediction 3: The acceptability pattern for extraction out of a *wh*-island is asymmetric. Argument extraction is more acceptable as it 'only' violates locality (or rather, Subjacency), whereas adjunct extraction is less acceptable because it also violates the Empty Category Principle.

Prediction 4: The pattern is not due to participant bias, neither expert bias (effect of being a linguist), nor repetition (effect of trial).

There might potentially be sociolinguistic factors that affect the acceptability judgments. It could be that there is variation between different varieties of English, or that there are overall differences between participants of different age or level of education. It is also possible that there is transfer from one language to another in bilinguals. To test for (and to control for) these possibilities, we also looked at the main effects of bilingualism, nationality, age, and level of education of the participants. However, since we did not have any theoretically or empirically motivated hypotheses about how any of these particular factors might specifically influence island extractions, we did not look for interaction effects. Their potential main effects were included as controls.

2. Experiment
2.1 Participants
The task description specified that participants must be native speakers of English, and the survey itself also contained a control question requiring participants to confirm they were native speakers. Only responses from native speakers were included in the analysis. In total, 122 persons participated in our online survey, which was sent to various Facebook forums for people interested in English (e.g. university English departments). In the analysis, we included only responses from people aged 11–100 with 8–29 years of education, and only nationalities with more than 10 participants. This filtering resulted in 100 native speakers of English (male 52, female 48; linguists 57, non-linguists 43); nationality: 45% from the UK, 45% from the USA, 10% from Canada; participants per list: 10, 13, 11, 13, 15, 38), mean age 42.6 years (range=18–81, SD=17.7) with a mean length of education of 19.5 years (range=12–27, SD=3.2).

2.2 Materials

The target stimuli consisted of 72 sentences embedded in a total set of 140 sentences: 12 sets of six target types as illustrated in Table 1: Baseline (no movement), Long ARG (argument extraction from the embedded clause), Long ADJ extraction (adjunct extraction from the embedded clause), Across ARG (island violation by argument extraction), Across ADJ (island violation by adjunct extraction), Anomaly (ungrammatical). All sentences were carefully constructed such that the matrix verb was incompatible with the *wh*-phrase in order to avoid (as far as possible) interpreting the sentences as local, matrix clause questions.

Example	Type
The mother explained that they should treat the children very leniently.	Baseline
Which children did the mother explain that they should treat very leniently?	Long ARG
How leniently did the mother explain that they should treat the children?	Long ADJ
Which children did the mother explain how leniently they should treat?	Across ARG
How leniently did the mother explain which children they should treat?	Across ADJ
The mother explained how leniently which children they should treat.	Anomaly

Table 1: Examples of the six types of sentences in the stimulus set.

All sentences were in the simple past tense, and the number of words was kept constant (except the interrogative structures which triggered the addition of dummy-*do*).

The sentences were distributed evenly over six lists, making sure that each participant saw each matrix verb only once (and hence, judged only one member of each quadruple). The same 20 fillers occurred on all lists, such that each list consisted of 40 sentences in randomized order. The six lists were presented as online surveys using Google Drive. Each participant chose a list based on the month of their birthday: January–February = list 6, March-April = list 5, etc.

2.3 Procedure

The task consisted of acceptability judgments on a five-point Likert scale from 1 (completely unacceptable) to 5 (completely acceptable). Participants were instructed to base their judgments on their own intuition, not on what they might expect to be correct or standard language, and to ignore punctuation. The instructions also included the following examples

of a completely unacceptable sentence (1 on the scale) and a completely acceptable one (5 on the scale), respectively:

(3) a.   *What kind of food did the truck explains that the mule died?
    b.   The child often broke the rules.

2.4 Results

Using R (R Core Team 2017) with the *lmerTest Package* (Kuznetsova, Brockhoff & Christensen 2017) and the *MASS Package* (Venables & Ripley 2002), the results were subjected to a linear mixed-effects analysis with sliding contrasts to compare the neighboring levels in the type factor. To control for effects of frequency of occurrence of the matrix verb (Christensen & Nyvad 2014), our model included the mean of the z-transformed frequencies of each verbs in the British National Corpus (Davies 2004) and in the Corpus of Contemporary American English (Davies 2008). The maximal model with all random intercepts and slopes (Barr et al. 2013) failed to converge as did the zero-correlation parameter model (Bates et al. 2015). The maximal converging model included random intercepts for participants and items and random slopes for trial by participants. The mean acceptability ratings are presented in Figure 1, and the results of the statistical analysis are presented in Table 2.
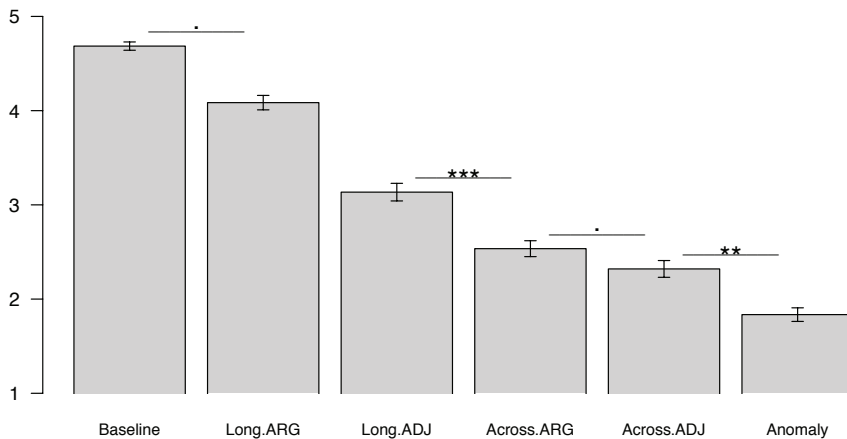

Figure 1: Mean acceptability ratings per type across items and participants. ***significant $p < 0.001$, **significant $p < 0.01$, *significant $p < 0.05$ (·marginal $p < 0.1$). Error bars ±1 standard error.

	Estimate	SE	df	*t*-value	*p*-value	
Long ARG – Baseline	-0.83	0.45	63.1	-1.82	0.074	·
Long ADJ – Long ARG	0.91	0.55	61.8	1.65	0.104	
Across ARG – Long ADJ	-2.25	0.58	62.4	-3.89	0.000	***
Across ADJ – Across ARG	0.83	0.47	59.9	1.76	0.083	·
Anomaly – Across ADJ	-1.31	0.42	62.2	-3.12	0.003	**
Age (in years)	-0.01	0.00	89.5	-2.24	0.027	*
Education (in years)	0.00	0.02	91.6	-0.01	0.996	
Bilingual (Yes – No)	0.04	0.11	91.5	0.36	0.722	
Nationality (UK – Canada)	0.25	0.19	91.4	1.30	0.197	
Nationality (USA – Canada)	0.07	0.19	90.3	0.37	0.711	
Baseline x Freq	-0.12	0.30	39.9	-0.40	0.689	
Long ARG x Freq	0.10	0.34	64.0	0.31	0.757	
Long ADJ x Freq	-0.92	0.38	58.3	-2.46	0.017	*
Across ARG x Freq	-0.07	0.37	62.8	-0.20	0.843	
Across ADJ x Freq	0.23	0.38	56.9	0.61	0.547	
Anomaly x Freq	-0.03	0.32	53.1	-0.08	0.933	
Baseline x Trial	0.01	0.01	60.0	0.71	0.481	
Long ARG x Trial	0.01	0.01	49.8	1.16	0.252	
Long ADJ x Trial	-0.06	0.02	55.9	-2.74	0.008	**
Across ARG x Trial	0.01	0.01	57.5	0.70	0.488	
Across ADJ x Trial	-0.03	0.01	52.6	-2.09	0.042	*
Anomaly x Trial	0.02	0.01	54.5	1.49	0.143	
Baseline x Linguist (Yes)	0.26	0.18	314.1	1.49	0.139	
Long ARG x Linguist (Yes)	0.53	0.17	312.9	3.06	0.002	**
Long ADJ x Linguist (Yes)	0.10	0.17	316.3	0.56	0.576	
Across ARG x Linguist (Yes)	0.23	0.17	314.1	1.30	0.194	
Across ADJ x Linguist (Yes)	-0.15	0.17	313.6	-0.88	0.381	
Anomaly x Linguist (Yes)	-0.31	0.17	302.8	-1.79	0.075	·

Table 2: Summary of fixed effects. 'Estimate' indicates the relationship between acceptability rating (the output) and each of the contrasts (between the sentence types) and interactions (between type and trial). SE= standard error, df=degrees of freedom, ***significant p<0.001, **significant p<0.01, *significant p<0.05, (·marginal p<0.1).

3. Discussion

Based on the existing theoretical and experimental literature on islands, we made four predictions:

Prediction 1: There are processing effects: Movement per se increases processing load which decreases acceptability (which is not an effect attributable to the specifics of the grammar of English). This was **confirmed**. There was a marginally significant drop in acceptability for long movement compared to the baseline condition. Unlike our previous studies on Danish (Christensen, Kizach & Nyvad 2013a; Christensen, Kizach & Nyvad 2013b), the difference between long argument and long adjunct movement was not statistically significant, though the trend is in the same direction. This is most probably because there was more variation in the adjunct condition than in the argument condition. Controlling for MVC, Nyvad, Kizach & Christensen (2014) also found that adjunct extraction was less acceptable than argument extraction (both [–MVC]) from an embedded declarative clause. The data could be taken to suggest that it is more difficult to integrate an incompatible [–MVC] adjunct, cf. also that agrammatic speakers seem to have an adjunction deficit: They prefer predicative adjectives over attributive ones, and they are significantly slower at integrating adjuncts than arguments (Lee & Thompson 2011; Meltzer-Asscher & Thompson 2014). (However, Nyvad, Kizach & Christensen (2014) found no difference for non-aphasic speakers in processing time for integrating arguments versus adjuncts.) In a similar vein, Hofmeister (2007: 56) states that adjunct questions "typically demand more effort for constructing the relevant existential presupposition and imagining an appropriate discourse for the question". The same intuition underlies the argument/adjunct asymmetry proposed in the theoretical syntax literature: In spite of the locality violation in (1)a, it is still possible to reconstruct the base-position for the extracted *wh*-element because it is selected by the embedded verb (the verb provides an identifiable empty slot in its argument structure); in (1)b, on the other hand, the base-position of *how* cannot as easily be reconstructed because, being an adjunct, it is not selected by the embedded verb and consequently, the base position is structurally indeterminate.

Prediction 2: Complement *wh*-clauses are islands, and extraction leads to consistent ungrammaticality or (at least) severely reduced acceptability. Therefore, (non-local) movement across a *wh*-element in the embedded spec-CP is significantly less acceptable than long (local, successive-cyclic) *wh*-movement. In addition, we assume that *wh*-islands are immune to lexical effects, and therefore, the level of acceptability of extraction from a *wh*-island does not correlate with the frequency of occurrence of the matrix verb. This prediction was **confirmed.** Our experiment showed that on average, the participants found extractions from a *wh*-island unacceptable (argument extraction was rated 2.54, adjunct extraction 2.32 on a scale from 1 to 5) but significantly better than the ungrammatical controls (the Anomaly condition, which was rated 1.84). Furthermore, frequency did not have a positive effect on acceptability. The only significant effect of frequency is negative. The more frequent the matrix verb, the more degraded our participants judged long adjunct extraction.

Prediction 3: The acceptability pattern for extraction out of a *wh*-island is asymmetric. Argument extraction is more acceptable as it 'only' violates locality (or rather, Subjacency), whereas adjunct extraction is also more difficult to reconstruct (because it also violates the Empty Category Principle). This was to some extent also **confirmed**: There is a marginally significant trend (p=0.083), which is in line with the standard pattern in theoretical syntax. The fact that is only marginally significant (the *p*-value is above 0.05 but below 0.1) fits the intuition that the difference between '??' and '*' is rather subtle. (As also pointed out by Hubert Haider, p.c., this acceptability asymmetry can also be reduced to a processing effect; as argued above, all things being equal, the base-position of an extracted argument is easier to reconstruct (there is an easily identifiable empty slot in the embedded argument structure) than the base-position of an extracted adjunct (which is not selected). This processing asymmetry is present in both licit and illicit contexts of extraction.)

Prediction 4: The pattern is not due to participant bias, neither expert bias (effect of being a linguist), nor repetition (effect of trial). This prediction was **confirmed**. The acceptability judgments for island extractions were not affected by being a linguist (expert bias). The linguists in our study rated long argument movement as more acceptable than the non-linguists did, as the linguists found the anomalies marginally worse

than the non-linguists did. This is in line with Culbertson & Gross (2009), who present data showing that linguists and students who have taken one or more classes in theoretical syntax show more consistent judgements as a group than naïve participants; see also Sprouse & Almeida (2013; 2017). Similarly, Christensen, Kizach & Nyvad (2013a: 58) found that students who had taken a course in generative syntax responded much faster (a full 1.1 second) and found long extractions (from islands and non-islands) slightly more acceptable (0.1 point on a 5-point scale) than students with no background in generative syntax did (there was no difference in the acceptability ratings for ungrammatical sentences). In short, linguists (and students with syntax training) are faster and more consistent in their judgments because they have better understanding of the nature of the task. This is also supported by the lack of significant effect of level of education, as well as age, which had a small but significantly negative effect; post hoc analysis revealed that this was driven by a decrease in acceptability of adjunct extraction from islands). There was also no ameliorating effect of trial. On the contrary. The two types of adjunct *wh*-movement were actually perceived as less acceptable over time. (Cf. also that Snyder (2000) found a ameliorating 'training' effect for *whether*-islands, but NOT for *wh*-islands.) Finally, the results also showed that the acceptability pattern is stable across different varieties of English (no effect of nationality). The effect of bilingualism was also not significant.

Taken together, the results from our studies on Danish and English strongly suggest that there is parametric variation between the two languages regarding the structure of the CP-domain. In non-V2 contexts, Danish allows a more elaborate structure in the CP domain by means of a recursive functional *c*P-layer which provides an escape hatch for extraction from *wh*-islands. English, on the other hand, only allows a single CP layer and since there is only one specifier position, which is filled by a *wh*-element, extraction out of the clause is effectively blocked. This contrast is illustrated in (4a) and (4b) below:

(4) a. English: *Ungrammatical*
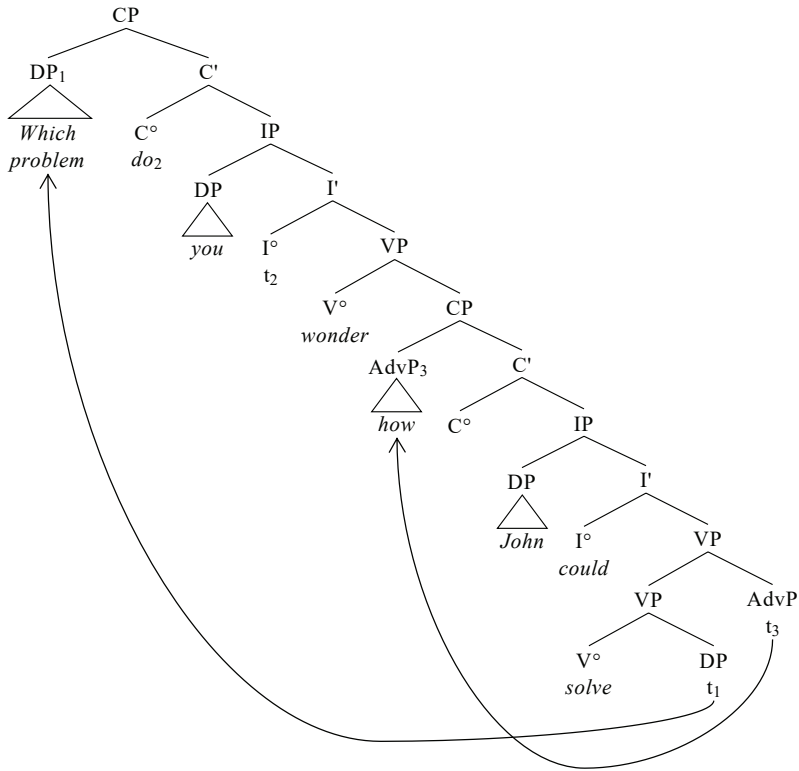
Ignoring differences due to verb movement, the key difference is the availability of an intermediate landing site for *wh*-movement in the Danish structure, which is not available in English. This is in line with the assumption that successive-cyclic movement is a universal principle. Without an intermediate landing site, movement from the embedded *wh*-clause is blocked. This analysis is fully compatible with the standard assumptions about clause structure in English as well as recent proposals about Mainland Scandinavian languages (Christensen, Kizach & Nyvad 2013a; Christensen, Kizach & Nyvad 2013b; Christensen & Nyvad 2014; Heinat & Wiklund 2015; Nyvad, Christensen & Vikner 2017; Vikner, Christensen & Nyvad 2017; Lindahl 2017).

(4) b. Danish: Difficult to process
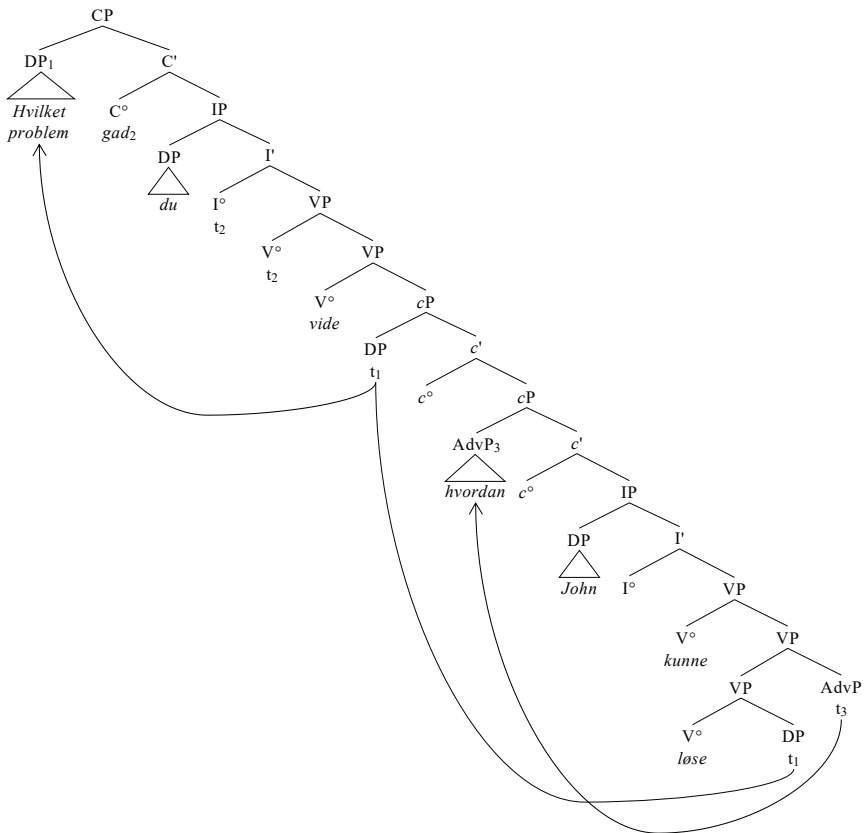
The option of *c*P-recursion may not be available in relation to all types of island constructions (adjunct islands, relative clauses, complex NPs, subject islands, *whether*-islands, etc.), as there appears to be some variation in the acceptability of extractions from these domains within and across the Mainland Scandinavian languages (Kush, Lohndal & Sprouse 2018; Kush, Lohndal & Sprouse 2019; Tutunjian et al. 2017).[2] Interestingly, however,

---

[2] As explained in the introduction, the option in Danish of a recursive functional *c*P-layer ('little' *c*P) that provides an extra specifier position as an escape hatch is available only in *subordinate* clause types. V2 clauses (embedded as well as main clauses), on the other hand, are 'big' CPs. The head of CP 'becomes' lexical when the finite verb moves into it. V2 is never selected (it is never required by a matrix verb), and it follows that there must be a projection above an embedded CP, namely a *c*P headed by a declarative comple-mentizer which does not provide an extra specifier. For details, see Nyvad, Christensen & Vikner (2017) and Vikner, Christensen & Nyvad (2017).

the present results from English native speakers corroborate (replicate) the intuitions from the 'armchair'.

4. Conclusions

Our results support the standard assumption in theoretical syntax that complement *wh*-clauses are weak islands in English. The argument/adjunct asymmetry, however, is only marginally significant, which could be taken as support for the assumption that the contrast is a subtle one between highly degraded (??) and ungrammatical (*). In conclusion, our results from the present experiment are compatible with the standard assumption in the generative syntax literature, namely that there is a universal island constraint that impedes extraction from finite complement *wh*-clauses in English. This confirmation, however, makes our results regarding *wh*-island structures in Danish all the more pertinent, and suggests that there may be parametric variation between English and Danish when it comes to the possibility of CP-recursion. The island is still there, and it is slightly better to extract an argument than an adjunct from it.

References

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. doi:10.1016/j.jml.2012.11.001.

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth & Harald Baayen. 2015. Parsimonious Mixed Models. *arXiv:1506.04967 [stat]*. http://arxiv.org/abs/1506.04967 (5 April, 2018).

Bohnacker, U. 2006. When Swedes begin to learn German: from V2 to V2. *Second Language Research* 22(4). 443–486. doi:10.1191/0267658306sr275oa.

Booth, Paul, Jon Clenton & Jo Van Herwegen. 2018. L1 – L2 semantic and syntactic processing: The influence of language proximity. *System* 78. 54–64. doi:10.1016/j.system.2018.07.011.

Branigan, Holly P. & Martin J. Pickering. 2017. An experimental approach to linguistic representation. *Behavioral and Brain Sciences* 40. doi:10.1017/S0140525X16002028.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, MA: MIT Press.

Christensen, Ken Ramshøj. 2019. On the need for experimental syntax. In Anne Mette Nyvad, Michaela Hejná, Anders Højen, Anna Bothe Jespersen & Mette Hjortshøj Sørensen (eds.), *A sound approach to language matters. Festschrift in honor of Ocke-Schwen Bohn*, 373–388. Aarhus University. http://ebooks.au.dk/index.php/aul/catalog/download/322/218/980-1?inline=1.

Christensen, Ken Ramshøj, Johannes Kizach & Anne Mette Nyvad. 2013a. Escape from the island: Grammaticality and (reduced) acceptability of *wh*-island violations in Danish. *Journal of Psycholinguistic Research* 42(1). 51–70. doi:10.1007/s10936-012-9210-x.

Christensen, Ken Ramshøj, Johannes Kizach & Anne Mette Nyvad. 2013b. The processing of syntactic islands – An fMRI study. *Journal of Neurolinguistics* 26(2). 239–251. doi:10.1016/j.jneuroling.2012.08.002.

Christensen, Ken Ramshøj & Anne Mette Nyvad. 2014. On the nature of escapable relative islands. *Nordic Journal of Linguistics* 37(1). 29–45. doi:10.1017/S0332586514000055.

Chung, Sandra & James McCloskey. 1987. Government, barriers, and small clauses in Modern Irish. *Linguistic Inquiry* 18(2). 173–237.

Crain, Stephen & Rosalind Thornton. 1998. *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics* (Language, Speech, and Communication). Cambridge, Mass: MIT Press.

Culbertson, Jennifer & Steven Gross. 2009. Are linguists better subjects? *The British Journal for the Philosophy of Science* 60(4). 721–736. doi:10.1093/bjps/axp032.

Dąbrowska, Ewa. 2010. Naive v. expert intuitions: An empirical study of acceptability judgments. *The Linguistic Review* 27(1). 1–23. doi:10.1515/tlir.2010.001.

Davies, Mark. 2004. *British national corpus (BNC)* (from Oxford University Press). https://www.english-corpora.org/bnc/.

Davies, Mark. 2008. *The corpus of contemporary American English (COCA): 560 million words, 1990-present*. https://www.english-corpora.org/coca/.

Featherston, Sam. 2009. Relax, lean back, and be a linguist. *Zeitschrift für Sprachwissenschaft* 28(1). 127–132. doi:10.1515/zfsw.2009.014.

Felser, Claudia. 2004. *Wh*-copying, phases, and successive cyclicity. *Lingua* 114(5). 543–574. doi:10.1016/S0024-3841(03)00054-8.

Ferreira, Fernanda, Kiel Christianson & Andrew Hollingworth. 2001. Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research* 30(1). 3–20. doi:10.1023/A:1005290706460.

Fodor, Janet Dean & Atsu Inoue. 1998. Attach anyway. In Janet Dean Fodor & Fernanda Ferreira (eds.), *Reanalysis in sentence processing*, 101–141. Springer Netherlands. doi:10.1007/978-94-015-9070-9_4.

Frazier, Lyn & Charles Clifton. 1989. Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4(2). 93–126. doi:10.1080/01690968908406359.

Gibson, Edward & Evelina Fedorenko. 2010. Weak quantitative standards in linguistics research. *Trends in Cognitive Sciences* 14(6). 233–234. doi:10.1016/j.tics.2010.03.005.

Gibson, Edward, Steven T. Piantadosi & Evelina Fedorenko. 2013. Quantitative methods in syntax/semantics research: A response to Sprouse and Almeida (2013). *Language and Cognitive Processes* 28(3). 229–240. doi:10.1080/01690965.2012.704385.

Gibson, Edward & Tessa Warren. 2004. Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax* 7(1). 55–78. doi:10.1111/j.1368-0005.2004.00065.x.

Haegeman, Liliane. 1994. *Introduction to government and binding theory*. 2nd edn. Oxford: Blackwell.

Heinat, Fredrik & Anna-Lena Wiklund. 2015. Scandinavian relative clause extractions apparent restrictions. *Working Papers in Scandinavian Syntax* (94). 36–50.

Henry, Alison. 1995. *Belfast English and Standard English: dialect variation and parameter setting*. New York: Oxford University Press.

Hofmeister, Philip. 2007. *Representational complexity and memory retrieval in language comprehension*. Standford CA: Standford University PhD dissertation.

Kayne, Richard & Jean-Yves Pollock. 1978. Stylistic inversion, successive cyclicity, and move NP in French. *Linguistic Inquiry* 9(4). 595–621.

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

Kush, Dave, Terje Lohndal & Jon Sprouse. 2019. On the island sensitivity of topicalization in Norwegian: An experimental investigation. lingbuzz/004442, ms. https://ling.auf.net/lingbuzz/004442 (24 April, 2019).

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

Lee, Jiyeon & Cynthia K. Thompson. 2011. Real-time production of arguments and adjuncts in normal and agrammatic speakers. *Language and Cognitive Processes* 26(8). 985–1021. doi:10.1080/01690965.2010.496237.

Lindahl, Filippa. 2017. *Extraction from relative clauses in Swedish*. Gothenburg: University of Gothenburg. https://gupea.ub.gu.se/handle/2077/51985 (14 May, 2018).

Marinis, Theodore, Leah Roberts, Claudia Felser & Harald Clahsen. 2005. Gaps in second language sentence processing. *Studies in Second Language Acquisition* 27(01). 53–78.

Meltzer-Asscher, Aya & Cynthia K. Thompson. 2014. The forgotten grammatical category: Adjective use in agrammatic aphasia. *Journal of Neurolinguistics* 30. 48–68. doi:10.1016/j.jneuroling.2014.04.001.

Nyvad, Anne Mette. 2016. Multiple complementizers in Modern Danish and Middle English. In Sten Vikner, Henrik Jørgensen & Elly Van Gelderen (eds.), *Let us have articles betwixt us – Papers in historical and comparative linguistics in honour of Johanna L. Wood*, 361–378. Aarhus: Dept. of English, School of Communication & Culture, Aarhus University. http://ebooks.au.dk/index.php/aul/catalog/view/119/107/477-1.

Nyvad, Anne Mette, Ken Ramshøj Christensen & Sten Vikner. 2017. CP-recursion in Danish: A *c*P/CP-analysis. *The Linguistic Review* 34(3). 449–477. doi:10.1515/tlr-2017-0008.

Nyvad, Anne Mette, Johannes Kizach & Ken Ramshøj Christensen. 2014. (Non-) arguments in long-distance extractions. *Journal of Psycholinguistic Research* 1–13. doi:10.1007/s10936-014-9300-z.

Phillips, Colin. 2009. Should we impeach armchair linguists? In Shoishi Iwasaki, Hajime Hoji, Patricia M. Clancy & Sung-Ock Sohn (eds.), *Japanese/Korean Linguistics 17*, 49–64. Stanford University: CSLI Publications. http://www.colinphillips.net/wp-content/uploads/2014/08/phillips2010_armchairlinguistics.pdf.

Poole, Geoffrey. 2011. *Syntactic theory*. 2nd edn. Houndmills: Palgrave Macmillan.

R Core Team. 2017. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Snyder, William. 2000. An experimental investigation of syntactic satiation effects. *Linguistic Inquiry* 31(3). 575–582. doi:10.1162/002438900554479.

Sprouse, Jon & Diogo Almeida. 2013. The empirical status of data in syntax: A reply to Gibson and Fedorenko. *Language and Cognitive Processes* 28(3). 222–228. doi:10.1080/01690965.2012.703782.

Sprouse, Jon & Diogo Almeida. 2017. Setting the empirical record straight: Acceptability judgments appear to be reliable, robust, and replicable. *Behavioral and Brain Sciences* 40. e311. doi:10.1017/S0140525X17000590.

Szabolcsi, Anna. 2006. Strong vs. weak islands. *The Blackwell Companion to Syntax*, 479–531. John Wiley & Sons, Ltd. doi:10.1002/9780470996591.ch64.

Torrego, Esther. 1984. On Inversion in Spanish and some of its effects. *Linguistic Inquiry* 15(1). 103–129.

Tutunjian, Damon, Fredrik Heinat, Eva Klingvall & Anna-Lena Wiklund. 2017. Processing relative clause extractions in Swedish. *Frontiers in Psychology* 8. doi:10.3389/fpsyg.2017.02118.

Venables, W. N. & B. D. Ripley. 2002. *Modern applied statistics with S* (Statistics and Computing). New York, NY: Springer New York. doi:10.1007/978-0-387-2.

Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.

Vikner, Sten, Ken Ramshøj Christensen & Anne Mette Nyvad. 2017. V2 and cP/CP. In Laura Bailey & Michelle Sheehan (eds.), *Order and structure in syntax I: Word order and syntactic structure*. 313–324. Berlin: Language Science Press. doi:10.5281/zenodo.1117724.