AN ACCEPTABILITY STUDY OF LONG-DISTANCE EXTRACTIONS IN SWEDISH

ANNA-LENA WIKLUND, FREDRIK HEINAT, EVA KLINGVALL AND DAMON TUTUNJIAN[1]

Abstract

The current study compares controlled acceptability judgments for Swedish relative clause extractions to extractions from *that*-clauses and extractions from non-restrictive relative clauses. It also compares each structure in both extracted and non-extracted form. The reported possibility of relative clause extraction in Swedish (and the other Mainland Scandinavian languages) has long presented a challenge to universal theories of constraints on extraction because the phenomenon is cross-linguistically very rare. In the off-line judgment data presented here, relative clause extractions are shown to pattern with extractions that are assumed to involve an island-like violation (non-restrictive relative clause extraction), thus contrasting with informal judgments reported in the literature. The data also appear to present a counterpoint to the conclusion reached in Tutunjian, Heinat, Klingvall and Wiklund (2017) from on-line eye-tracking measures, regarding the representational status of this structure as being more in line with that of a licit extraction. Potential explanations for the obtained patterns of the result are discussed, presenting avenues for further investigations.

---

[1] All authors contributed equally to this work. Author affiliations: E.K., D.T., and A.-L.W., Lund University. F.H., Linnæus University. Corresponding author: Damon Tutunjian, Center for Languages and Literature, Lund University, S-221 00 Lund (Sweden). E-mail: damon.tutunjian@englund.lu.se

1. Introduction

In both psycholinguistics and theoretical linguistics, considerable effort has been spent on investigating long distance filler-gap dependencies (e.g. Boeckx, 2012; Chomsky, 1973, 1986; Hofmeister, Casasanto & Sag, 2013; Kluender & Kutas, 1993; Omaki & Schulz, 2011; Phillips, 2006; Rizzi, 1990; Ross, 1967). Such dependencies can theoretically span an indefinite number of clauses and are therefore referred to as unbounded. Thus, in the well-formed sentence in (1a), the filler, *what*, has been extracted from a *that*-clause (*that Liz bought _*) − (where '_' indicates the gap) − which is embedded in yet another *that*-clause (*that John claimed*). At the same time, however, there are structural constraints that severely restrict these dependencies. In (1b) the extraction of the filler, *what*, from a *that*-clause (*that Liz bought _)* renders the sentence ill-formed, since the clause in this case is modifying a noun (*rumors*). Sentence (1c) is another example of an illicit extraction from a complex NP, i.e. a noun with a modifying relative clause. The extracted filler is the object of the verb *sells* in the relative clause. The relevant constraint has been referred to as the complex NP constraint in the literature (Ross, 1967), cf. Boeckx (2012) and Bošković (2014) for recent discussions.

(1) a. *What book did Mary say that John claimed that Liz bought _?*
   b. **What book did you hear rumors that Liz bought _?*
   c. * *What book did Mary know the man that sells _?*

The structural configurations that resist extractions as in (1b) and (1c) are called islands (Ross, 1967). The degraded acceptability that such extractions result in is sometimes called "island effects" (Sprouse & Hornstein, 2013). Island effects seem to be relatively stable cross-linguistically and therefore constitute phenomena that have been central to claims about universal principles of grammar. However, a small number of languages appear to exist as exceptions to this universality. For example, the Mainland Scandinavian languages (Danish, Norwegian and Swedish) allow extraction from complex NPs with relative clauses as in (1c) from Swedish (Allwood, 1976; Andersson, 1982; Christensen & Nyvad, 2014; Creider, 1987; Engdahl, 1997; Erteschik-Shir, 1973; Taraldsen, 1982). The topicalized NPs in (2) have all been extracted out of relative clauses (naturally occurring examples from Engdahl, 1997, p. 11).

(2) a. *Den teorin känner jag ingen som tror på _.*
 that theory know I nobody who believes in _
 "That theory, I don't know anyone who believes in it."

 b. *Där har jag en moster som bor _.*
 there have I an aunt that lives _
 "That place, I have an aunt who lives there."

 c. *Det finns det ingen som kan hjälpa mig med _.*
 that is there nobody that can help me with _
 "That thing, there is no one who can help me with that."

Research on relative clause extraction in Mainland Scandinavian shows that, given pragmatic well-formedness, it is possible to extract from relatives embedded under definite as well as indefinite head nouns (Engdahl, 1997) and from subject as well as object relatives (Heinat & Wiklund, 2015). Extraction of adjuncts/non-referential material is possible (Engdahl, 1997) and the matrix verb need not be a light verb or a small clause predicate (Christensen & Nyvad, 2014; Müller, 2015). In other words, there is no reason to suggest that relative clause extraction is marked in any obvious way in these languages. There is no consensus on how to analyze these structures. Attempts have been made to derive the facts from a discourse-based approach (Erteschik-Shir & Lappin, 1979), an island obviation approach (Cinque, 1990), and a small clause analysis (Kush, Omaki & Hornstein, 2013). However, none of these syntactic (or semantic) approaches has succeeded in accounting for the acceptability of Swedish RCEs (cf. Boeckx, 2012; Engdahl, 1997; and Müller, 2015, respectively), and thus Mainland Scandinavian continues to present a challenge for the universality of the ban on extraction and for theories of island constraints in general.

    One recent study provides some insights regarding the processing of Swedish relative clause extractions. Using an eye-tracking-while-reading paradigm, Tutunjian et al. (2017) investigated the processing of relative clause extraction (extracted RCs) in Swedish (cf. (2)) compared to extraction from *that*-clauses (extracted TCs), which is unproblematic along the lines of English (1a), and extraction from non-restrictive relative clauses (extracted NRCs), reported to be impossible in Swedish (Platzack, 2000; Teleman, Hellberg & Andersson, 1999; Smits, 1988) and hence arguably an island-like violation, (cf. (3) from Engdahl, 1997, p. 58).

(3) **Den teorin känner jag en man som för övrigt tror på _.*
 that theory know I a man that by-the-way believes in _
 "That theory, I know a man that, by the way, believes in."

In their materials, a speaker-oriented adverbial förresten "by the way", comparable to för övrigt in (3), was used to create a non-restrictive reading of the relative clause. The addition of this adverb signals upcoming new but secondary information, characteristic of non-restrictive relatives (Asher, 2000). It also renders extraction from the relative clause less acceptable (see the references above). This difference between NRCs and RCs is not unique to Mainland Scandinavian. In most other languages (e.g. English), extraction from both RCs and NRCs is impossible. However, differences between the two can still be seen when looking at the licensing of negative polarity items inside the relative, and the possibility of variable binding into the relative – impossible in NRCs but possible in RCs (cf. De Vries, 2007). Thus, extracted NRCs were good candidates for a less acceptable comparison point to extracted RCs, with the advantage of being surface-identical to extracted RCs, and thus reducing the possibility of any confounding factors.

Tutunjian et al. (2017) found that extracted RCs did not show similar patterning to the extracted NRC islands at the embedded RC verb where filler-gap integration is first possible, and instead often patterned together with extracted TCs. Their results have thus generally conformed to the informal judgments reported in the literature on Swedish relative clause extractions. However, it is still unknown to what degree processing results and informal judgments will conform to controlled acceptability judgments (but cf. Müller, 2015).

In the current study, we compare controlled acceptability judgments for extracted Swedish RCs to extracted island-like and non-island structures, using the same three extracted structures used in Tutunjian et al. (2017): RC, TC, and NRC. In our stimuli, the relativization present in RCs and NRCs is assumed to create a filler-gap dependency between the head noun and the subject gap in the relative clause (mediated by the relativizer). Filler-gap dependencies are known to tax the processor (Frazier & Clifton Jr, 1989; Gibson, 1998; Hawkins, 1999; Hofmeister & Sag, 2010; Kluender & Kutas, 1993; Lewis & Vasishth, 2005; Wanner & Maratsos, 1978), and we thus expect that RCs and NRCs will involve a layer of complexity not present in *that*-clauses. Note that we only assess RC and NRC holistically. We thus conflate the presence or absence or type of relativization with structure (TC, RC, and NRC), rather than examining relativization and non-restrictiveness components independently. We also manipulate extraction (extracted, non-extracted), allowing us to have comparison points which are free of any island-like violation (expected to be present at least for extracted NRCs). Extraction introduces a new level of complexity, but it is assumed to apply across the

board to all levels of structure.

Given the above, we expect the following: (i) that RCEs will elicit higher ratings than NRCEs; (ii) that all structures will be rated worse when in their extracted forms; (iii) that both relative clause variants will be rated worse than *that*-clauses in extracted and non-extracted forms, though there may also be differences between the two, given our assumption that they comprise different structures.

1.1. Method

1.1.1 Participants

Twenty-seven self-reported native speakers of Swedish, all students at Lund University, took part in the study in exchange for a cinema ticket. The age of the participants ranged from 19 to 32 years (mean 22.1, median 22).

1.1.2 Materials

In order to test our hypotheses, we constructed 30 experimental items, as in (4). We wanted to investigate two factors: structure (with three levels) and extraction (two levels), and each item appeared in six variant forms. The three levels of structure were: *that*-clauses (4a, b); restrictive relative clauses (4c, d); and non-restrictive relative clauses (4e, f). The two levels of extraction were: non-extracted (4a, c, e), and extracted (4b, d, f).

(4) a. *Vi noterade att en bekant   gärna odlade ovanliga blommor på*
we noted   that an acquaintance gladly grew rare   flowers on
kolonilotten.
allotment-the
"We noted that an acquaintance gladly grew rare flowers on the
allotment."

b. *Ovanliga blommor noterade vi att en bekant   gärna odlade på*
rare   flowers noted   we that an acquaintance gladly grew   on
kolonilotten.
allotment-the
"Rare flowers, we noted that an acquaintance gladly grew on the
allotment."

c. *Vi noterade en bekant   som gärna odlade ovanliga blommor på*
   we noted   an acquaintance who gladly grew rare   flowers   on
   kolonilotten.
   allotment-the
   "We noted an acquantaince who gladly grew rare flowers on the
   allotment."

d. *Ovanliga blommor noterade vi en bekant   som gärna odlade på*
   rare   flowers   noted   we an acquaintance who gladly grew on
   kolonilotten.
   allotment-the
   "Rare flowers, we noted an acquantaince who gladly grew on the
   allotment."

e. *Vi noterade en bekant   som förresten odlade ovanliga blommor på*
   we noted   an acquaintance who by-the-way grew rare   flowers on
   kolonilotten
   allotment-the
   "We noted an acquaintance who, by the way, grew rare flowers on the
   allotment."

f. *Ovanliga blommor noterade vi en bekant   som   förresten   odlade*
   rare   flowers   noted   we an acquaintance who by-the-way grew
   på kolonilotten
   on allotment-the
   "Rare flowers, we noted an acquaintance who, by the way, grew on the
   allotment."

In order to force a non-restrictive reading of the relative clause, the adverb *förresten* "by the way" was inserted in the relative clause. To make the sentences as similar as possible, we added an adverb in the other structures as well, e.g. *gärna* "gladly" in the examples above.

The sentences were rotated across three presentation lists, so that each participant saw one extracted and one non-extracted sentence from each item, although not of the same structural variant. The critical items were interspersed with 30 distractor items, of which ten were grammatical, as in (5a), and 20 ungrammatical, as seen in (5b). The ungrammatical sentences all involved word order violations where the verb was preceded by two constituents rather than just one, thereby violating the verb-second constraint (Holmberg, 2015).

(5) a. *Uttjänta mobiltelefoner reparerade han inte gärna   utan   att få*
   used-up mobile-phones repaired   he   not gladly   without to get
   rejält betalt.
   considerably paid
   "Mobile phones that had seen their best days, he didn't want to repair
   unless he was well-paid."

   b. **Frostskadade vinbär   en dam plockade på landet   för att*
   frostbitten   red-currants a lady picked   in countryside-the for to
   ha   till vintern.
   have for winter-the
   "Frostbitten red currents a lady picked in the countryside to have for the
   winter."

Since the experimental lists were implemented in Google Forms, it was
not possible to truly randomize the items. To reduce the possibility of
ordering effects, each of the three presentation lists received three different
pseudo-randomizations for its items, yielding a total of nine list variants.
The first three sentences on each list were special warm-up items
constructed to display varied acceptability to make the participants
familiar with the task and the range of possible ratings. In total, each list
included 93 sentences of which 60 were critical sentences, 30 were
distractor sentences, and three were warm up items.

1.1.3 Procedure

The experimental lists were implemented in Google Forms and completed
under experimenter supervision in a group computer room. Each form
began with a short description of the rating task in which participants were
instructed to rate the sentences at their own pace using a seven-point
Likert scale, ranging from completely impossible [1] to fully possible [7].
The instructions included example sentences (unrelated to our
manipulation) illustrating how different points on the scale might be used.

1.1.4 Results

One participant was excluded for later reporting not being a native speaker
of Swedish, and one participant was excluded for misunderstanding the
task, leaving a total of nine participants for one list, and eight each for the
remaining two lists.
   For our analysis, we used the lme4 package for linear mixed models
(Bates, Maechler, Bolker, Walker et al., 2014). We thus followed
Hofmeister and Sag (2010) and Sprouse, Wagers and Phillips (2012) in
using parametric tests for the analysis of Likert scale data. It should be

noted that Likert scale data are often assumed to be inherently ordinal in nature. As such, conducting parametric tests would require that the scale values be treated as having equal intervals, which could have implications for the normality of the distribution of the data (Agresti & Min, 2002). However, a number of researchers have challenged such claims, asserting that parametric tests such as linear mixed models can sufficiently accommodate any non-normality that might be present (Carifio & Perla, 2007; Norman, 2010) and that parametric tests may actually be more accurate than other approaches (Kizach, 2014). Based on this prior research, we proceeded with the assumption that linear mixed models provide a suitable means of analyzing ratings data.

To fit the fixed component of our models, we began with a "beyond optimal model" (Zuur, Ieno, Walker, Saveliev & Smith, 2009) which included two categorical predictors, Structure and Extraction, and their interaction as fixed factors; random intercepts for subject and items, and by-item and by-subject slopes for structure. However, the inclusion of random intercepts gave rise to convergence issues, thus requiring their removal from the models. The fixed component was then determined by comparing Log likelihood (ML) in a backwards stepwise fashion (Zuur et al., 2009) between models until we identified the best fit model (cf. Table 1 for the final model).

To investigate the hypothesis that sentences involving RCEs will be rated higher than NRCEs, a set of contrasts was constructed using the glht function in the multcomp package (Hothorn, Bretz, Westfall, Heiberger & Schuetzenmeister, 2011). To reduce the likelihood of type I error, we enlisted a conservative, "single-step" adjustment for multiple comparisons. This approach accounts for the correlations between parameter estimates and yields smaller p-values than the Bonferroni test (Bretz, Hothorn & Westfall, 2011). Results from the three different structures in their extracted and non- extracted forms are presented in Table 1 and Figure 1.

There was a simple effect of structure for the non-extracted sentences: TC and RC displayed significantly higher ratings than NRC. Furthermore, ratings for TC were significantly lowered when the structure appeared in its extracted form. A cost for extraction was also visible for both RC and NRC via the three significant interactions between structure and extraction. Thus, all extracted structures were rated significantly lower than their non-extracted versions. However, as the interactions show, the effect of extraction was not equivalent across levels of structure. Both RC and NRC incurred a higher cost of extraction than TC; for RC this higher cost was more pronounced than it was for NRC. This can be seen via RCs having a significantly more negative slope to extraction cost than NRC

(see Figure 1). As a result, RC no longer displayed a significant difference from NRC in extracted form.

Table 1. *Linear effects model analysis of Likert scale (1-7) ratings.*

Non-extracted

Linear Hypotheses	Estimate	SE	z - value
StructRC - StructTC = 0	- 0.2298	0.1293	- 1.777
StructNRC - StructTC = 0	- 1.3454	0.1293	- 10.404***
StructRC - StructNRC = 0	1.1157	0.1293	8.627***
ExtractE (intercept structure: TC) = 0	- 1.6414	0.1293	- 12.693***
RC:ExtractE - TC:ExtractE = 0	- 1.9448	0.1829	- 10.632***
StructNRC:ExtractE - StructTC:ExtractE = 0	- 1.0949	0.1829	- 5.986***
StructRC:ExtractE - StructNRC:ExtractE = 0	- 0.8499	0.1829	- 4.646***

Extracted

Linear Hypotheses	Estimate	Std. Error	z - value
StructRC - StructTC = 0	- 2.1746	0.1293	- 16.815***
StructNRC - StructTC = 0	- 2.4403	0.1293	- 18.870***
StructRC - StructNRC = 0	0.2658	0.1293	2.055†

Model: lmer (Rating ~ Struct*Extract + (1 | Subject) + (1 | Item))

p-values: '***' 0.001 '**' 0.01 '*' 0.05 '†' 0.1 - adjusted for multiple comparisons using the "single-step" method (Bretz et al., 2011).
Note: Intercept level for structure contrasts is always the term to the right. Intercept level for extraction contrasts is as indicated in italics. ExtractE = "extracted".
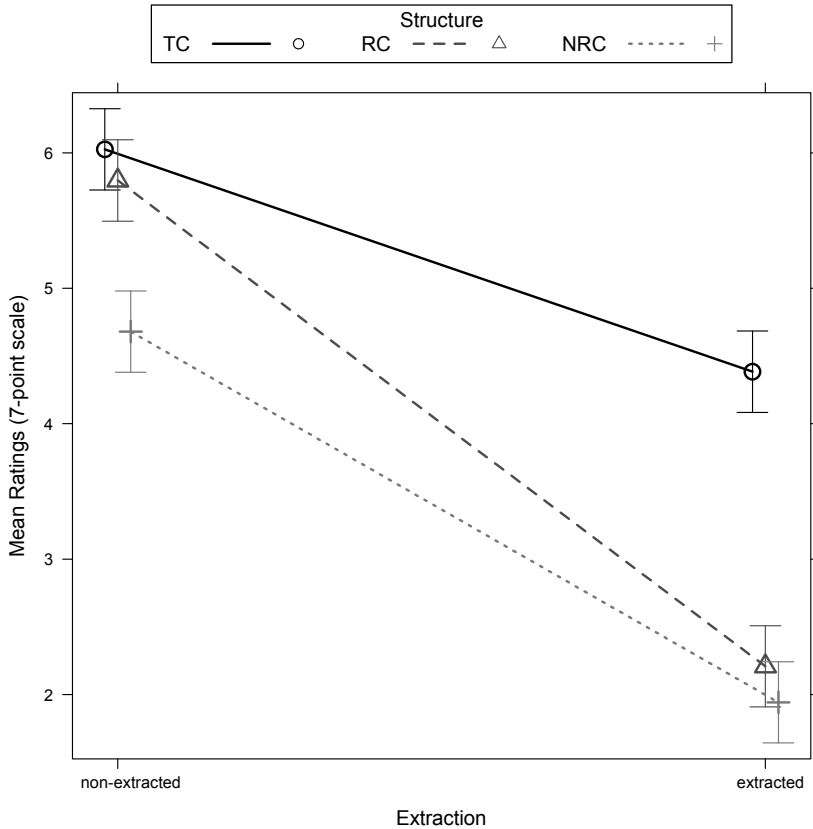
Figure 1. Mean Likert score ratings: extraction by structure interaction. Error bars represent the 95% confidence interval (CI) of the mean.

2. Discussion

The purpose of this study was to obtain controlled acceptability judgment data for three structures in their extracted and non-extracted forms: relative clauses (RC), *that*-clauses (TC), and non-restrictive relative clauses (NRC), to investigate whether the extracted form of RCs would be rated in a similar manner to extracted TCs, or instead pattern closer to extracted NRCs (which were expected to induce island-like effects on judgments). We expected that off-line judgments for these structures would align with the on-line measures obtained via eye-tracking in Tutunjian et al. (2017),

whereby RCEs patterned between NRCEs and TCEs or together with TCEs. We also expected that all structures would be rated worse in their extracted forms and that both relative clause variants would be rated worse than *that*-clauses in both their extracted and non-extracted forms.

Conforming to our expectations, sentences with extractions were rated as less acceptable than their non-extracted versions, irrespective of structure. We take this to be a reflection of the findings that filler-gap dependencies affect processing (Frazier & Clifton Jr, 1989; Gibson, 1998; Hawkins, 1999; Kluender & Kutas, 1993; Lewis & Vasishth, 2005; Wanner & Maratsos, 1978). Note that our stimuli were not preceded by context sentences and thus the present study cannot distinguish effects of extraction from those resulting from lack of (inter-sentential) contextual cues. Context has been shown to influence the acceptability of sentences where information structural processes have modified their base structures, as has been shown for passivation (Engelkamp & Zimmer, 2005, p. 60). This is likely to also be the case for extraction.

Regarding structural manipulation, which was the main focus in the present study, RCs' behavior in extracted form was unexpected. An interaction between structure and extraction was observed in which RCs incurred a greater cost of extraction than NRCs. This was driven by the fact that NRCs were rated as much worse than RCs in the non-extracted condition and that RCs drew closer to NRCs in extracted form, effectively resulting in non-significant differences in extracted RC mean ratings. The pattern displayed in the extracted form is in sharp contrast to informal judgments that extracted RCs are more acceptable than extracted NRCs. It is also in conflict with the eye-tracking data presented in Tutunjian et al. (2017).

The non-extracted RC/NRC difference may relate to the information of the structural status of NRCs as introducing peripheral or supplementary information. In our stimulus sentences, the head noun was always indefinite (thus introducing a new discourse referent). It is possible that RCs are preferred over NRCs with indefinite head nouns, as restrictive relative clauses have the effect of narrowing the domain of reference for the head noun. Given that the peripheral information expressed by the NRC does not help with the identification of the referent, speakers may find NRCs to be less acceptable than RCs with indefinites, at least from a pragmatic point of view.

The unexpectedly low rating for the extracted RCs is of greater interest for this study. A conclusion one might draw from these data is that RCEs show island-like effects in Swedish, despite previous reports to the contrary. However, a number of alternative explanations need to be

considered prior to accepting such a conclusion.

First, judgment tasks are inherently different from on-line tasks such as eye-tracking while reading. Task demands are more overt in judgment data, rather than in just passively reading sentences, since participants were asked to actively engage in a self-reflective activity that focused on their own language intuitions. Judgment data may therefore reflect metalinguistic performance rather than true competence (cf. Schütze, 1996, for discussion), and may allow for the influence of response biases (Graeff, 2002, 2003; Schütze, 1996) or other effects on cognitive processes arising from a self-awareness of mental states (e.g. Nelson, 1995). With such complex sentences, it is difficult to predict how a difference in method might translate to a difference in results, especially given that the structures in question are relatively under-studied from both methodologies. Moreover, a number of studies have shown that the grammaticality of a structure is no guarantee that it will elicit a high rating, especially when the structures are highly complex (Alexopoulou & Keller, 2007; Gibson & Thomas, 1996; Omaki & Schulz, 2011). Thus, off-line acceptability judgments potentially do not provide sufficient means of identifying differences between RC and NRC in the extracted forms, both being equally complex. Alternatively, acceptability judgments do provide adequate granularity to see the differences but are more susceptible than on-line measures to the presence or absence of contextual cues in the stimuli (cf. Bornkessel, Schlesewsky & Friederici, 2011). This possibility opens up a mismatch between the patterns of on-line measures and the patterns of off-line judgments.

In addition, the use of a pre-determined rating scale introduces the possibility of ceiling or floor effects. Extracted NRCs, for example, may not have had enough experimental space to display a difference from their extracted RC counterparts. This, however, does not appear to be a significant concern with our data, since low and high rating extremes for extracted NRCs never reached the informal threshold of 70% for identifying floor effects (Hyland, 2003).

A second possibility is that the acceptability of extracted RCs has changed from better to worse between generations. Most descriptions of extractions from relative clauses are from the early 1970s to the mid-1990s (for example: Allwood, 1976; Andersson, 1982; Engdahl, 1997; Erteschik-Shir, 1973). With one exception, the participants in our study were all born in the 1990s. Controlling for this factor would potentially enable the difference between RCE and NRCE to come out as significant for the older generation. Notably though, participants of the same younger age group did show the expected differences in Tutunjian et al. (2017).

Although such an explanation thus seems unlikely, controlling for the age of the participants is necessary to exclude this factor as contributing to off-line judgment patterns.

A third explanation relates to the filler NP which, in our stimuli, always consisted of an indefinite plural preceded by an adjective (e.g. ovanliga blommor, "unusual flowers"). It has been noted that the addition of elements that limit the set of possible referents to the filler enable or improve extraction from certain types of structure. This is referred to as D(iscourse)-linking in the literature (Rizzi, 2001; Szabolcsi, 2006). Notably, this effect has been observed for weak island-extractions. The effect of D-linking on strong islands is less clear despite literature claims (cf. Hofmeister & Sag, 2010). In view of a recent proposal that extracted RCs are weak island-extractions (Lindahl, 2015), the addition of a D-linking element such as a demonstrative element or a *wh*-element like *vilka* "which" (implying the existence of a set of contextually-determined entities from which to choose) could have potentially had a greater impact on the acceptability of the extracted RCs than on the acceptability of the extracted NRCs (assumed to be more island-like). This would allow for a divergence between the two relativized conditions. Thus, issues related to context and D-linking in our materials appear to offer one viable explanation, alternative to RCs being island-like, regarding why the difference between RCs and NRCs in their extracted forms did not come out.

Returning to the pattern of results, we also expected that sentences involving relativization (RC and NRC) should be rated as less acceptable than those involving *that*-clauses (TC), because of the filler-gap dependency created by relativization. This hypothesis was partly supported by the results. In the non-extracted condition, the acceptability of sentences involving NRCs was significantly lower than those involving TCs. However, the small difference between RCs and TCs was not significant (cf. Sprouse et al., 2012, for a similar result), suggesting that restrictive relativization is not costly enough to significantly lower the acceptability of RC in relation to TC.

3. Conclusion

The current study compared controlled acceptability judgments for Swedish relative clause extractions to extractions from *that*-clauses and extractions from non-restrictive relative clauses. It also compared each structure in both its extracted and non-extracted form. In the off-line judgment data presented here, relative clause extractions were seen to

pattern with extractions assumed to involve an island-like violation (non-restrictive relative clause extraction), in contrast to the informal judgments reported in the literature. The data also appear to present a counterpoint to the conclusion reached in Tutunjian et al. (2017) from on-line eye-tracking measures, regarding the representational status of this structure as being more in line with that of a licit extraction. The difference in patterns between off-line and on-line measures may be because off-line acceptability judgments do not provide sufficient means of identifying differences between RC and NRC. Alternatively, acceptability judgments may be more susceptible than online measures to the presence or absence of contextual cues in the stimuli. The implication for future research on relative clause extractions is that parallels should not immediately be assumed when drawing connections between off-line research (e.g. Hofmeister & Sag, 2010; Sprouse et al., 2012) and on-line research, and that particular attention should be paid to differences in stimuli constructions between studies. A good avenue for further research would be to investigate the impact of contextual cues on off-line patterns and to examine whether adding complexity to the filler (D-linking) would yield any difference in the pattern of results.

Acknowledgements

References

Agresti, A. & Min, Y. (2002). Unconditional small-sample confidence intervals for the odds ratio. *Biostatistics*, 3(3), 379–386.

Alexopoulou, T. & Keller, F. (2007). Locality, cyclicity, and resumption: At the interface between the grammar and the human sentence processor. *Language*, 110–160.

Allwood, J. (1976). The complex NP constraint as a non-universal rule and some semantic factors influencing the acceptability of Swedish sentences which violate the CNPC. In J. Stillings (Ed.), *University of Massachusetts Occasional Papers in Linguistics II*. Amherst, Ma.1-20.

Andersson, L.-G. (1982). What is Swedish an exception to? Extractions and island constraints. In Engdahl & Ejerhed (Eds.), *Readings on unbounded dependencies in Scandinavian Languages*, Almqvist & Wiksell International, Stockholm. 33–45.

Asher, N. (2000). Truth conditional discourse semantics for parentheticals. *Journal of Semantics*, 17(1), 31–50.

Bates, D., Maechler, M., Bolker, B., Walker, S. et al. (2014). *lme4: Linear Mixed-Effects Models Using Eigen and s4*. *R Package Version*, 1(7).

Boeckx, C. (2012). *Syntactic Islands*. Cambridge: Cambridge University Press.

Bornkessel, I., Schlesewsky, M. & Friederici, A. D. (2011). Contextual information modulates initial processes of syntactic integration: The role of inter-versus intrasentential predictions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(29), 871–882.

Bošković, Ž. (2014). From the Complex NP constraint to everything: On deep extractions across categories. Ms., University of Connecticut, Storrs. Downloadable at http://ling. auf. net/lingbuzz/002278.

Bretz, F., Hothorn, T. & Westfall, P. (2011). *Multiple Comparisons Using R*. Boca Ratón, FL: CRC Press.

Carifio, J. & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about likert scales and likert response formats and their antidotes. *Journal of Social Sciences*, 3, 106–116.

Chomsky, N. (1973). Conditions on Transformations. In S. Anderson & P. Kiparsky (Eds.), *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston. 232-286.

—. (1986). *Barriers*. Cambridge, MA: MIT Press.

Christensen, K. R. & Nyvad, A. M. (2014). On the nature of escapable relative islands. *Nordic Journal of Linguistics*, 37, 29–45.

Cinque, G. (1990). *Types of A'-Dependencies*. Cambridge, MA: MIT Press.

Creider, C. A. (1987). Structural and Pragmatic Factors Influencing the Acceptability of Sentences with Extended Dependencies in Norwegian. *Working papers in Linguistics* (4), 1–141.

De Vries, M. 2007. Invisible Constituents? Parentheses as B-Merged Adverbial Phrases. In: N. Dehé & Y.Kavalova (Eds.) *Parentheticals*, Amsterdam: John Benjamins. 203-234.

Engdahl, E. (1997). Relative clause extractions in context. *Working Papers in Scandinavian Syntax*, 60, 51–79.

Engelkamp, J. & Zimmer, H. D. (2005). *Dynamic Aspects of Language Processing: Focus and Presupposition*. New York: Springer.

Erteschik-Shir, N. (1973). *On the Nature of Island Constraints*. Unpublished doctoral dissertation, Cambridge, MA: MIT Press.

Erteschik-Shir, N. & Lappin, S. (1979). Dominance and the functional explanation of island phenomena. *Theoretical Linguistics*, 6(1), 41–86.

Frazier, L. & Clifton Jr, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, 4(2), 93–126.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1), 1–76.

Gibson, E. & Thomas, J. (1996). The processing complexity of English center-embedded and self-embedded structures. *Proceedings of the NELS 26 workshop on language processing*. MIT Working Papers in Linguistics.

Graeff, T. R. (2002). Uninformed response bias in telephone surveys. *Journal of Business Research*, 55, 251–259.

—. (2003). Exploring consumers' answers to survey questions: are uninformed responses truly uninformed? *Psychology and Marketing*, 20, 643–667.

Hawkins, J. A. (1999). Processing complexity and filler-gap dependencies across grammars. *Language*, 75(2), 244-285.

Heinat, F. & Wiklund, A.-L. (2015). Scandinavian relative clause extraction: Apparent restrictions. *Working Papers in Scandinavian Syntax*, 94(1).

Hofmeister, P., Casasanto, L. S. & Sag, I. A. (2013). Islands in the grammar? Standards of evidence. *Experimental Syntax and Island Effects*, 42.

Hofmeister, P. & Sag, I. A. (2010). Cognitive constraints and island effects. *Language*, 86(2), 366–415.

Holmberg, A. (2015). Verb second. In T. Kiss & A. Alexiadou (Eds.), *Syntax – Theory and Analysis* (Vol. 1). Berlin: Mouton de Gruyter. 342-382.

Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. & Schuetzenmeister, A. (2011). Multcomp: simultaneous inference in general parametric models. r package version 1.2-hosted on the Vienna university of economics and business website.

Hyland, M. E. (2003). A brief guide to the selection of quality of life instrument. *Health and Quality of Life Outcomes*, 1(1), 1.

Kizach, J. (2014). Analyzing likert-scale data with mixed-effects linear models–a simulation study. Poster presented at the Linguistic Evidence, 14.

Kluender, R. & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, 8(4), 573–633.

Kush, D., Omaki, A. & Hornstein, N. (2013). Microvariation in islands? In J. Sprouse & N. Hornstein (Eds.), *Experimental syntax and island effects* Cambridge: Cambridge University Press. 239-264.

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

Lindahl, F. (2015). Swedish relative clauses as weak island. In Lsa annual meeting extended abstracts (Vol. 6, pp. 13–1).

Müller, C. (2015). Against the Small Clause Hypothesis: Evidence from Swedish relative clause extractions. *Nordic Journal of Linguistics*, 38(1), 67–93.

Nelson, T. (1995). *Metacognition: Core Readings*. Allyn and Bacon.

Norman, G. (2010). Likert scales, levels of measurement and the awsf statistics. *Advances in Health Sciences Education*, 15(5), 625–632.

Omaki, A. & Schulz, B. (2011). Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition*, 33(04), 563–588.

Phillips, C. (2006). The real-time status of island phenomena. *Language*, 82(4), 795–823.

Platzack, C. (2000) A complement-of-N account of restrictive and non-restrictive relatives: the case of Swedish, A. Alexiadou, P. Law, A. Meinunger and C. Wilder (Eds.), *The Syntax of Relative Clauses*. Amsterdam: John Benjamins. 265-308.

Rizzi, L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.

—. (2001). Reconstruction, weak island sensitivity and agreement. In C. Cecchetto et al. (Eds.), *Semantic Interfaces: Reference, Anaphora and Aspect*, Chicago: CSLI. 145–176.

Ross, J. R. (1967). Constraints on variables in syntax. Unpublished doctoral dissertation, Cambridge, MA: MIT Press. ([published as 'Infinite syntax'. 1986])

Schütze, C. T. (1996). *The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology*. Chicago: University of Chicago Press.

Smits, R. (1988). The relative and cleft constructions of the Germanic and Romance languages. Unpublished doctoral dissertation, Catholic University of Brabant.

Sprouse, J. & Hornstein, N. (Eds.) (2013). *Experimental syntax and Island Effects*. Cambridge University Press.

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

Szabolcsi, A. (2006). Strong vs. weak islands. *The Blackwell Companion to Syntax*, 4, 479–531.

Taraldsen, K. T. (1982). Extraction from relative clauses in Norwegian. *Readings on unbounded dependencies in Scandinavian languages*, 43, 205–221.

Teleman, U., Hellberg, S. & Andersson, E. (1999). *Svenska Akademiens Grammatik*. Svenska Akademien/Norstedts ordbok (distr.).

Tutunjian, D., Heinat, F., Klingvall, E. & Wiklund, A.-L. (2017). *Processing Relative Clause Extractions in Swedish* (Under Review).

Wanner, E. & Maratsos, M. (1978). An ATN approach to comprehension. In M. Halle, J. Bresnan & G. Miller (Eds.), *Linguistic Theory and Psychological Reality*. Cambridge, MA: MIT Press. 119-161.

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A. & Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. New York: NY: Springer.