

## Two (non-) islands in Slovenian: A study in experimental syntax

Arthur Stepanov, Manca Mušič and Penka Stateva

Center for Cognitive Science of Language, University of Nova Gorica

*Abstract:* There exists a controversy in the literature and among the speakers of Slovenian concerning the grammaticality of wh-island and subject island constructions in this language. We conducted an acceptability rating study of wh-islands and subject islands in Slovenian, using the factorial definition of island as developed in Sprouse et al. (2012) and other works. This definition provides for a capability to isolate a true island effect while controlling for two complexity factors that potentially interfere in speakers' evaluation of the relevant sentences: the *length* of the respective movement dependency and the presence of an *island structure* itself. We found that 1) Slovenian speakers do judge the wh-island sentences worse than the respective controls, but the observed degradation cannot be attributed to a true island effect; 2) subject extraction out of a wh-island leads to a so called *reverse island effect* whereby the acceptability is higher than expected even if the above two complexity factors are taken into consideration; and 3) speakers are sensitive to the subject island effect, as predicted by the mainstream theories of syntactic locality. The results of our study contribute to establishing a solid empirical base for further theoretical investigations of the island effects, and raise new questions about the role of processing factors in speakers' evaluation of island constructions.

*Keywords:* syntactic island, experimental syntax, Subjacency, Empty Category Principle, Slovenian

### 1. Introduction

In this study we ask whether Slovenian observes basic wh-island and subject island effects. In the previous literature, the relevant constructions were reported as acceptable for Slovenian speakers, in contrast to their counterparts in English (Golden 1995, Golden 1996, Golden 1997a, 1997b). Other island constructions, such as adjunct or complex NP islands, were reported as unacceptable. If proved sustainable, this state of affairs may have interesting consequences for the theories of syntactic locality that have been at the heart of generative syntax for many years. More specifically, claims concerning circumventing island effects a priori challenge the universal character of constraints on syntactic movement and/or constraints on licensing the traces of movement such as Subjacency and/or the Empty Category Principle (see Section 2.2). Hence it is important to establish whether the reported aberrations are real – and then to face a task of adapting the existing theories of syntactic islands to accommodate these languages (e.g. Rizzi 1982), – or not, in which case factors that might potentially affect speakers' judgments of island sentences into the direction of their better acceptance must be determined.

To our knowledge, no systematic studies of wh-island and subject island effects in Slovenian have been reported since Golden's work. An elucidation and strengthening of the empirical paradigm on wh-islands and subject islands in Slovenian is, therefore, called for not only in order to clarify relevant syntactic judgments in this language, but also with a prospective theoretical goal in mind, to see whether and how the syntactic, as well as processing-related accounts of island structures currently developed in the literature may fare against those data. These considerations provided the main impetus for our present study.

In order to test for wh-islands and subject island effects in Slovenian we employ the *factorial definition* of islandhood (see below for details). The factorial definition enables one to isolate an island effect on the basis of a fine-grained experimental design informed by syntactic theory as well as certain processing considerations that enter into speakers' evaluation of island structures. It has been previously shown in the literature to be particularly suitable for tasks like ours, whereby empirical patterns concerning islandhood ought to be clarified on a larger scale and isolation of a true island effect becomes possible by controlling the relevant processing factors.

The article is organized as follows. Section 2 presents a short background on syntactic islands and current theories of syntactic locality aimed to capture the island phenomena. In Section 3 we consider some Slovenian data used in Golden's work to test whether Slovenian observes the standard island constraints. We then concentrate on the wh-islands and subject islands specifically. Section 4 outlines the factorial design. The experimental study itself is reported in Sections 5 and 6. In Section 7 we discuss our results and their theoretical implications for theories of syntax and sentence processing. Section 8 offers concluding remarks.

## 2. A short background on syntactic islands

### 2.1. Island effects

A remarkable fact about human language is the existence of *dependencies*, whereby two syntactically related elements in a sentence are separated by some intervening material. Wh-questions, in which a wh-phrase can be separated infinitely long from its base position, provide a classic case of such dependency:

- (1) a. Who(m) did they elect \_ as president?  
b. Who(m) did John think that Sara said that Homer should meet \_?

Since Chomsky (1964) and Ross (1967), the common understanding was that there are structural constraints on such extraction. Specifically, extraction is precluded from the following types of structural domains termed islands (not an exhaustive list):

- (2) a. WH-ISLAND:  
??What do you wonder [whether John bought \_\_ ]?  
b. SUBJECT ISLAND  
\*Who did [pictures of \_\_ ] attract everybody's attention?  
c. ADJUNCT ISLAND  
\*What did John go to bed [after he saw \_\_ on TV]?  
d. COMPLEX NP ISLAND  
\*What did you make [the claim that John bought \_\_ ]?

Structural constraints on extractions have been a central topic in the generative syntactic theory. The importance of this issue is highlighted by the fact that these seem to be one of the most uncontroversial manifestations of purely syntactic mechanisms operating in human language (note that the sentences in (2) are semantically fully coherent). The corresponding (un)acceptability patterns have then been replicated in great many languages, and the theoretical effort aiming at their proper and formally rigorous explanation has been driving the field of generative syntax since Ross' seminal work up to this day.

## 2.2. Locality principles

An island effect is standardly taken to signal a violation of some locality principle(s) of grammar. These principles are not (only) about linear locality: as (1b) demonstrates, a wh-phrase can in principle be separated from its base position by an indefinitely long linear distance. Rather, locality has to do with a ban on (wh-)movement crossing a structural boundary of a certain kind. For instance, in (2a), movement is precluded out of an embedded clause headed by a wh-phrase such as *whether* (technically speaking, *whether* is in Spec-CP of the embedded clause), in (2b) movement is preclude out of an NP in the subject position, in (2c) out of a clause (CP) which is a structural adjunct, and in (2d) out of a clause (CP) that is a complement to a noun. In early generative transformational theories, this special structural boundary was termed a bounding node (Chomsky 1973). The *Subjacency Principle* forbids crossing a certain amount of bounding nodes in one step of movement (Chomsky 1973), thus regulating how local the movement in each case must be. In long-distance dependencies involving several embeddings, as in (1b), Subjacency forces movement to proceed in several successive cyclic steps, each step crossing not more than the allowed number of bounding nodes. The inventory of bounding nodes was claimed to be subject to parametric variation: while the set [NP, IP (=inflection phrase)] was postulated for languages like English, a different set [NP, CP (=complementizer phrase)] was postulated in Italian, on the basis of (Rizzi 1982). It was reported there that Italian sentences with a single embedded CP did not manifest the expected wh-island and Complex NP island effects, while sentences with more than one embedded CP clauses did manifest these effects.

Descriptively speaking, “non-crossable” structural domains can either be complements of some lexical head, as is the case of wh-islands, or non-complements, such as subjects and adjuncts. Concerning the latter, Huang's (1982) proposed Condition on Extraction Domains (CED) predicts that extraction out of structural subjects as in (2b) and out of structural adjuncts as in (2c) must be treated uniformly, and lead to a similar kind of degradation, compared to extraction out of objects, e.g. complement clauses. The concepts of Subjacency and CED were later unified in the Barriers system of (Chomsky 1986) by introducing the notion of barrier, a structural node that fails to be licensed in a certain way in order to allow extraction out of it. Subjacency was then reformulated in the Barriers system as a principle that precludes movement across a barrier, if the latter emerges on the way of an otherwise licensed movement.

Alongside with Subjacency, however, there is a need for an additional principle to account for the asymmetry that arises with respect to the structural type of the extracted phrase. In particular, it is well known that extraction of adjuncts in English leads to a worse ungrammaticality than extraction of direct objects or subjects. Thus (3b) and (3c) feel markedly worse than extraction of the wh-object *what* in (3a). (e.g. Chomsky 1986; Lasnik and Saito 1992):<sup>1</sup>

---

<sup>1</sup> If crossing the wh-island boundary is not an initial step of wh-movement (cf. (3c)), wh-subjects actually tend to pattern with wh-objects in that extraction of either of those out of a wh-island leads to a somewhat lesser degree of degradation than extraction of wh-adjuncts. This is shown in (i) (see Lasnik and Saito 1992 for the details of the Barriers-based ECP account that predicts this change of pattern). We thank an anonymous reviewer for a helpful reminder concerning this issue.

- (i) a. ??What do you wonder [whether Mary thinks John bought \_\_\_]  
b. ??/?\*Who do you wonder [whether Mary thinks \_\_\_ bought that house]  
c. \*Why do you wonder [whether Mary thinks John bought the house \_\_\_]

- (3) WH-ISLAND:
- a. ??What do you wonder [whether John fixed \_\_\_ ]?
  - b. \*How do you wonder [whether John fixed the car \_\_\_ ]?
  - c. \*Who do you wonder whether \_\_\_ fixed the car]?

Since in all cases movement crosses the same amount of (bounding) nodes, this asymmetry in judgments cannot be attributed (only) to Subjacency. In Chomsky's (1986) Barriers-based system or Lasnik and Saito's (1992) modified version of it, the more degraded status of (3b) and (3c) is attributed to the *Empty Category Principle* (ECP), a principle requiring proper licensing of movement traces. Proper licensing (more correctly, proper government in the Government and Binding framework) of a trace requires either licensing it by a lexical head such as a verb, under theta-marking, or by some relevant antecedent in the structure, such as a moved wh-phrase within the same clause. In (3a), by hypothesis, the trace is licensed by the theta-marking verb *fixed*. In (3b) and (3c), however, there is neither theta-marking, nor possible antecedent licensing, since the only available potential licenser (*how* or *who*) is not in the same clause as its trace. Hence, an ECP violation ensues, in addition to a Subjacency violation. In Chomsky's (1986) system, the ECP thus operates over and above the Subjacency principle. On the assumption that violating two structural constraints amounts to stronger ungrammaticality than violating only one such constraint, the asymmetry in (3) can then be accounted for. Alternatively, in Lasnik and Saito's (1992) system, Subjacency is encoded into the definition of antecedent government (needed for satisfying the ECP), as both concepts are formulated in terms of a unifying concept of barrier. Thus, in Lasnik and Saito's system, a failure of antecedent government responsible for ECP implies the lack of Subjacency at the same time. In this sense, only one structural constraint is violated in (3c), namely, the ECP. It should also be noted that ECP violations alone (presumably, as in *\*Who do you think that \_\_\_ left?*) are generally felt by speakers to be more degraded than mere Subjacency violations. We refer the reader to the above two works for the details of the respective accounts.

The closely related Relativized Minimality approach to syntactic locality prohibits movement of an element across another phrase of the same type, viz. head across head, A- or A-bar elements across another A- or A-bar element, respectively (Rizzi 1990; Chomsky 1995). Thus, for instance, a wh-phrase cannot move over another wh-phrase. Among other things, this approach predicts wh-island effects a priori (cf. (2a)), which would then be instances of such illegitimate crossing. Any empirical deviations from the predicted pattern would then raise non-trivial concerns regarding Relativized Minimality and its interaction with other components of the grammar.

More recent, minimalist, models attempt to derive the effects of Subjacency, CED and Relativized Minimality in island structures from independent considerations pertaining to the core computational properties of syntactic derivation (Chomsky 1995, 2005; Uriagereka 1999b, among others), properties of the syntax-semantics interface (Truswell 2007; Bianchi and Chesi 2014), or both (Stepanov 2007) with a reasonable degree of success. For instance, the effects of Relativized Minimality can be attributed to a syntactic computation principle such as Attract Closest, whereby some feature of the head of a projection constituting the target of movement (for instance the interrogative Q feature of the complementizer), is matched with the closest corresponding feature down the syntactic tree, such as the wh-feature of the wh-phrase. This matching feature would then serve to satisfy some licensing requirement of the target (Chomsky 1995 and later works). A promising line of recent minimalist research aims to derive many properties of syntactic islands from the mechanism of constituent labeling in the bare phrase structure (e.g. Bošković 2015, Bošković 2016;

Cecchetto and Donati 2015). Some recent proposals also attempted to derive certain effects of the ECP (see e.g. Krivochen and Kosta 2013) although there is notably less progress in this area compared to the Subjacency/CED domains. Nevertheless, for the present purposes we will continue to use the terms Subjacency, CED and ECP at least in the metaphorical/descriptive sense without going into the details of the respective proposals, as a reference to whatever (system of) principles and properties underlies the relevant empirical contrasts.

### 2.3. *Variability in subject islands*

There is a growing understanding in the syntactic literature, based on empirical studies of many different languages, that not all islands are created equal. Specifically, while some island effects tend to persist on a cross-linguistic basis, other island effects seem to be more language-dependent, showing diverging acceptability patterns depending on the language. The first group of island effects includes at least 1) adjunct islands with finite clause adjuncts, and 2) complex NP islands. In these structural domains, wh-extraction (as well as other type of extraction, such as topicalization) seems to be universally precluded (see, however, Truswell 2007 for certain caveats regarding the typology of relevant adjunct island effects).

The second, language-dependent, group of island effects includes primarily subject islands. Stepanov (2007) provides a representative typology of languages in which the predictions of the CED account (and its later derivatives) do not appear to hold: while finite clause adjuncts in these languages yield domains opaque for extraction leading to strict ungrammaticality (similarly to (2c)), extraction out of subject in these languages in counterparts of (2b) is judged quite acceptable, and in any case, better than the former. Subsequent research pursued two alternative scenarios. A number of researchers have argued that the amelioration effect in subject islands is only apparent, and with proper adjustment of the relevant examples, the CED-predicted pattern is restored, in at least some of the languages considered in Stepanov's work (e.g. Jurka 2010; Uriagereka 2012). The other strand of research takes these effects as real, and tries to derive them from independent, possibly interacting, properties of grammars of the respective languages (Wexler and Culicover 1981; Takahashi 1994; Stepanov 2001). The reader is referred to these works for the details of the relevant theoretical accounts.

Yet another recently developing line of research maintains that variation in the subject island effect can be observed even intra-linguistically, or within one language. Polinsky et al. (2013), on the basis of off-line acceptability judgment tasks as well as online self-paced reading tasks on the English and Russian materials, argue that a relevant factor that affects the islandhood of the subject domains is unaccusativity of the corresponding verb: acceptability in judging extraction out of subjects falls down the path unaccusatives >> unergatives >> transitives, the sentences with unaccusatives being judged the best, and transitives the worst. In other words, unaccusative subjects are more transparent to extraction than the other. The bulk of their empirical support for this conclusion comes from the Russian data. The results of a similar study of English by these authors supported this conclusion only partially: while they observed a significant unaccusativity contrast in the on-line task (reading times), the global acceptability judgments only supported a marginal significance of the unaccusativity factor. This leaves open the question as to whether sensitivity to this factor is language-specific, or a result of a general property, in which case further clarifications are called for.

The present study is in line with a growing number of works suggesting that correct identification of an island effect may become quite a subtle empirical matter, and so requires an attentive and careful evaluation of acceptability of the relevant structures. Such proper evaluation may need to rely on data over and above the results based on the methodology of

eliciting judgments from individual speakers, standardly adopted in theoretical syntactic research. Whatever the chosen strategy of theoretical argumentation, a careful construction and robustness of the relevant empirical paradigm is paramount for the ultimate success of the inquiry into the nature of subject islands and their observed cross-linguistic variability.

### 3. Islands in Slovenian: State of the art

Slovenian is a South Slavic language that has so far received relatively little attention in the literature on syntactic islands. To our knowledge, Golden (1995), Golden (1996), Golden (1997a: ch.8), Golden (1997b) are the only systematic studies to date that discuss the Slovenian data in the context of Subjacency and ECP-based accounts of syntactic locality. Specifically, Golden reports the following pattern of acceptability in the four above mentioned island types (the examples and translations are from Golden 1995; the glosses are ours):

(4) WH-ISLAND:

- a. Kaj se je Peter spraševal, kako je Špela popravila \_\_?  
 what REFL AUX Peter wonder-PST how AUX Špela fix-PST  
 ‘\*What did Peter wonder how Špela fixed?’
- b. Kako se je Peter spraševal, kaj je Špela popravila \_\_?  
 how REFL AUX Peter wonder-PST what AUX Špela fix-PST  
 ‘\*How did Peter wonder what Špela fixed?’
- c. Kdo se je Peter spraševal, kako \_ je popravil avto \_?  
 who REFL AUX Peter wonder-PST how AUX fix-PST car  
 ‘\*Who did Peter wonder how fixed the car?’
- d. [Proti kateri teoriji] se je Peter trudil spomniti, komu je  
 against which theory REFL AUX Peter try-PST remember-INF who-DAT AUX  
 Špela mimogrede omenila, kdaj je Tone ugovarjal \_?  
 Špela in-passing mention-PST when AUX Tone argue-PST  
 ‘\*Against which theory did Peter try to remember to-whom Špela mentioned in  
 passing, when Tone argued?’

(5) SUBJECT ISLAND

- a. Čigavim predlogom se mu je [ugovarjati \_ na oddelčnih sestankih] zdelo  
 whose proposals REFL him AUX discuss-INF at departmental meetings seem-PST  
 nesmiselno?  
 pointless  
 ‘\*Whose proposals did to discuss at the departmental meetings seem pointless to him?’
- b. Zakaj se mu je [odpustiti kogarkoli z dela \_] zdelo nesprejemljivo \_?  
 why REFL him AUX fire-INF anyone from job seem-PST unacceptable  
 ‘(\*)Why did it seem to him unacceptable to fire anyone from work?’

(6) ADJUNCT ISLAND

- \*Kaj se oglasi pri nas, preden kupiš \_?  
 what REFL drop-by at us before buy-2SG.  
 ‘\*What do you drop by before you buy?’

(7) COMPLEX NP ISLAND

- \*Koga so vsi kritizirali [Petrov načrt [da povabi na srečanje \_]] ?  
whom AUX all criticize-PST Peter's plan that invite at meeting  
‘\*Whom did they all criticize Peter's plan to invite to the meeting?’

From the perspective of theories of islands based on Subjacency, CED and ECP, as well as their contemporary minimalist counterparts, the data pertaining to the adjunct island and complex NP island in (6)-(7) fall into the expected pattern of ungrammaticality. However, the wh-island and subject island acceptability data in (4)-(5) are quite unexpected. Consider these data in more detail.

In terms of Subjacency, the full grammaticality of (4) and (5) is not predicted, as the fronted wh-phrase crosses an embedded wh-island boundary in (4), and a subject island boundary in (5). One may attribute the acceptability of (5) to the unaccusative status of the relevant verb, along the lines of similar observations for Russian in the study of (Polinsky et al. 2013), but the question remains why this pattern is different from that in English and other languages.

From the point of view of CED, the Slovenian data are again puzzling, as they manifest a diverging pattern of acceptability between extraction out of adjuncts in (7) and out of subjects in (5). A priori, Slovenian seems to qualify for the group of languages that present a challenge to the unification idea of the CED (Stepanov 2007), or at least require a special explanation for this diverging pattern.

Another puzzle is presented by (4b) and (4c) from the point of view of the ECP. (4b) provides a direct contrast with the English (3b). In both cases the adjunct *how/kako* moves to the matrix clause past a wh-island boundary. The trace of *how/kako* is not properly licensed since there is no suitable lexical head (e.g. by theta-marking, see above), or an antecedent wh-phrase in the same clause that could license it. By the ECP, this state of affairs should lead to ungrammaticality. Furthermore, while violations of Subjacency are known to be often tolerated by native speakers, especially in the case of wh-islands (see Lasnik and Saito 1992 for discussion), the ECP has always been understood as an absolute constraint on trace licensing, whose violation leads to strong ungrammaticality. However, the Slovenian example (3b) is reported as fully acceptable. In Slovenian, there thus seems to be no contrast between an ECP configuration, Subjacency configuration, and control examples not featuring an island structure. Since the ECP is not known to be subject of parametric variation in cases of adjunct extraction, this state of affairs is a priori unexpected. Similar considerations apply with respect to extraction of subjects, as in (4c) vs. (3c).

The Slovenian data thus present an a priori striking constellation of empirical puzzles. If proved sustainable, the observed empirical generalizations are likely to become interesting material for theoretical advancements and provide new insights into the nature of island constraints. Unfortunately, the data reported above have so far not been scrutinized in later studies and it is not yet clear whether they can receive a status of proper empirical generalizations that can be utilized as input to syntactic theories. There seems to be little consensus both in the theoretical literature and among the native speaker community, concerning the status of wh-island and subject island structures in Slovenian. Our own preliminary inquiries revealed that, while some Slovenian speakers generally agree with Golden's reported judgments, other speakers, both linguists and not, often show a greater conservativity with regard to the wh-island and subject island structures. Many such speakers judge sentences such as (4)-(5) as unacceptable. Yet other speakers feel uncertain and/or

ambivalent with regard to their status (all of our polled speakers agree with Golden’s reported judgments on adjunct and complex NP islands as in (6)-(7)). Therefore, we felt that, prior to setting out on a theoretical quest for an interpretation of the reported patterns of syntactic locality in this language, it would be methodologically instructive to validate the relevant empirical contrasts on a larger scale, taking into account possible individual variability and contributing to their robustness.

#### 4. The factorial design for testing island effects

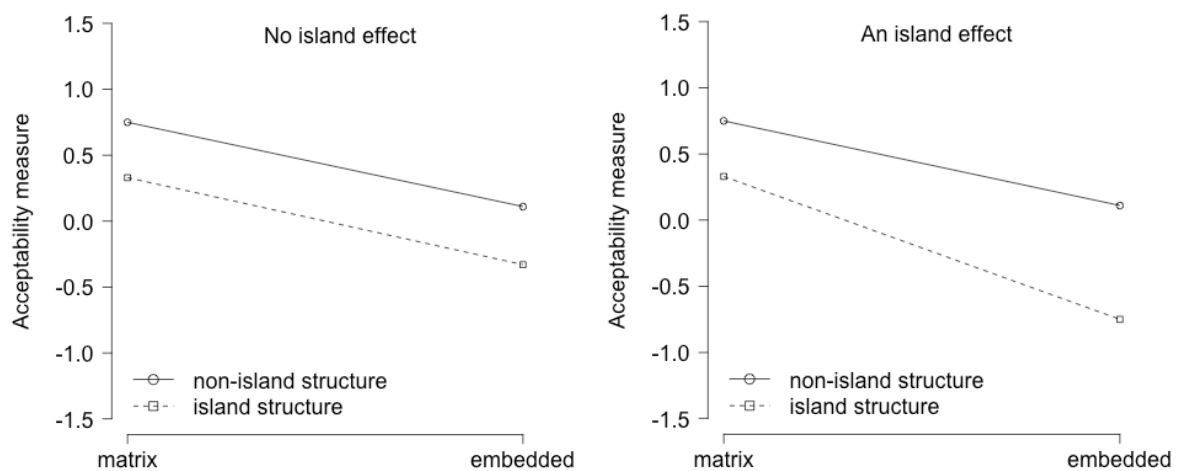
Empirical investigations on syntactic islands have recently received a boost by utilizing the factorial definition of island, advanced in a number of recent studies (Sprouse et al. 2011, Sprouse et al. 2012, Sprouse et al. 2016). The factorial design recognizes two (processing) factors that may enter into comprehension of island sentences. The first factor is the length of the respective movement dependency, that is, a distance between the moved wh-phrase and the gap (matrix, embedded), which we will refer to as LENGTH. The second factor is presence or absence of the complex island STRUCTURE itself. Each of these factors may contribute to the lower acceptability of the sentence, regardless of whether a true island effect obtains in a crucial sentence or not. The idea behind inclusion of the LENGTH factor is that longer sentences, or sentences with a greater distance between the head and the tail of a long-distance dependency, are more difficult to process, than shorter distance dependencies. Similarly, inclusion of the STRUCTURE factor is motivated by similar considerations that some complex structures induce more complex meanings (e.g. complex NPs or embedded wh-interrogatives) which, in turn, affects acceptability. This factorial definition of island then allows for an experimental design for testing for basic island effects in larger populations of speakers. A factorial design is thus a 2x2 design where these two factors are crossed (examples from Sprouse et al. 2016):

- |     |  |                       |
|-----|--|-----------------------|
| (8) | a. Who __ thinks [that John bought a car]?       | NON-ISLAND   MATRIX   |
|     | b. What do you think [that John bought __ ]?     | NON-ISLAND   EMBEDDED |
|     | c. Who __ wonders [whether John bought a car]?   | ISLAND   MATRIX       |
|     | d. What do you wonder [whether John bought __ ]? | ISLAND   EMBEDDED     |

This experimental model predicts that if there is no true island effect over and above each of these two factors, then the degree of degradation of the critical sentence (8d) should amount to the sum of the degradations incurred by the LENGTH and STRUCTURE factors, or  $(8a-8d) = (8a-8b)+(8a-8c)$ . If, on the other hand, a true island effect exists, it should reveal itself over and above the sum of effects incurred by each of the above two factors alone. In other words, a true island effect in this design amounts to a super-additive effect, whereby the overall measure of degradation of the critical sentence is greater than the sum of the degradations imposed by the length and island structure factors  $((8a-8d) > (8a-8b)+(8a-8c)$ ; see Sprouse et al. 2012 for details). This super-additive effect can be identified in at least three different ways. First, the numerical difference in the inequality above can be calculated by using the measure of *differences-in-differences* (DD), that is,  $DD = (8a-8d) - ((8a-8b)+(8a-8c))$ . If  $DD=0$ , there is no island effect; if  $DD > 0$ , there is an island effect (though the authors of the method do not provide a metric of islandhood based on the DD score). The second way of identifying an island effect in this design is visual: if there is no island effect, then an interaction based on the four conditions in (8) plotted according to the two factors will reveal two parallel lines, whereas the presence of an island effect will reveal two non-parallel lines (Figure 1). The third way of detecting an island is statistical: a super-additive effect reveals



itself as a statistically significant interaction of the LENGTH and STRUCTURE factors, in an ANOVA-type or linear (mixed-effect) model analysis.



**Figure 1.** Sample patterns illustrating the absence and presence of a true island effect, independent of the LENGTH and STRUCTURE factors (see Sprouse et al 2012 for details)

Sprouse et al. (2012) and Sprouse et al. (2016) apply the factorial design for experimental investigations of a number of island phenomena in English and Italian, demonstrating that super-additive effects obtain in a number of standard island configurations, in accord with the predictions of the syntactic theory. Thus the factorial design is well suited to empirical investigations of island effects in languages in which a more profound, fine-grained and quantitatively robust inquiry is called for. It also has the potential to identify novel and interesting patterns of data that have not been noticed before. To take one example, the results of the study in Sprouse et al. (2016) suggest that speakers of Italian are more sensitive to wh-island structures, than was previously thought (Rizzi 1982). We believe the factorial design is also appropriate for a fine-grained investigation of the island effects in Slovenian, which is why it was chosen for the present study.

## 5. The present study

### 5.1. Goals

In the context of the factorial design as outlined above, here we ask whether Slovenian wh-island and subject island structures actually manifest super-additive effects over and above the cumulative impact of the processing factors LENGTH and STRUCTURE. Irrespective of whether the relevant island effects obtain or not in Slovenian, we predict that length and structure should independently affect the acceptability of the respective sentences in this language, a natural expectation on the assumption that the cognitive mechanisms engaged in processing island structures are subject to the same parsing constraints and limitations (e.g. related to working memory) in any language, under the strictly incremental character of syntactic processing. However, if Golden's claim regarding the absence of island effects in Slovenian is on the right track, this absence should reveal itself in the formal ways described above. In particular, the respective DD score should equal or be around 0. Statistically, we expect that there will be no significant interaction between the two processing factors. In

contrast, if the data does show an island effect, then we expect  $DD > 0$  and a significant interaction between the two factors.

In addition, in our study we were interested in replicating the amelioration effect in subject island sentences with unaccusative verbs compared to other verb types, based on data from Russian, as reported in Polinsky et al. (2013).

### 5.2. Design, materials and questionnaires

In accord with the goals outlined above, we targeted the following four construction types:

- (9)
  - a. Wh-islands, featuring extraction of wh-objects (*kaj, koga* “what, whom”)
  - b. Wh-islands, featuring extraction of wh-subjects (*kdo* “who”)
  - c. Subject islands, featuring sub-extraction of *kakšen* “what kind of” or *čigav* “whose” out of subject NPs, in constructions with transitive verbs
  - d. Same as c., in constructions with unaccusative verbs.

The design of the present study was organized as a combination of two 2x2x2 subdesigns, one subdesign testing wh-islands and one testing subject islands.

Consider first the wh-island subdesign. Previous studies using the factorial design commonly targeted embedded *whether*-clauses to instantiate a wh-island configuration. Since in Slovenian there is no proper wh-counterpart of *whether* (its closest counterpart *ali* also means *or*), we used embedded *kdaj* “when”-clauses for this purpose. Sentences featuring extraction of wh-objects out of wh-islands were intended to serve as material for testing potential Subjacency-type effects. Sentences featuring extraction of wh-subjects were targeted with a purpose to investigate a potential ECP effect on top of the Subjacency effect (see Section 2.2. for discussion and references).

In this subdesign, we crossed factors LENGTH (matrix, embedded), STRUCTURE (island, non-island) and POSITION (subject, object). This resulted in 8 respective conditions. Examples of the constructed sets of target sentences are given below:

- (10) WH-ISLAND / OBJECT EXTRACTION
  - a. Kdo \_ je omenil, da je Maja brala revijo? [NON-ISLAND | MATRIX]  
 who AUX mention-PST that AUX Maja read-PST journal  
 ‘Who mentioned that Maja read a journal?’
  - b. Kaj je Jure omenil, da je Maja brala \_? [NON-ISLAND | EMBEDDED]  
 what AUX Jure mention-PST, that AUX Maja read-PST  
 ‘What did Jure mention that Maja read?’
  - c. Kdo \_ se je čudil, [kdaj je Maja brala revijo]? [ISLAND | MATRIX]  
 who REFL AUX wonder-PST when AUX Maja read-PST journal  
 ‘Who wondered when Maja read a journal?’
  - d. Kaj se je Jure čudil, [kdaj je Maja brala \_]? [ISLAND | EMBEDDED]  
 what REFL AUX Jure wonder-PST when AUX Maja read-PST  
 ‘What did Jure wonder when Maja read?’
- (11) WH-ISLAND / SUBJECT EXTRACTION
  - a. Kdo \_ je omenil, da je Tina pokosila travo? [NON-ISLAND | MATRIX]  
 who AUX mention-PST that AUX Tina mow-PST grass  
 ‘Who mentioned that Tina mowed the grass?’

- b. Kdo je Janez omenil, da je \_ pokosil travo? [NON-ISLAND | EMBEDDED]  
 who AUX Janez mention-PST that AUX mow-PST grass  
 ‘Who did Janez mention that mowed the grass?’
- c. Kdo \_ se je čudil, [kdaj je Tina pokosila travo] [ISLAND | MATRIX]  
 who REFL AUX wonder-PST when AUX Tina mow-PST grass  
 ‘Who wondered when Tina mowed the grass?’
- d. Kdo se je Janez čudil, [kdaj je \_ pokosil travo]? [ISLAND | EMBEDDED]  
 who REFL AUX Janez wonder-PST when AUX mow-PST grass  
 ‘Who did Janez wonder when mowed the grass?’

Note that the stimuli in conditions a. in both object extraction and subject extraction sentences have identical structural characteristics. It might therefore appear that our wh-island subdesign is slightly “defective” in that the POSITION factor does not fully distribute over the matrix level of the LENGTH factor. However, we do not see this as a methodological obstacle. The main consequence of this situation is simply that our constructed 2x2x2 model will not be fully informative without additional planned post-hoc comparisons (including estimating 2x2 sub-models as well as pairwise comparisons). But, as a matter of fact, in designs like ours, this was always going to be the case. In particular, the interaction term of the 2x2x2 model, showing whether the three factors at hand interact, merely indicates whether the two 2x2 interactions in question have different sizes. Since there is currently no linguistically-grounded theory of differences in sizes in the empirical domain that we are interested in, this interaction term would not be easily interpretable anyway. Instead, we will need to run individual 2x2 models, which is what the “defective” nature of the design requires too. Recall that, in the factorial design, potential island effects are interpreted as an interaction of two factors (see above). Given the focus of the present study, we might then anticipate that the most linguistically interesting inferences must indeed come from comparing smaller 2x2 models and perhaps some other pairwise comparisons, in addition to evaluating the 2x2x2 model as whole (see Section 6 for details; thanks to an anonymous reviewer for helpful suggestions on this point).

We now turn to the subject island subdesign. Here we crossed factors LENGTH (matrix, embedded), STRUCTURE (island, non-island) and verb TYPE (transitive, unaccusative). With regard to the island structure, following Polinsky et al. (2013), we constructed examples of subject islands which involve subextraction of *what-kind-of-* and *whose-* phrases as instances of left-branch extraction. In Slovenian, these wh-phrases are adjectival. Left-branch extraction of adjectival phrases is a somewhat limited phenomenon in Slovenian in that it does not occur as freely as in other Slavic languages such as Russian or Bosnian/Croatian/Serbian (see Bošković 2009; Franks 2014 and references therein for discussion).<sup>2</sup> Nevertheless, wh-interrogatives involving the kind of subextraction we are interested in here seem to be largely acceptable. Golden (1996: fn.5) reports examples like (11) as acceptable in colloquial speech, and our informants agree on this and similar examples. For the present purposes, we assume that left-branch extraction is available with *what-kind-of-* and *whose-* phrases in Slovenian.<sup>3</sup>

---

<sup>2</sup> As Bošković (2005) argues, even in languages that allow left-branch extraction, the presence of an additional adjective very often affects its availability.

<sup>3</sup> See also Stepanov et al. (to appear) for a follow-up study of extraction out of subject NPs in Slovenian in contexts that, crucially, do not involve left-branch extraction. The results of that study largely corroborate those reported here.

- (12) Čigave mi prinašaš [\_ pozdrave]?  
 whose me-DAT bring-2SG. greetings  
 ‘Whose greetings are you bringing me?’

The verb TYPE was manipulated with the aim to investigate the potential unaccusativity factor claimed to play a role in speakers’ acceptability of subject islands in Polinsky et al. (2013). In choosing the verbs for the respective sentences, we used the ability of unaccusative, but not transitive, verbs to form past participles as a diagnostic of unaccusativity in Slovenian (Marvin 2000). This is illustrated in (13) (Marvin’s examples (4b,c)):

- (13) a. Videl sem žensko, prispelo danes zjutraj [unaccusative]  
 see-PST 1SG woman-ACC arrive- PST.PTCP.ACC today morning  
 ‘I saw a woman who arrived this morning.’  
 b. \*Videl sem žensko, napisalo knjigo [transitive]  
 see-PST 1SG woman-ACC write-PST.PTCP.ACC book-ACC  
 ‘I saw a woman who wrote a book’

Examples of the constructed sets of target sentences are given below:

- (14) SUBJECT ISLAND / UNACCUSATIVE
- a. Kakšna gospa \_ je mislila, da so študenti prišli na žur?  
 what-kind-of lady AUX think-PST that AUX students come-PST on party  
 ‘What kind of lady thought that students came to the party?’
- b. Kakšni študenti je gospa mislila, da so \_ prišli na žur?  
 what-kind-of students AUX lady think-PST that AUX come-PST on party  
 ‘What kind of students did the lady think that came to the party?’
- c. Kakšna gospa \_ je mislila, da so [dobri študenti] prišli na žur?  
 what-kind-of lady AUX think-PST that AUX good students come-PST on party  
 ‘What kind of lady thought that good students came to the party?’
- d. Kakšni je gospa mislila, da so [\_ študenti] prišli na žur?  
 what-kind-of AUX lady think-PST that AUX students come-PST on party  
 \*’What kind of \_ did the lady think that students came to the party?’
- (15) SUBJECT ISLAND / TRANSITIVE
- a. Kakšna tajnica \_ je mislila, da je učitelj kaznoval dijake?  
 what-kind-of secretary AUX think-PST that AUX teacher punish-PST students  
 ‘What kind of secretary thought that a teacher punished the students?’
- b. Kakšen učitelj je tajnica mislila, da je \_ kaznoval dijake?  
 what-kind-of teacher AUX secretary think-PST that AUX punish-PST students  
 ‘What kind of teacher did the secretary think that punished the students?’
- c. Kakšna tajnica \_ je mislila, da je [strog učitelj] kaznoval dijake?  
 what-kind-of secretary AUX think-PST that AUX strict teacher punish-PST students  
 ‘What kind of secretary thought that a strict teacher punished the students?’
- d. Kakšen je tajnica mislila, da je [\_ učitelj] kaznoval dijake?  
 what-kind-of AUX secretary think-PST that AUX teacher punish-PST students  
 \*’What kind of \_ did the secretary thought that a teacher punished the students?’

We constructed four sets of target sentences for wh-island / object extraction (cf. (10)) and four sets for wh-island / subject extraction (cf. (11)). In addition, we constructed eight sets of

sentences for subject islands of which half were with unaccusative verbs as in (14), and half with transitive as in (15). There were 64 target sentences in total. These target sentences were divided into two equal size lists (32 targets in each). These two lists were equally distributed among the participants, so that each participant rated two tokens of each target condition. Each list of targets was then combined with 32 sentence fillers of various non-interrogative syntactic types, to provide some diversity to the testing materials and minimize chances of participants developing item-specific strategies in their ratings. Half of the fillers were grammatical, the other half ungrammatical, according to native speakers' judgment. The ungrammaticality of the fillers was due to violation of various grammatical constraints, such as subject-verb agreement, the clitic-second requirement of Slovenian, position of the auxiliaries, word order etc. The fillers were matched by length and plausibility to the target items. This resulted in two versions of the questionnaire, with 64 total items in each version. The order of items was pseudo-randomized for each list, and it was ensured that the first four items in each list are fillers.

### *5.3. An acceptability rating task*

The task was rating acceptability of sentences using the technique of magnitude estimation. Magnitude estimation is a method of subjective evaluation originally developed in psychophysics whereby a participant evaluates some gradable property (e.g. intensity of light) relative to some available standard, by assigning a numerical value on the basis of a subjective judgment (Stevens 1975), in relation to the numerical value assigned to the standard. The subjects are not restricted either in the range of numerical values that they are allowed to give (on the positive number scale) or in the granularity of the numerical scale adopted by each participant for the purposes of the experiment. We used a version of the magnitude estimation task adapted for judging acceptability of sentences (Bard et al. 1996). This task is well suited to the present study because of its capability to capture a potentially greater variability and range of acceptability ratings by using the unbounded positive number line.

The task began with a training session. The goal of the training session was to acquaint and familiarize the participants with the concept of magnitude estimation. During the training session, the subjects were offered to estimate the length of seven straight lines relative to the given line to which the numerical score 100 was assigned. Subjects were instructed to rely only on their subjective intuitions in evaluating the length of the lines; if the line seemed, for instance, twice as large as the standard, they were encouraged to give a value 200, and if it seemed about one third as large, then they would give a value 30. Both whole and decimal numbers could be used.

The training was followed by a sentence rating questionnaire. In our study, participants were presented with a reference sentence, or a standard, and a numeric value representing its acceptability. In our case, the sentence was (16), and it was assigned the value 100 (note that the number itself does not imply any particular acceptability status; this point was also stressed in the instructions). The participants were then instructed to indicate the acceptability of each of the subsequent sentences relative to the score assigned to the standard.

- (16) Proti kateremu pravilu je Klara mimogrede rekla, da je Cene protestiral?  
Against which rule AUX Klara in-passing say-PST that AUX Cene protest-PST  
'Against which rule did Klara say in passing that Cene protested?'

The study was conducted in the form of a paper survey. The standard remained visible throughout the entire procedure by being placed on top of each page of the questionnaire and separated by a line from the rest of the stimuli. Participants were under no time constraints to complete the task. On average, the surveys were completed within 20 minutes.

#### *5.4. Participants*

Sixty-four native speakers (thirty-nine females) of Slovenian, aged 18-54, participated in the experiment. Forty of those were high school students from the Srečko Kosovel School Center in Sežana, Slovenia. All participants had normal or corrected to normal vision and were naïve to the purposes of the study. The experiment in the school center was conducted over two visits, the rest of the participants completed the task individually.

The data from four participants were removed from consideration, two because of their indicated non-native Slovenian and two for the failed line test. This left the data from 60 subjects to be used for subsequent analyses.

#### *5.5. Statistical procedures*

Prior to analysis, the raw numerical ratings from each participant were z-score transformed. The z-score transformation converts each participant's ratings to a standardized score, in which each transformed rating represents the number of standard deviations by which the corresponding raw rating is different from that participant's mean rating. This kind of conversion eliminates potential scale biases between participants (such as choosing different ranges of values among participants or using one end of the scale), and therefore allows for a cleaner comparison of the participants' performance.

For the statistical analyses, we used linear mixed-effects models (Baayen et al. 2008), with LENGTH, STRUCTURE, POSITION as fixed factors in the wh-island subexperiment, and LENGTH, STRUCTURE and TYPE as fixed factors in the subject island subexperiment. For both series of models, subject and items were entered as random factors. Where possible, our constructed models utilized the maximal random effect structure with random intercepts for subjects and items, and random slopes for the fixed effect terms in subjects and items (Barr et al. 2013). Pairwise comparisons were performed as Tukey post-hoc estimations in our models. We report main effects as  $\chi^2$  and  $p$  values based on the likelihood-ratio test used in an ANOVA-like procedure which compares models including a fixed factor in question with models without that factor. Analyses were performed using the "lme4" package (Bates et al. 2015) in R (R Core Team 2014).

Finally, following a similar procedure in Sprouse et al. (2016) we computed DD scores for each participant, on the basis of which we then calculated mean DD scores for each island as a (non-standardized) effect-size measure for the island types under question.

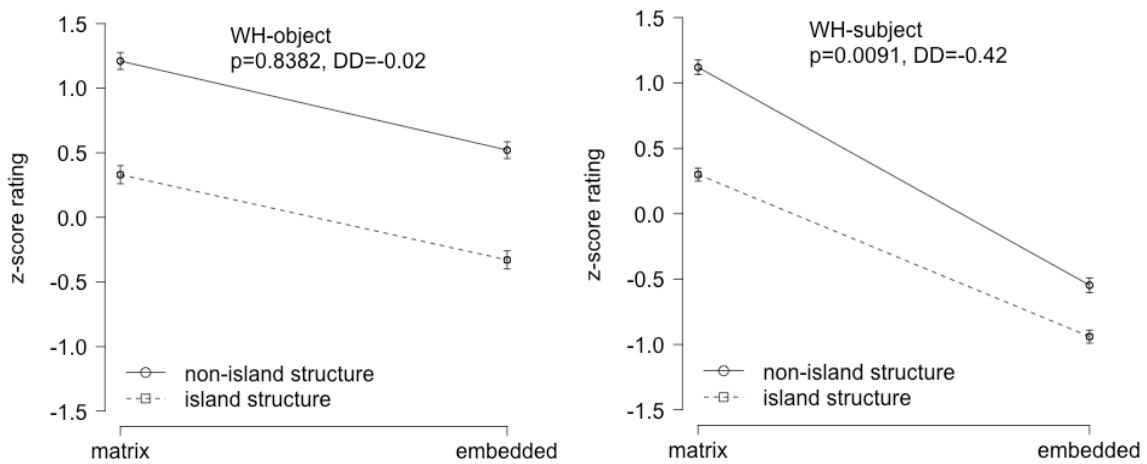
## **6. Results**

### *6.1. Wh-island sub-experiment*

We first evaluated the overall 2x2x2 model for the wh-island sub-experiment. Main effects were observed for each of the three factors involved, that is, LENGTH, STRUCTURE and POSITION. In other words, each of these factors emerged as a significant predictor of the acceptability scores. There was also a (marginal) interaction among these three factors. The results are summarized in Table 1 and Figure 2.

**Table 1.**  $\chi^2$ ,  $t$  and  $p$ -values for the three-way linear mixed-effects models in the wh-island sub-experiment

<i>FULL 2X2X2 MODEL</i>	$\chi^2$	$p$
main effect of LENGTH	52.804	<.0001
main effect of STRUCTURE	35.309	<.0001
main effect of POSITION	17.692	<.0001
interaction LENGTH x STRUCTURE x POSITION	3.848	<b>.0498</b>



**Figure 2.** An interaction plot of the 2x2x2 model in the wh-island sub-experiment.

Recall that the main question that we are interested in is whether the structures under consideration manifest true island effects, which, in the classical factorial design, emerge as a super-additive interaction of the LENGTH and STRUCTURE factors. In this respect, we considered two 2x2 models crossing these two factors and pertaining to extraction of wh-object and wh-subject, respectively. With respect to wh-island/object extraction, linear mixed-effects modeling revealed a main effect of LENGTH, as well as a main effect of STRUCTURE, but *no* interaction between these two factors.

With respect to wh-island/subject extraction, we again observed main effects of LENGTH and STRUCTURE. Furthermore, our models revealed that there is significant interaction between these two factors. However, instead of a super-additive effect that would be expected of a true island effect, in the case of wh-island/subject extraction we actually observed a *sub-additive* effect whereby the overall measure of degradation of the critical sentence is *less* than the sum of the degradations imposed by the length and island structure alone ( $DD < 0$ , cf. Figure 1).

We thus have a reason to believe that both length and island structure affect the acceptability score in the form of respective processing costs in the wh-island constructions. Following the method suggested in Sprouse et al. (2012), we identify these costs by computing the relevant pairwise comparisons: the length cost was identified with a pairwise comparison of NON-ISLAND | MATRIX and NON-ISLAND | EMBEDDED conditions, and the structure cost was identified with a pairwise comparison of NON-ISLAND | MATRIX and NON-ISLAND | EMBEDDED conditions. The cost effects of LENGTH and STRUCTURE were significant for both kinds of wh-island structure. The results of these planned comparisons are summarized in Table 2.

**Table 2.**  $\chi^2$ ,  $t$  and  $p$ -values for the two-way LENGTH x STRUCTURE linear mixed-effects models in the wh-island sub-experiment

	WH-OBJECT		WH-SUBJECT	
<i>2X2 MODELS</i>	$\chi^2$	$p$	$\chi^2$	$p$
main effect of LENGTH	32.03	<.0001	52.15	<.0001
main effect of STRUCTURE	39.08	<.0001	28.17	<.0001
interaction LENGTH x STRUCTURE	.04	<b>.8382</b>	6.80	<b>.0091</b>
<i>PAIRWISE COMPARISONS</i>	$t$		$t$	
LENGTH (STRUCTURE=non-island)	6.31	<.0001	14.89	<.0001
STRUCTURE (LENGTH=matrix)	7.94	<.0001	6.82	<.0001

We also constructed models that crossed factors LENGTH and POSITION, separately for island and non-island structures. Here again, we observed main effects of each of these factors. Furthermore, linear mixed-effects models revealed their significant interaction, both in the case of island sentences, as well as non-island sentences. This signals super-additive effects also in the case of these two factors. Furthermore, pairwise comparisons at the embedded conditions revealed significant differences in acceptability scores both in the case of island as well as non-island sentences. The results of these comparisons are summarized in Table 3.

**Table 3.**  $\chi^2$ ,  $t$  and  $p$ -values for the two-way LENGTH x POSITION linear mixed-effects models in the wh-island sub-experiment

	NON-ISLAND		ISLAND	
<i>2X2 MODELS</i>	$\chi^2$	$p$	$\chi^2$	$p$
main effect of LENGTH	54.713	<.0001	44.785	<.0001
main effect of POSITION	38.757	<.0001	23.592	<.0001
interaction LENGTH x POSITION	25.981	<b>&lt;.0001</b>	15.867	<b>&lt;.0001</b>
<i>PAIRWISE COMPARISONS</i>	$t$		$t$	
POSITION (LENGTH=embedded)	10.105	<b>&lt;.0001</b>	5.865	<b>&lt;.0001</b>

We do not report the results of the POSITIONxSTRUCTURE models here for reasons of space, and also because no particularly interesting and/or linguistically interpretable results could be obtained from these models in addition to the effects and interactions already reported above.

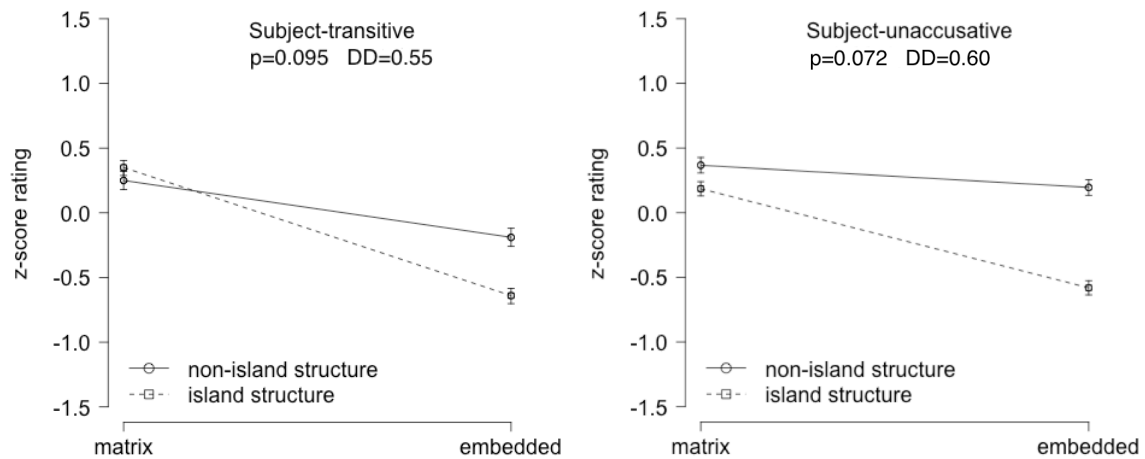
### 6.2. Subject island sub-experiment

Linear mixed models revealed main effects of LENGTH and STRUCTURE in the overall model of the subject island sub-experiment. There was, however, no verb TYPE main effect, and no interaction between the three factors. In other words, only the first two factors emerged as significant predictors for the acceptability scores. The results of the full model are shown in Table 4 and Figure 3.



**Table 4.**  $\chi^2$ ,  $t$  and  $p$ -values for the three-way linear mixed-effects models in the subject island sub-experiment

<i>FULL 2X2X2 MODEL</i>	$\chi^2$	$p$
main effect of LENGTH	11.92	.0005
main effect of STRUCTURE	3.98	.045
main effect of TYPE	.5001	.4794
interaction LENGTH x STRUCTURE x TYPE	0.01	.92



**Figure 3.** An interaction plot of the 2x2x2 model in the subject island sub-experiment.

Turning to the embedded 2x2 models, the relevant results are summarized in Table 4. In the subject island case, linear mixed models revealed main effects of LENGTH and STRUCTURE, as well as interaction between them, as a super-additive effect. However, in both unaccusative and transitive structures, as well as cumulatively for the subject island constructions, the interaction effect is marginal, as the corresponding  $p$  value falls slightly outside the usual .05 range. The corresponding DD scores are indicated on the graphs in Figure 2.

Again, the emergence of main effects of LENGTH and STRUCTURE suggests that these factors affect the acceptability score in the form of respective processing costs in the subject island construction. Similarly to the wh-island case, we asked if we could identify these costs by computing the relevant pairwise comparisons. It turned out that the cost effects of neither LENGTH nor STRUCTURE were significant for the subject island.

**Table 4.**  $\chi^2$ ,  $t$  and  $p$ -values for the two-way linear mixed-effects models the subject island sub-experiment

	SUBJECT-TRANSITIVE		SUBJECT-UNACCUSATIVE		SUBJECT-CUMULATIVE	
<i>2X2 MODELS</i>	$\chi^2$	$p$	$\chi^2$	$p$	$\chi^2$	$p$
main effect of LENGTH	11.16	0.003	7.39	0.025	13.903	.0009
main effect of STRUCTURE	4.15	0.125	7.83	0.019	6.53	.0368
interaction LENGTH x STRUCTURE	2.78	<b>0.095</b>	3.22	<b>0.072</b>	3.106	<b>.078</b>
<i>PAIRWISE COMPARISONS</i>	$t$		$t$		$t$	
LENGTH (STRUCTURE=non-island)	-1.14	0.67	0.32	0.98	-1.342	.5431
STRUCTURE (GAP=matrix)	0.27	0.99	0.40	0.97	-0.17	.9982

## 7. Discussion

### 7.1. Extraction of objects out of wh-islands

We begin the discussion of our results with wh-islands featuring extraction of objects. Golden (1995) and other works report such sentences as fully acceptable, concluding that a wh-island effect is somehow obviated in Slovenian (see (4) above). The results of our study allow for a more fine-grained qualification of this claim. The significant cost effect of the LENGTH and STRUCTURE factors suggest that Slovenian speakers are in fact sensitive to each of these factors in their evaluations. Each of these factors thus independently contributes to a lower acceptability of the relevant sentences involving extraction out of wh-islands.

At the same time, the absence of a super-additive effect in Slovenian wh-islands with object extraction, confirmed by all three metrics made available by the factorial design (no significant interaction between the LENGTH and STRUCTURE factors, the DD score around 0, parallel lines on the graph), strongly indicates that there is no true wh-island effect. In other words, the cumulative degradation in acceptability amounts to a sum of the degradations induced by the LENGTH and STRUCTURE factors independently, but not more than that. In this, our results support Golden's previous conclusion that there may be no wh-island effect in Slovenian, despite the fact that Slovenian speakers reported lower acceptability on wh-island sentences.

The case of object extraction out of wh-islands in Slovenian demonstrates, among other things, how the factorial design allows for a fine-grained distinction between sources of potential lowered acceptability of the island sentences. This distinction is particularly important in the context of the long-standing debate in the syntactic and sentence processing literature as to whether syntactic island effects may be reducible to independent processing factors that enter into a global evaluation of respective sentences, such as length of a dependency and the structural and/or semantic complexity of an island structure itself. For instance, some authors (e.g. Kluender and Kutas 1993; Hawkins 1999) suggested that processing constraints may be responsible for island effects in general. Viewed in the context of the factorial design with length and structure as chosen processing factors, this suggestion would entail that islands should manifest no-interaction effects similar to what we observed in the Slovenian case across-the-board. However, our results for Slovenian wh-islands contrast with Sprouse's et al. (2016) results obtained for wh-islands in English and Italian wh-questions where super-additive effects were properly manifested (see this work for details). The factorial design thus allows us to distinguish these two cases, attributing the ungrammaticality purely to processing factors in the Slovenian case, but not in the English or

Italian cases, even though on the surface the respective sentences may receive similar (reduced) acceptability scores.

### 7.2. *Extraction of subjects out of wh-islands*

Recall that the classical generative theories of locality of syntactic movement predict that extraction of wh-subjects (and of wh-adjuncts) out of island structures must lead to a greater degradation in acceptability than extraction of wh-objects. In syntactic terms, extraction of wh-objects out of islands amounts to a mild Subjacency violation, whereas extraction of wh-subjects and adjuncts amounts to violations of both Subjacency and the ECP, the former being a constraint on movement, the latter a constraint on licensing the trace(s) of movement. On the assumption that violating two constraints lead to a stronger ungrammaticality than violating of only one, the latter type of extraction should in general be judged much worse than the former (e.g. Chomsky 1986). Given these considerations, we wanted to see whether this markedly different kind of source of island-related unacceptability ratings could possibly reveal itself in the factorial design.

The observed main effects of LENGTH and STRUCTURE once again suggest that processing factors play a role in reducing the overall acceptability of the relevant structures. Individual costs of each of these processing factors were identified by running pairwise comparison tests (see Table 2). Thus the fact that Slovenian speakers generally tend to judge the wh-island sentences worse than their non-island counterparts runs against Golden's reported pattern of judgments regarding extraction of wh-subjects.

Furthermore, in contrast with the case of wh-island/object extraction, now we observe a significant interaction of the two fixed factors. However, the pattern of this interaction is sub-additive, rather than super-additive ( $DD < 0$ ). One can possibly argue that this sub-additive pattern is basically a floor effect observed on the embedded/island condition (the average z-score = -0.94, SE= 0.05, cf. Figure 1). In order to see whether this might be a potential explanation, we ran additional tests to compare the participants' performance on the embedded/island condition with their performance on the filler sentences (recall that half of these sentences were grammatical, and half ungrammatical). Among the subset of ungrammatical fillers, average z-scores ranged between -0.76 and -1.17 (standard errors 0.07 and 0.04, respectively). Importantly, 4 out of 16 (25%) of these filler sentences were evaluated by the participants significantly lower than the embedded/island condition ( $p < 0.05$  by pairwise t-tests). Per participant average z-scores on the set of ungrammatical fillers ranged between -0.35 and -1.34 (standard errors 0.06 and 0.05, respectively), with 40% of the participants showing average z-scores lower than -1.00 on these fillers. Across the entire set of filler sentences, the participants reported judgments corresponding to z-scores equal or lower than -1.00 in the total of 51% trials, with the lowest z-score reaching as low as -2.82. Furthermore, 23% of the total trials received a z-score equal or less than -1.25, a hypothetical threshold value for the embedded/island condition if the effect were completely additive, that is, in case  $DD=0$  (cf. Figure 1). We thus believe that a floor effect is an unlikely explanation of the observed sub-additive pattern on the embedded/island condition.

This state of affairs points to an additional systematic factor that affects the pattern of acceptability scores in the direction of their amelioration. Before we suggest what this additional factor might be, we will briefly review a similar pattern of results and their explanation in the previous literature.

Sprouse et al. (2011) investigate island effects in multiple wh-questions in English, using the same factorial approach. In particular, they cross factors STRUCTURE (island, non-island) and WH (single, double) resulting in four conditions exemplified in (17) for wh-islands (their Table 2):

- |      |  |                     |
|------|--|---------------------|
| (17) | a. Who __ thinks that John bought a car?     | NON-ISLAND   SINGLE |
|      | b. Who thinks that John bought what?         | NON-ISLAND   DOUBLE |
|      | c. Who __ wonders whether John bought a car? | ISLAND   SINGLE     |
|      | d. Who wonders whether John bought what?     | ISLAND   DOUBLE     |

The results of their acceptability rating study revealed a sub-additive interaction effect (alongwith the main effects of the two factors), whereby the difference in the score between the conditions (17a) and (17c) was larger than the difference in the score between the conditions (17b) and (17d). The authors refer to this as a *reverse island effect*. They offer an explanation of this as an additional processing effect, or rather, a combination of two additional processing strategies. First, the authors propose that parsing the wh-phrase in situ in multiple wh-questions (e.g. *what* in (17b,d)) requires a backward search for an antecedent, which in the case of English and other overt wh-movement languages is the fronted wh-phrase (*who* in (17)). According to the authors, this backward search for an antecedent is necessary for reasons such as online computation of scopal properties of the wh-phrase in situ. This search is in many ways similar to the forward search for a gap site that occurs in single wh-questions, known as the Active Filler strategy (Frazier and Clifton Jr 1989; see also Fodor 1978; De Vincenzi 1991). Furthermore, this kind of search requires reactivation of previously parsed material in the working memory in order to locate the wh-phrase in the sentence-initial position. Reactivation of the previously parsed material is a costly process processing-wise, and is likely to be subject to resource optimization and efficiency considerations. Second, the parser sometimes does not follow syntactic rules precisely as it builds syntactic structure online, but may commit various syntactic errors, as long as the resulting structure is “good enough” in some sense (Ferreira and Patson 2007; cf. also Townsend and Bever 2001). Taken together, these two considerations allow for a possibility that a backward search of the antecedent of *what* in (17d) may mistakenly end or be truncated not at the matrix clause-initial position hosting *who*, but earlier, at the embedded clause-initial position where the wh-phrase *whether* is located and which, at least feature- and scope-wise, may potentially serve as an antecedent for *what*. Such early closure or truncation of the search procedure might be justified by the need to minimize the amount of previously parsed material that needs to be reactivated and therefore reduces relevant processing costs. Under this analysis, the backward search across a single clause in (17d) proves less costly than the backward search across two clauses in (17b) in terms of length, therefore yielding an “illusory” grammaticality and a higher acceptability rating (a similar effect and an explanation was discussed in the case of adjunct islands in multiple wh-questions).

We believe a similar kind of explanation may be offered in the case of wh-subject extraction in Slovenian wh-island sentences. Consider (11) again, repeated here:

- (11) WH-ISLAND / SUBJECT EXTRACTION
- |    |  |                         |
|----|--|-------------------------|
| a. | Kdo _ je omenil, da je Tina pokosila travo?      | [NON-ISLAND   MATRIX]   |
|    | who AUX mention-PST that AUX Tina mow-PST grass  |                         |
|    | ‘Who mentioned that Tina mowed the grass?’       |                         |
| b. | Kdo je Janez omenil, da je _ pokosil travo?      | [NON-ISLAND   EMBEDDED] |
|    | who AUX Janez mention-PST that AUX mow-PST grass |                         |

- ‘Who did Janez mention that mowed the grass?’
- c. Kdo \_ se je čudil, [kdaj je Tina pokosila travo] [ISLAND | MATRIX]  
 who REFL AUX wonder-PST when AUX Tina mow-PST grass  
 ‘Who wondered when Tina mowed the grass?’
- d. Kdo se je Janez čudil, [kdaj je \_ pokosil travo]? [ISLAND | EMBEDDED]  
 who REFL AUX Janez wonder-PST when AUX mow-PST grass  
 ‘Who did Janez wonder when mowed the grass?’

As pointed out above, in the case of single *wh*-questions we are dealing with a forward search for a gap initiated by the moved *wh*-phrase. This search proceeds so that upon encountering the moved *wh*-phrase (“filler”) the parser temporarily stores it in the working memory and, upon subsequent incremental parsing, continuously probes potential gap sites as to their availability and suitability to successfully integrate the filler (e.g. Stowe 1986). Storing the filler in the working memory incurs an additional processing cost (Chen et al. 2005; Phillips et al. 2005, among others). Integration decisions are driven by various sources of information available to the parser, such as thematic information, phrase structure rules, plausibility etc. In *wh*-islands, specifically the cases of *wh*-object extraction, the STRUCTURE factor plays a role impeding integration of the filler and completion of an appropriate “filler-gap dependency” across the relevant structural boundary of a *wh*-island. Despite this obstacle, there is no choice for the parser other than to complete this association. The parser is in a way forced to complete it, which, as we saw above, incurs a special processing cost thereby (in addition to the LENGTH factor).

In the case of *wh*-subject extraction as in (11d), however, the situation is different. Slovenian is a *pro*-drop language that in principle allows null subjects. Note that the embedded *wh*-island structure *kdaj je \_ pokosil travo* /*when \_ mowed the lawn* is by itself a completely grammatical sentence in Slovenian meaning *when did he mow the lawn?* If the parser mistakenly takes the gap in the embedded clause as a silent *pro* subject, it no longer faces the task of associating the *wh*-filler (moved *wh*-phrase) with a gap inside the *wh*-island structure. The parser may choose this alternative *pro*-strategy to “save” the parse, driven by similar considerations of resource optimization and efficiency, while nevertheless adhering to the grammatical rules available in the language.

More specifically, we suggest that in this case a sort of reanalysis takes place whereby the parser actually ends up associating the moved *wh*-phrase with a silent *pro* subject. This kind of association is, in principle, allowed by the grammatical rules and many languages are known to exercise it as part of their grammars. Furthermore, as is known from syntactic literature, the structural conditions on associating the moved *wh*-phrase with a *pro* may substantially differ from conditions on associating the moved *wh*-phrase with its trace. In particular, the former may be exempt from syntactic islands, while the latter may not. This is true, for instance, in cases of *resumptive pronouns* that serve as a last resort grammatical device to save the structure from an island violation (18). In the absence of an island structure, resumptive pronouns are not licensed, as seen (19) (cf. Safir 1986):

- (18) a. \*This is the senator<sub>i</sub> that I don’t know what **t<sub>i</sub>** wrote  
 a. This is the senator<sub>i</sub> that I don’t know what **he<sub>i</sub>** wrote
- (19) a. That’s the senator<sub>i</sub> that I voted for **t<sub>i</sub>**  
 b. \*That’s the senator<sub>i</sub> that I voted for **him<sub>i</sub>**

This grammatically-licensed configuration, which allows for circumventing an island effect, is likely to be reflected also in the processing system.<sup>4</sup> We can interpret this to the effect that processing-wise, the factor STRUCTURE does not play the same role in *pro*-drop environments, as it does in non-*pro*-drop environments (viz. in wh-object extraction). It may thus be possible that the parser activates a rule similar to the one(s) implementing a resumptive-like strategy to associate the wh-phrase with the subject *pro* to circumvent additional processing costs related to the STRUCTURE factor. The factor LENGTH, of course, still incurs an independent processing cost. If this much is on the right track, we might be dealing here with another instance of a locally “good enough” structure which leads to an illusion of grammaticality and, consequently, to a higher acceptability rating.

### 7.3. *An alternative perspective on sub-additivity in wh-subject extraction*

A potential account of sub-additivity in the wh-subject island extraction outlined in the previous section is predicated on the underlying assumption that the participants’ rating on the embedded/island condition is higher than what would be expected if the LENGTH and STRUCTURE factors did not interact at all. There is an alternative way of looking at the apparent sub-additivity, suggested to us by an anonymous reviewer. This alternative view focuses not (only) on the embedded/island condition, but also on the embedded/non-island condition and its potential role in inducing this subadditivity. Consider this perspective in slightly more detail.

The logic of this alternative view is suggested by our results in Table 3 concerning two super-additive effects that emerged in our LENGTH x POSITION models for island as well as non-island sentences (see Table 3). For the moment, we put aside a super-additive effect for the island sentences, but we return to it in the end of this section. Concerning the super-additive effect in non-island structures, note the main effect of POSITION and a significant interaction between these two factors, suggesting that it matters for the speakers whether what is being extracted is a wh-subject or wh-object. The acceptability of embedded wh-subject extraction from non-islands is significantly lower than the extraction of embedded wh-objects. This subject-object asymmetry is also reinforced by respective pairwise comparisons. The reason why this effect is a priori surprising is that, on null hypothesis, subjects are structurally closer than objects, so the LENGTH effect by itself would entail that extraction of wh-subjects and of wh-objects would result either in equal acceptability rating (if the relevant length is calculated in clauses), or that extraction of wh-subject would be more acceptable than extraction of wh-object (if the relevant length is calculated in words). This suggests additional factor X that causes lowered acceptability in wh-subject sentences in Slovenian, irrespective of islandhood. If apparent sub-additivity is then taken into account, we may further speculate that factor X is somehow “turned off” in island sentences. This raises an obvious question about the possible origin of this factor. Several possibilities could be considered here, but for reasons of space we only mention one that seems most viable to us in this regard.

---

<sup>4</sup> This type of island-repairing resumption, which Sells (1984) refers to as “intrusive” resumption, should be distinguished from “true” resumption, not restricted to island contexts (see, e.g. McCloskey 2005). The former kind of resumption is likely to be at least partially processing-driven. In particular, the processing factor of length plays a role here (see also the discussion in the main text). As an anonymous reviewer correctly points out, examples like (i) are more acceptable than (19b):

(i) (?)That’s the senator<sub>i</sub> that you claim that Mary voted for him<sub>i</sub>

In contrast, the true resumption does not (obviously) depend on these processing considerations. See Sells (1984), Kroch (1981), Erteschik-Shir (1992), among others, concerning the “intrusive” and related types of resumption. Concerning true resumption, see Aoun et al. (2001), Shlonsky (1992), McCloskey (2005), Salzmann (2006), Hladnik (2015) among many others. Hladnik (2015) also discusses both types of resumptive strategy in Slovenian, although his discussion is confined mostly to relative clauses.

We are inclined to think that this factor is grammatical in nature. Specifically, we might be dealing here with a latent *that*-trace effect in Slovenian. *That*-trace (or, rather, comp-trace) effects with *wh*-subject extraction so far have not been reported in the literature on Slovenian syntax (e.g. Golden 1996), and this was taken into consideration when preparing the stimulus items for this study (cf. (11b)). However, it would not be surprising if such effects were ultimately discovered in Slovenian. A well-documented precedent in the literature concerns German, for which the standard, consensus view was that it is a language in which a classical *that*-trace effect does not obtain, in contrast to English (see, e.g. Bayer 1990, Haider 1993, Lutz 1996, among others). However, Featherston (2005), in a large-scale experimental study involving the magnitude estimation technique, demonstrated that *that*-trace effects actually obtain in German *wh*-extraction and topicalization structures whereby extraction takes place from verb-final sentences (German is an SOV language with verb-second in root clauses). The question why this particular effect emerges only in larger-scale studies but may remain elusive at the level of individual judgments is a non-trivial one. It requires a better understanding of linguistic principles behind the *that*-trace effect, on the one hand, and psychological mechanisms behind introspective judgments, on the other (see, among others, Bošković 2011; Chomsky and Lasnik 1977; Culicover 1993; Kandybowicz 2006; Pesetsky and Torrego 2001; on the former, and Chomsky 1965; Schütze 1996; Sprouse and Almeida 2013; on the latter). But, as far as we can see, postulating that a *that*-trace effect obtains in Slovenian does not contradict current knowledge of the Slovenian syntax, and so is a good potential candidate for factor X in the above sense.<sup>5</sup>

If this is indeed so, then the apparent sub-additivity may actually be due to the difference in acceptability between the respective island and non-island structures, whereby the former do not contain a *that*-trace configuration, but the latter does. This, of course, raises a further question as to how the sub-additive pattern may be affected if the *that*-trace factor is controlled for. More specifically, should we expect non-interaction of the LENGTH and STRUCTURE factors, as in the case of *wh*-object, or, rather, their interaction and a full super-additive effect, in other words, a true island effect? A follow-up study should elucidate this issue further.

There is also another super-additive effect identified in the LENGTH x POSITION models, pertaining to the island structures per se (see Table 3). Here, too, we observe a main effect of POSITION, as well as a significant interaction of these two factors, suggesting that it matters whether what is extracted out of a *wh*-island is a *wh*-object or *wh*-subject. In particular, embedded *wh*-subject extractions are evaluated significantly lower than embedded *wh*-object extractions. Again, some additional factor seems to be at play here, which affects subjects more than objects. Similarly to the case above, our hunch is that this factor, too, is grammatical and related to syntactic locality. Both island structures instantiate configurations relevant to Subjacency in the sense of Section 2.2., so Subjacency cannot be the factor distinguishing the *wh*-subject and *wh*-object extraction. However, the ECP, another locality

---

<sup>5</sup> A potential objection might be that Slovenian is a *pro*-drop language and the absence of *that*-trace effect seems to correlate with the *pro*-drop property, as was demonstrated, for instance, for Italian and Spanish (Rizzi 1990). However, *pro*-drop may be a necessary, but not a sufficient condition for voiding *that*-trace effects: in Romance, there is also free subject-verb inversion, which, in Rizzi's proposal, is crucial for licensing subject positions in *pro*-drop environments. Since there is currently no conclusive evidence that Slovenian features a Romance type subject-verb inversion, proposals like Rizzi's would actually predict that *that*-trace effects should obtain in this case (see also Szczegielniak 1999). More recent research suggests that there is a multitude of factors that may affect the presence or absence of *that*-trace effects in a language (see the references in the main text).

principle, distinguishes between subjects and objects in terms of licensing their movement traces. It might be, therefore, that in the case of island structures the observed super-additive effect is due to the ECP, in which case we would be led to conclude that this principle is fully operative in Slovenian and induces an effect predicted by the grammatical theory (see Section 2.2. for details). This possibility, too, must be further explored on a wider scale.

#### 7.4. Subject islands and processing costs

Golden (1995) and other works report examples of subextraction out of clausal subjects as acceptable (6). In our study, we found that similar examples involving subextraction out of NP subjects are less than acceptable. Furthermore, the observed super-additive effect ( $DD = 0.54$ ) points to a true island effect. However, the effect obtained in this study (cf. Table 4) has a somewhat marginal statistical significance so the issue of whether this is a true island effect is, therefore, a matter of interpretation. It is instructive to note in this regard that the marginality of the subject island effect (reported  $p=0.065$ ) was also a result of the Sprouse et al. (2016) factorial design study using materials in English and Italian, the languages where subject island effects are standardly believed to obtain, on the basis of informal acceptability judgments. Similarly to our study, their materials involved sub-extraction out of NP. Sprouse et al. (2016) interpret their result as significant. For the present purposes, we, too, opt to interpret our resulting effect as significant. This conclusion is reinforced by the results of the follow-up study reported in Stepanov et al (to appear) who report a significant subject island effect, using materials similar, though not identical, to those in the present study (see also fn. 3). Taken together, these considerations suggest that there is a subject island effect in Slovenian.

What is also interesting about our results is that neither the LENGTH nor STRUCTURE factor appears to incur independent processing costs affecting acceptability of the respective sentences. This suggests that the observed true island effect in subject islands in Slovenian is due to reasons other than these processing factors. Using similar materials Sprouse et al. (2012) observed a significant interaction/super-additive effect with subject islands in English. However, the authors reported in that study that the factor STRUCTURE did not incur an independent processing cost (as estimated by pairwise comparison tests). This is consistent with our results here (see Table 1). An explanation of this state of affairs might lie in the constructed syntactic complexity of the relevant noun phrases that also affects processing complexity. Indeed, the actual structures used to exemplify the NON-ISLAND and ISLAND values of the factor STRUCTURE had a shape such as *what* vs. *the speech about global warming*, respectively. The structure of the phrases *the gift* and *what* is of comparable syntactic complexity (presumably, each a DP with D selecting an N as a sister), so the only added complexity in the ISLAND condition comes from the prepositional phrase (PP) *about global warming*. Similarly, in the Slovenian materials that we used, the contrast between the two conditions lies in the presence of an adjective, e.g. *šudenti* / "students" vs. *dobri šudenti* / "good students" (cf. (14)). Perhaps this added syntactic complexity is not sufficient to incur a significant processing cost. This is different, for instance, from wh-islands, which all represent a clausal piece of structure, hence, presumably, are a priori more syntactically complex.

Our results also indicate that the factor LENGTH does not incur an independent processing cost in the subject island sentences. In the English study of Sprouse et al. (2012), LENGTH was found to incur such independent cost. The origin(s) of this emerging cross-linguistic contrast are less clear, given that this factor was controlled for in a manner similar to that in the English study. There is, however, one notable difference between our materials and those



used in the English study. Our materials involve D(iscourse)-linked *wh*-phrases, that is, (Slovenian counterparts of) *which*- and *whose*-phrases, whereas in the English study bare *wh*-words such as *what* and *who* are used. Since Pesetsky (1987), D-linked phrases are generally known to be subject to more liberal constraints on extraction than bare *wh*-words (though see Bošković 2008 for some potentially exceptional cases), which might also affect the length of respective dependencies. Processing-wise, it has been demonstrated that items that are richer in featural composition leave a longer and more robust trace in the working memory, and consequently are subject to a slower memory decay compared to items that have less relevant features (e.g. Hofmeister and Vasishth 2014). Thus a richer featural make-up potentially allows these items to linger in the memory for a longer time, overcoming potential effects of dependency length. In the case of *wh*-phrases, a phrase such as *which student* has a much larger set of relevant features to be encoded in the working memory (e.g. the semantic features of *student*, +*wh*, +animate, +referential, +singular, +presupposed, in Slovenian that would include also a gender feature value), than a bare *who* which presumably includes a rather limited set such as [+*wh*, +animate]. Thus a D-linked *wh*-phrase is predicted to linger in the working memory longer than a non-D-linked *wh*-phrase. The observed lack of the independent processing cost incurred by LENGTH could possibly be due to that.

Our results also suggest that, contrary to the claims in Polinsky et al. (2013), unaccusativity is not a relevant factor affecting the strength of the subject island effects. At least with respect to global acceptability judgments, Slovenian seems to pattern more with English than with Russian in being largely insensitive to this factor. We have not yet tested passives that, in theory, should pattern with unaccusatives with respect to their associated syntactic structure, and so there is still a logical possibility that passives would behave differently from the other verb types that we tested. However, Slovenian equivalents of English verbal passives generally have a rather marked character and tend to be avoided by the speakers, and this markedness could potentially confound a respective inquiry. A more feasible possibility would be to test *se*-constructions as equivalent to passives, and we leave that for future research.

### 7.5. Potential consequences for the theories of islands

On the basis of this study, Slovenian emerges as a language whose behavior with respect to island structures could largely be explained by the syntactic theories of islands based on Subjacency, CED, ECP and their minimalist descendants (see Section 2.2), but also raise some outstanding questions. Without going into detail, we will briefly sketch what we consider to be the points of major significance with respect to these theories.

First, Slovenian appears to be well behaved with respect to the subject island effect, which places this language on a par with many other languages including other Slavic languages (see Stepanov 2007; Polinsky et al. 2013; and the references therein).

Second, the fact that Slovenian manifests a subject island effect, taken together with the reported degradation of adjunct island sentences (see (7)), strongly suggests that Slovenian is well-behaved with respect to the CED in its original formulation (Huang 1982; see also Stepanov et al. to appear). Further experiments might be necessary to reinforce the significance of the subject island effect in other types of subject island constructions in Slovenian, such as sentential subject islands, in order for this claim to be sustainable. Nevertheless, we have reasonable grounds now to claim that, contrary to previous reports, there are no unexpected anomalies with regard to the uniform pattern of (un)acceptability in these two island configurations.

Third, in the context of Polinsky et al. (2013), the fact that subject island effects in Slovenian are not affected by the unaccusativity factor places Slovenian alongside with English, but leaves Russian as a puzzle in need of a more detailed explanation. Polinsky et al (2013) used Russian to argue against a specific type of accounts of subject islands that they referred to as a “generalized freezing” account. In this account, a subject island effect follows from a ban on subextraction from a moved phrase. In the standard clausal architecture, a subject originates within the VP and then moves to a higher functional domain (Spec-IP) before *wh*-movement. On the “generalized freezing” account, after the subject has moved, its internal structure is no longer available for further manipulation (such as subextraction). Russian is a language with a fairly flexible word order, with various linear combinations of verbal arguments possible. Capitalizing on this property, Polinsky et al. (2013) compared, in particular, subject island sentences in SV (preverbal subject) and VS (postverbal) orders. They found that the unaccusativity-driven pattern of acceptability is not affected by the word order in the target materials. On the assumption that at least one of these two constituent orders must be derived (by movement), they argue that the Russian data speak against a “generalized freezing” account.

Polinsky et al. (2013) further argue that their results are compatible with a derivationally-based account whereby a subject becomes an island, roughly, if it originates as an external argument of the predicate (e.g. Spec-vP. This is the case in unergative verbs like *jump* and transitives. In unaccusatives like *fall* or *arrive*, the surface subject is an underlying (direct) object, hence an internal argument, a sister of the verb. Therefore, on this account, unaccusative subject domains are not subject to island effects. But, as we saw, English and Slovenian present a problem for this otherwise elegant account. What we seem to face is some kind of parametric variation between English and Slovenian on the one hand, and Russian on the other. A source of the explanation for this variation, we suspect, lies in a more precise characterization of movement processes in the Russian clause. It might well be that the generalized freezing account is nevertheless correct, but, in addition, English and Slovenian require some further syntactic movement of the subject that Russian does not require. In that case unaccusativity will be the only factor that will regulate subextraction out of subject domains in Russian, but it will not be relevant in determining the islandhood of subjects in English and Slovenian, because all subjects would move regardless of their base position in the argument structure, and therefore become equally opaque for subextraction as moved domains. Of course, this requires additional clarification of the base position of subjects in the SV and VS orders which under this alternative perspective would both count as base-generated. One possibility that could be considered in this regard is that the subject always stays in-situ, i.e. in Spec-vP, whereas the verb can either undergo movement above vP, or stay in-situ, i.e. within vP (we thank an anonymous reviewer for suggesting this possibility). We believe this direction offers a promising way of reconciling an otherwise puzzling set of cross-linguistic facts in the three languages.

Slovenian also presents two other challenges for the theories of syntactic locality and islands. In contrast with subject islands, the absence of a *wh*-island effect in Slovenian emerges as a real phenomenon, despite the lowered acceptability data obtained in the course of our study. From the point of view of Subjacency-based theories, as well as more recent theories based on Relativized Minimality (Rizzi 1990; Chomsky 1995), this result is surprising. However, it is not unprecedented: apparent violations of *wh*-islands have been reported previously in Italian (Rizzi 1982), Bulgarian (Rudin 1988; Richards 1997), Swedish (Engdahl 1986), Spanish and Catalan (Torrego 1984), Danish (Erteschik-Shir 1973), Icelandic (Maling 1978),

Norwegian (Taraldsen 1982), Palauan (Georgopoulos 1991), Hebrew (Reinhart 1981, Preminger 2009), Romanian (Comorovski 1986) and a number of other languages. In a comprehensive evaluation of the patterns observed in Albanian, Bulgarian, Hebrew, Icelandic, Norwegian, Romanian and Swedish, Bošković (2008) argues that this group of languages manifests selective sensitivity to *wh*-islands, whereby the possibility of extraction out of an *wh*-island depends on a number of factors, including a) the type of the relevant extraction domain, viz. a main question or a relative clause, b) D-linking (whether the moving phrase is D-linked or not, and also whether or not the *wh*-island itself contains a D-linked phrase); as well as c) the thematic type of the element being extracted, viz. argument or adjunct. In addition, Bošković notes that the above languages allow extraction out of more than one *wh*-island, in contrast to languages like Italian, which do not (cf. Rizzi 1982). It is at present not clear if there exists a common explanatory denominator that predicts the amelioration effects for all *wh*-island violating languages. Factors that were suggested in the literature to play a role in this amelioration in different (groups of) languages include subject-verb inversion (Torrego 1984), a combination of subject-verb inversion and *Wh*-criterion (Rizzi 1996; Barbosa 1995), *pro*-drop (Uriagereka 1999a), (intermediate) adjunction to *Comp/Spec-CP* (Rudin 1988) and affixal articles (Bošković 2008). Slovenian thus offers a potentially interesting testing ground for further advancement of the Subjacency and Relativized Minimality-based theories of *wh*-island effects, with a hope to uncover further intricate aspects of the generative system that regulates the application of structural constraints on these configurations.

Our final comment concerns the ECP. This constraint on trace licensing played an important role in the pre-minimalist theories of locality (Chomsky 1981), but it is not quite clear how it should be restated in the current minimalist framework. Its proper minimalist formulation requires further understanding of its empirical aspects, and Slovenian seems to provide two interesting insights in this domain. If the reverse island hypothesis outlined in Section 7.2 is correct, it suggests that an empirical effect of the ECP, a strong grammatical constraint, can be modulated by processing considerations in the direction of amelioration of a global acceptability judgment, thus creating an “illusion” of grammaticality and causing a higher acceptability rating. We hypothesized that this ameliorating process may be mediated by an independent grammatical property of Slovenian, namely *pro*-drop. This state of affairs, therefore, presents an intriguing perspective for future grammatical as well as processing treatments of the ECP, suggesting that structural constraints imposed by the ECP could potentially be affected and/or overridden by the parser, on the grounds of “good enough” strategies and considerations of resource optimization and efficiency. If, on the other hand, our explanation of the super-additive interaction of LENGTH and POSITION within the island structures is on the right track, then Slovenian offers an interesting possibility to isolate an ECP effect irrespective of the Subjacency effect, the latter being absent in Slovenian, as evidenced from our results of *wh*-object extraction in *wh*-islands. Further experimental research on these and other theoretical possibilities may prove to provide valuable empirical support to our understanding of this fundamental principle.

## 8. Concluding remarks

In this work we were interested in elucidating patterns of grammaticality of sentences involving *wh*-islands and subject islands in Slovenian. In our empirical inquiry, we applied the factorial definition of syntactic island recently developed in the literature, which allows separating a potential true island effect from the influence of two processing factors that

might independently affect speakers' acceptability rating on the island sentences, namely, the length of the resulting dependency, and the presence of an island structure itself. Using this fine-grained technique, we found that, contrary to claims in the previous literature, Slovenian speakers manifest a true island effect in sentences featuring subextraction out of subject islands. Furthermore, we found that there is no true wh-island effect in Slovenian, even though Slovenian speakers judge corresponding sentences lower than their non-island containing counterparts. This lowered acceptability is thus likely to be due to the two processing factors indicated above. In addition, we discovered a "reverse island effect" in wh-island sentences involving extraction of wh-subjects, and suggested a processing explanation of it in terms of considerations of parsing efficiency and resource optimization. Overall, our results place Slovenian on a par with languages that seem to circumvent the wh-island constraint, similarly to languages like Italian, Bulgarian, Spanish and Catalan. Therefore, similarly to these cases, one may legitimately ask why the wh-island is circumvented in Slovenian, thereby setting the ground for further theoretical work in this domain. Another direction of further research is the potential interaction of the grammatical constraints such as ECP with processing strategies that in our case led to a reverse island effect; that, too, needs further analysis and understanding. In addition to these theoretical issues, our study can be seen as a methodological test especially targeting languages that present a priori challenges for the current theories of syntactic locality.

### Acknowledgments

We are grateful to the participants of the 8<sup>th</sup> Conference on Syntax, Phonology and Language Analysis (SinFonJA 8) at the University of Ljubljana, and to two *Linguistics* reviewers for their very helpful comments concerning both the theoretical and the experimental parts of this work. We also thank Dušan Štolfa and Marija Vidmar at the Srečko Kosovel School Centre in Sežana, Slovenia, for help in recruiting the participants for this study. This work was partially supported by the EU 7<sup>th</sup> Framework project "Advancing the European Multilingual Experience" (ATHEME, Grant Agreement 613465), and by the Slovenian Research Agency (ARRS) program No. P6-0382.

### References

- Aoun, Joseph, Lina Choueiri & Norbert Hornstein. 2001. Resumption, movement, and derivational economy. *Linguistic Inquiry* 32. 371–403.
- Baayen, R Harald, Douglas J. Davidson, & Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390–412.
- Barbosa, Pilar. 1995. *Null subjects*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Bard, Ellen Gurman, Dan Robertson & Antonella Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72. 32–68.
- 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. 255–278.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67. 1–48.
- Bianchi, Valentina & Cristiano Chesi. 2014. Subject Islands, Reconstruction, and the Flow of the Computation. *Linguistic Inquiry* 45. 525–569.

- Bošković, Željko. 2009. The NP/DP analysis and Slovenian. In Ljiljana Subotić (ed.), *Proceedings of the University of Novi Sad Workshop on Generative Syntax*, 53–73. Novi Sad: Filozofski fakultet, University of Novi Sad.
- Chen, Evan, Edward Gibson & Florian Wolf. 2005. Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language* 52. 144–169.
- Chomsky, Noam. 1964. *Current issues in linguistic theory*. The Hague: Mouton.
- Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson and Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 232–286. New York: Holt, Rinehart & Winston.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT press.
- De Vincenzi, Marica. 1991. *Syntactic parsing strategies in Italian: The minimal chain principle*. Dordrecht: Springer.
- Featherston, Sam. 2005. That-trace in German. *Lingua* 115. 1277–1302.
- Ferreira, Fernanda & Nikole D. Patson. 2007. The 'good enough' approach to language comprehension. *Language and Linguistics Compass* 1. 71–83.
- Fodor, Janet D. 1978. Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9. 427–473.
- Franks, Steven. 2014. The Slovenian orphan accusative, empty pronouns and noun phrase structure. In Lilia Schurcks, Anastasia Giannakidou & Urtzi Etxeberria (eds.), *The nominal structure in Slavic and beyond*, 129–182. Berlin: De Gruyter.
- Frazier, Lyn & Charles Clifton Jr. 1989. Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4. 93–126.
- Golden, Marija. 1995. Interrogative wh-movement in Slovene and English. *Acta Analytica* 14. 145–187.
- Golden, Marija. 1996. K-premik in skladijski otoki v slovenski skladnji. *Razprave SAZU, Razred II*, 15. 237–253.
- Hawkins, John A. 1999. Processing complexity and filler-gap dependencies across grammars. *Language* 75. 244–285.
- Hofmeister, Philip & Shravan Vasishth. 2014. Distinctiveness and encoding effects in online sentence comprehension. *Frontiers in Psychology* 5. doi:10.3389/fpsyg.2014.01237.
- Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Jurka, Johannes. 2010. *The importance of being a complement: CED effects revisited*. College Park, MD: University of Maryland dissertation.
- Kluender, Robert & Marta Kutas. 1993. Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8. 573–633.
- Krivochen, Diego Gabriel & Peter Kosta. 2013. *Eliminating empty categories*. Frankfurt am Main: Peter Lang.
- Lasnik, Howard & Mamoru Saito. 1992. *Move a: Conditions on its application and output*. Cambridge, MA: MIT Press.
- Marvin, Tatjana. 2000. Past participles in reduced relatives. In Marjo van Koppen, Erica Thrift, Erik Jan van der Torre & Malte Zimmermann (eds.), *Proceedings of ConSOLE IX*, 141–156. Lund: University of Lund.
- Phillips, Colin, Nina Kazanina & Shani H. Abada. 2005. ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22. 407–428.
- Polinsky, Maria, Carlos G. Gallo, Peter Graff, Ekaterina Kravtchenko, Adam Milton Morgan & Anne Sturgeon. 2013. Subject islands are different. In Jon Sprouse (ed.),

- Experimental syntax and island effects*, 286–309. Cambridge: Cambridge University Press.
- R Core Team. 2014. R: A Language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. <http://www.R-project.org/>.
- Richards, Norvin. 1997. *What moves where when in which language*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Rizzi, Luigi. 1982. Violations of the wh-island constraint and the subjacency condition. In Luigi Rizzi (ed.), *Issues in Italian syntax*, 49–76. Dordrecht: Foris.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
- Rizzi, Luigi. 1996. Residual verb second and the Wh-criterion. In Adriana Belletti & Luigi Rizzi (eds.), *Parameters and Functional Heads*, 63–90. Oxford: Oxford University Press.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Rudin, Catherine. 1988. On multiple questions and multiple *wh*-fronting. *Natural Language and Linguistic Theory* 6. 445–501.
- Sprouse, Jon, Shin Fukuda, Hajime Ono & Robert Kluender. 2011. Reverse island effects and the backward search for a licenser in multiple *wh*-questions. *Syntax* 14. 179–203.
- Sprouse, Jon, Matt Wagers & Colin Phillips. 2012. A test of the relation between working-memory capacity and syntactic island effects. *Language* 88. 82–123.
- Sprouse, Jon, Ivano Caponigro, Ciro Greco & Carlo Cecchetto. 2016. Experimental syntax and the variation of island effects in English and Italian. *Natural Language and Linguistic Theory* 34. 307–344.
- Stepanov, Arthur. 2001. *Cyclic domains in syntactic theory*. Storrs, CT: University of Connecticut dissertation.
- Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10. 80–126.
- Stevens, S.S. 1975. *Psychophysics: Introduction to its perceptual, neural and social prospects*. New York: John Wiley.
- Stowe, Laurie A. 1986. Parsing *wh*-constructions: Evidence for on-line gap location. *Language and Cognitive Processes* 1. 227–245.
- Takahashi, Daiko. 1994. *Minimality of movement*. Storrs, CT: University of Connecticut dissertation.
- Torrego, Ester. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15. 103–129.
- Townsend, David J. & Thomas G. Bever. 2001. *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.
- Truswell, Robert. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117. 1355–1377.
- Uriagereka, Juan. 1999. Multiple Spell-Out. In Samuel D. Epstein and Norbert Hornstein (eds.), *Working minimalism*, 251–282. Cambridge, MA: MIT Press.
- Wexler, Kenneth & Peter Culicover. 1981. *Formal principles of language acquisition*. Cambridge, MA: MIT Press.