Resumption and locality: a crosslinguistic experimental study

Theodora Alexopoulou and Frank Keller
University of Edinburgh

1 Introduction

We present two magnitude estimation studies that seek to establish the interaction between island violations, resumption and embedding in English and Greek. In this section we introduce the phenomena in question and in Sections 2 and 3 we present the experimental results.

1.1 Strong and weak islands

Sensitivity to island violations has long been recognized as a property of movement or gap constructions. Thus, unlike *that*-clauses in (1-a), indirect questions in (1-b) and complex NPs containing a relative clause in (1-c) are islands for movement. Indirect questions are viewed as "weak" islands since examples like (1-b) are judged better than examples like (1-c) which involve "strong" islands.

(1) a. Who does John think Mary will choose?
     b. ?*Who did Mary wonder whether they will fire?
     c. *Who did John meet the girl who will marry?

Strong islands constrain movement crosslinguistically while languages vary with respect to the status of weak islands. For example, various authors take the focus and *wh*-movement out of the indirect questions in (2) to be grammatical in Greek (Tsimpli, 1995; Alexopoulou, 1999).

(2) a. dhen ksero   sti   Maria pios   milise
         not know-1sg to-the Maria who-nom talked-3sg
         'I don't know who talked to Maria.'
     b. pion   anarotithikes an   tha ton   apolisoune
         who-acc wondered-2sg whether/if will him-acc fire-3pl
         'Who did you wonder whether they will fire?'

The first aim of this study is to establish the nature of these violations and provide a concrete empirical basis for terms like "weak" and "strong" by quantifying the effect of such violations. Furthermore, we compare the status of such violations in English and Greek with emphasis on the status of weak islands which have been argued to be different in the two languages.

1.2 Islands and resumptive pronouns

Resumption is often viewed as a "last resort" device to "save" island violations, at least in cases like (3) (from Haegeman 1991), which involve weak islands.

(3)   a.   This is the man whom Emsworth told me when he will invite him.
        b.   This is the man whom Emsworth made the claim that he will invite him.

The second aim of this paper is to investigate the nature of the interaction between resumption and islands, establish the "saving" effect of resumptives and investigate these interactions from a crosslinguistic perspective. In particular, we compare English and Greek. Both languages resist resumptives in ordinary *wh*-questions (4).

(4)   a.   Who did you fire (*him)?
        b.   pion   (*ton) apelises
            who-acc (*him) fire-2sg

However, unlike English, which has no productive resumptive strategy, Greek allows resumption in a number of constructions such as Clitic Left Dislocation in (5-a) and Null Operator Structures (see (5-b) from Tsimpli 1999). The crosslinguistic comparison therefore allows us to investigate the effect the availability of an independent resumptive strategy might have on the interaction between resumption and islands.

(5)   a.   ta klidia ta   stilame sti   Maria
            the keys them sent   to-the Maria
            'We sent the keys to Maria.'
        b.   i   Maria ine omorfi na   *(tin)   kitas
            the Maria is   pretty subj her-acc look-at-2sg
            'Maria is pretty to look at.'

1.3 Resumptives and embedding

Resumption has also been argued to interact with the "distance" or depth of embedding separating the extracted phrase from the extraction site. For example, Erteschick-Shir (1992) argues that a resumptive pronoun becomes more acceptable as the extraction site becomes more deeply embedded, a claim that she illustrates with the examples in (6).

(6) a. This is the girl that John likes 0/*her.
     b. This is the girl that Peter said that John likes 0/??her.
     c. This is the girl that Peter said that John thinks that Bob likes 0/?her.
     d. This is the girl that Peter said that John thinks that yesterday his mother had given some cakes to ?0/her.

Similarly, it has been argued for Greek that, unlike (4-b), a resumptive is acceptable when embedded at least one *that*-clause away from the matrix as shown in (7) (Tsimpli, 1999).

(7) pion   ipoptefthike i   Maria oti tha ton   kalesoume
     who-acc suspect-3sg the Maria that will him-acc invite-1pl
     'Who did Maria suspect we will invite?'

Note that in (6) and (7) there is no grammatical violation and therefore no grammar principle could explain the availability of resumptives here.[1]

The final aim of this study is again to establish the validity of these observations and compare interactions between resumption and islands with interactions between resumption and ordinary *that*-clause embedding. The overall aim of this comparison is to provide a concrete empirical basis for identifying the source of locality restrictions on extraction that, on the basis of the observations presented in this section, appear as the result of both grammatical and processing restrictions.

In the following sections we report results of two magnitude estimation studies in English and Greek. We use object extraction in *wh*-questions to investigate the interaction between resumption, islands and embedding. Most discussions in the literature relate to relative clauses rather than *wh*-questions. However, relative clauses are ambiguous between a restrictive and non-restrictive interpretation and, under the latter, they do not obey islands. This ambiguity could have introduced a confounding factor and so *wh*-questions were adopted instead of relatives. It should be pointed out though that nothing in the literature suggests that the effect of resumptives in islands is restricted to relative clauses.

2   Experiment 1: resumptives in English

The first experiment investigated how embedding and island constraints interact with resumption in English. Four different types of islands were used: complement clause without *that* (no island), complement clause with *that* (no island), complement clause with *whether* (weak island), relative clause (strong island). Two levels of embedding were tested: single embedding (one complement clause or relative clause) and double embedding (one complement clause embedding another complement clause or a relative clause).To have a standard of comparison, we also in-

---

[1]Indeed Erteschick-Shir (1992) and Dickey (1996) attribute such effects to memory constraints on the human parser. Due to space limitations we cannot discuss their proposals.

cluded sentences without embedding (control condition, zero embedding). Example sentences are given in (8)–(10).

(8) **No island violation**
    a.   **Zero embeddings:** Who will we fire 0/him?
    b.   **One embedding:**   Who does Mary claim that we will fire 0/him?
    c.   **Two embeddings:** Who does Jane think that Mary claims that we will fire 0/him?

(9) **Weak island violation**
    a.   **One embedding:**   Who does Mary wonder whether we will fire 0/him?
    b.   **Two embeddings:** Who does Jane think that Mary wonders whether we will fire 0/him?

(10) **Strong island violation**
    a.   **One embedding:**   Who does Mary meet the people that will fire 0/him?
    b.   **Two embeddings:** Who does Jane think that Mary meets the people that will fire 0/him?

2.1 Predictions

Our hypotheses are based on the theoretical literature where the central claim is that resumptive pronouns in English can save weak island violations.

(a) In the no island condition, there is no grammatical violation and therefore gaps should be equally acceptable at all levels of embedding and better than resumptives which are also not expected to interact with embedding.

(b) In the weak island condition, resumptives are expected to save the island violation and, hence, be more acceptable than gaps. At the same time, resumptives are expected to be more acceptable in single and double embedding than in the control condition, while the opposite should hold for gaps (as they incur a weak island violation).

(c) In the strong island condition, resumptives cannot save the island violation. Hence we predict gaps and resumptives to be both unacceptable. Also, gaps should be less acceptable in single and double embedding than in the control condition (as they incur a strong island violation).

2.2 Method

2.2.1 Subjects

Thirty-six subjects were recruited over the Internet by postings to newsgroups and mailing lists. All subjects were self-reported native speakers of English. Linguists and students of linguistics were excluded from the sample.

2.2.2 Materials

The design crossed the following factors: *Embedding* (single or double embedding), *Island* (complement clause without *that*, complement clause with *that*, complement clause with *whether*, relative clause), and *Resumption* (gap or resumptive). This resulted in *Embedding* $\times$ *Island* $\times$ *Resumption* $= 2 \times 4 \times 2 = 16$ cells. As controls, we included stimuli without embedding (gap or resumptive), resulting in a total of 18 cells. Nine lexicalizations were used for each cell, yielding a total of 162 stimuli.

The stimulus set was divided into nine subsets of 18 stimuli by placing the items in a Latin square. A set of 18 fillers was used, covering the whole acceptability range.

2.2.3 Procedure

The method used was magnitude estimation (ME) as proposed by by Bard et al. (1996) and Cowart (1997).

Subjects first saw a set of instructions that explained the concept of numerical ME using line length. Subjects were instructed to make length estimates relative to the first line they would see, the reference line. They were told to give the reference line an arbitrary number, and then assign a number to each following line so that it represented how long the line was in proportion to the reference line. Several example lines and corresponding numerical estimates were provided to illustrate the concept of proportionality. Then subjects were told that linguistic acceptability could be judged in the same way as line length. Examples of sentences of varying acceptability were used to illustrate the task.

After reading the instructions, subjects took part in a training phase designed to familiarize them with the task. In the training phase, subjects were ask to use ME to judge the length of a set of lines. Then, a set of practice items (similar to the experimental items) were administered to familiarize subjects with applying ME to linguistic stimuli. Finally, subjects had to judge the experimental items. Each subject judged one set of 18 experimental stimuli and all 18 fillers, i.e., a total of 36 items.

Keller and Alexopoulou (2001) present a detailed discussion of the safeguards that WebExp puts in place to ensure the authenticity and validity of the data collected, and also present a validation study comparing web-based and lab-based judgment data (for a WebExp validation study using sentence completion data, see Corley and Scheepers 2002).

2.3 Results

The data were normalized by dividing each numeric judgment by the modulus value that the subject had assigned to the reference sentence. This operation creates a common scale for all subjects. Then the data were transformed by taking the decadic logarithm. This transformation ensures that the judgments are normally distributed and is standard practice for ME data (Bard et al., 1996). All analyses and figures are based on normalized, log-transformed judgments. Figures 1–4 graph the mean judgments for all four island conditions.

Extraction from no island (no-that-clause)
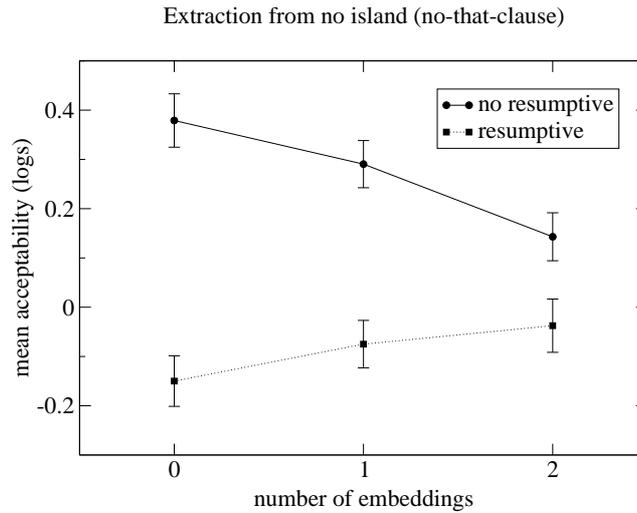
Figure 1: Effect of embedding and resumption on extraction in English (no-*that*-clause condition)

An ANOVA yielded significant main effects of *Embedding* ($F_1(1,35) = 6.867$, $p = .013$; $F_2(1,8) = 13.291$, $p = .007$), *Island* ($F_1(3,105) = 30.217$, $p < .0005$; $F_2(3,24) = 36.393$, $p < .0005$), and *Resumption* ($F_1(1,35) = 31.323$, $p < .0005$; $F_2(1,8) = 26.069$, $p = .001$). We are mainly interested in an interaction of *Island* and *Resumption*, as this indicates that the acceptability of resumptives is sensitive to island violations. This interaction was significant ($F_1(3,105) = 13.792$, $p < .0005$; $F_2(3,24) = 7.857$, $p = .001$). All other interactions were only significant by subjects: *Island/Embedding* ($F_1(3,105) = 2.923$, $p = .037$; $F_2(3,24) = 1.066$, $p = .382$), *Embedding/Resumption* ($F_1(1,35) = 7.197$, $p = .011$; $F_2(1,8) = 3.038$, $p = .120$), and *Island/Embedding/Resumption* ($F_1(3,105) = 3.550$, $p = .017$; $F_2(3,24) = 1.631$, $p = .208$).

We conducted a Tukey post-hoc test to further investigate the interaction of *Island* and *Resumption*. This test allows us to determine in which of the island conditions a gap is more acceptable than a resumptive. We found that in both no island conditions (complement clause with or without *that*), the gap was more accept-
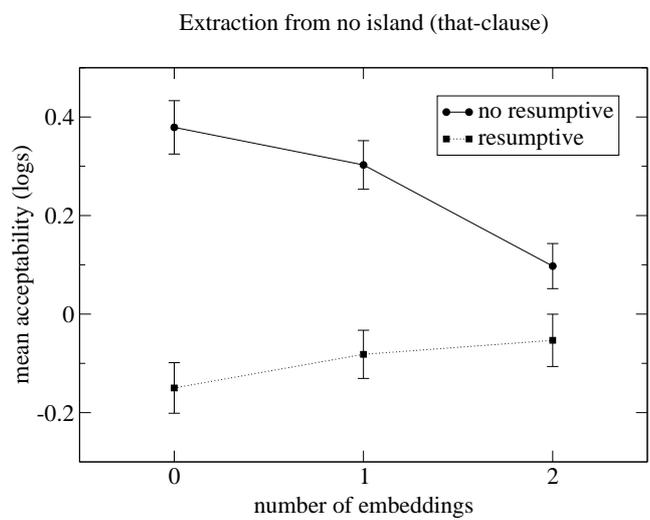
Extraction from no island (that-clause)

Figure 2: Effect of embedding and resumption on extraction in English (*that*-clause condition)

Extraction from weak island (whether-clause)
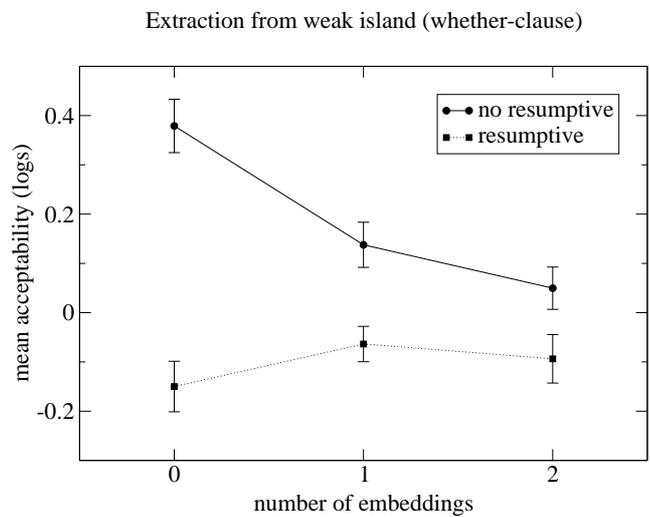
Figure 3: Effect of embedding and resumption on extraction in English (*whether*-clause condition)
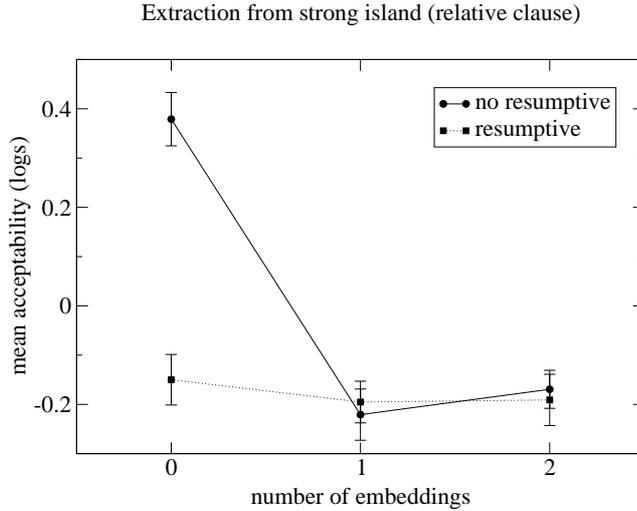
Extraction from strong island (relative clause)

Figure 4: Effect of embedding and resumption on extraction in English (relative clause condition)

able than the resumptive ($\alpha < .01$ in both cases). Also in the weak island condition (complement clause with *whether*), the gap was more acceptable than the resumptive ($\alpha < .01$). In the strong island condition, both the gap and the resumptive were equally unacceptable.

As a next step, we compared the conditions with single and double embedding to the control (no embedding). The appropriate statistic is Dunnett's test for comparing multiple conditions to a control condition. We will first report the results of comparing the gapped stimuli to the gapped control condition. For both no island conditions, there was no significant difference between control and the single embedding condition, while the double embedding was significantly less acceptable than the control, by subjects only ($t_{d_1}(35,9) = 3.016, p < .01; t_{d_2}(8,9) = 2.794, p > .05$ and $t_{d_1}(35,9) = 4.860, p < .01; t_{d_2}(8,9) = 2.681, p > .05$). In the weak island condition, we found that both the single and the double embedding condition were less acceptable than the control, by subjects only ($t_{d_1}(35,9) = 4.278, p < .01; t_{d_2}(8,9) = 2.500, p > .05$ and $t_{d_1}(35,9) = 5.429, p < .01; t_{d_2}(8,9) = 3.007, p > .05$). Also in the relative clause condition, singly and doubly embedded stimuli were less acceptable than the control ($t_{d_1}(35,9) = 7.767, p < .01; t_{d_2}(8,9) = 7.905, p < .01$ and $t_{d_1}(35,9) = 8.766, p < .01; t_{d_2}(8,9) = 8.362, p < .01$).

In a separate test, we compared the resumptive stimuli to the resumptive control condition. We failed to find significant differences between the single or double embedded stimuli and the control (this holds for both no island conditions, the weak island condition, and the strong island condition).

2.4 Discussion

Our results appear to partially confirm and partially disprove the predictions adopted earlier.

Prediction (a). No interaction between embedding and the acceptability of gaps was expected in the no-island condition. As a consequence, resumptives were predicted to be less acceptable than gaps for all levels of embedding. These predictions were largely confirmed by the experimental results (see Figures 1 and 2). Firstly, gaps were significantly more acceptable than resumptives. Secondly, for resumptives, the single and double embedding conditions were not significantly different from the control. For gaps, the single embedding condition was not different from the control, but the double embedding condition was less acceptable than the control. This is an unexpected result since there is no obvious grammatical violation here.

All these results hold for both no island conditions; there does not seem to be a difference between complement clauses with *that* and ones without *that*.

Prediction (b). For weak islands, we found that singly and doubly embedded gaps were less acceptable than the control condition, as predicted. However, resumptives fail to "save" or improve a weak island violations, since gaps are always significantly more acceptable than resumptives (see Figure 3). Furthermore, there was no difference between the singly or double embedded resumptives and the control.

Prediction (c). In the strong island condition, resumptives were not expected to save the island violation. Both resumptives and gaps were predicted to be unacceptable. Also, gaps were predicted to be less acceptable than the control in both the single and double embedding condition, due to the strong island violation they incur. These predictions were born out (see Figure 4). We found that there was no significant difference between gaps and resumptives; both were highly unacceptable. At the same time, the comparison with the control showed that both singly and double embedded gaps were less acceptable than the control. There was no significant difference between singly and double embedded resumptives and the resumptive control.

3 Experiment 2: resumptives in Greek

The purpose of the present experiment was to test the crosslinguistic validity of our findings from English. We tested Greek, which, as mentioned earlier, differs from English in the following ways: (i) indirect questions in Greek are not considered islands; (ii) a resumptive is available in *wh*-questions when embedded at least one *that*-clause away from the matrix (Tsimpli, 1999); (iii) unlike English, Greek exhibits productive resumption in Clitic Left Dislocation and Null Operator Structures.

3.1 Predictions

Based on the theoretical literature on resumptives and extraction in Greek, we arrived at the following predictions:

(a) As in English, gaps should be as acceptable in the single and double embedding as in the control condition. Resumptives should be more acceptable in single and double embedding than in the control (where they are predicted to be unacceptable). Furthermore, resumptives should be as acceptable as gaps in the single and double embedding condition.

(b) Since indirect questions are not considered islands in Greek, the acceptability of gaps and resumptives in this condition should be the same as in the no island one.

(c) Greek respects strong islands, hence we expect the same picture as in English here. Both gaps and resumptives should be unacceptable and gaps should less acceptable than the control, both for single and double embedding.

3.2 Method

3.2.1 Subjects

Fifty-nine subjects were recruited over the Internet by postings to newsgroups and mailing lists. All subjects were self-reported native speakers of Greek. Linguists and students of linguistics were excluded from the sample.

3.2.2 Materials

The design mirrored the one for English, but left out the no-*that*-clause condition (the complementizer is obligatory in Greek). This resulted in *Embedding* $\times$ *Island* $\times$ *Resumption* $= 2 \times 3 \times 2 = 12$ cells. As controls, we included stimuli without embedding (gap or resumptive), resulting in a total of 14 cells. Seven lexicalizations were used for each cell, yielding a total of 98 stimuli.

The stimulus set was divided into seven subsets of 14 stimuli by placing the items in a Latin square. A set of 14 fillers was used, covering the whole acceptability range.

3.2.3 Procedure

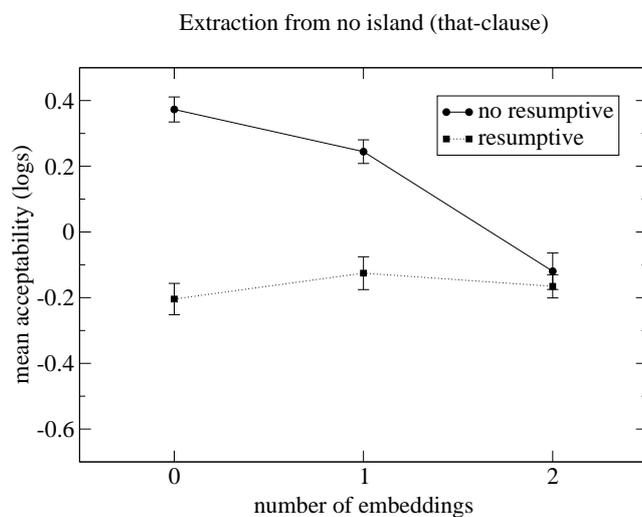The same procedure as in Experiment 1 was used.

Figure 5: Effect of embedding and resumption on extraction in Greek (*that*-clause condition)
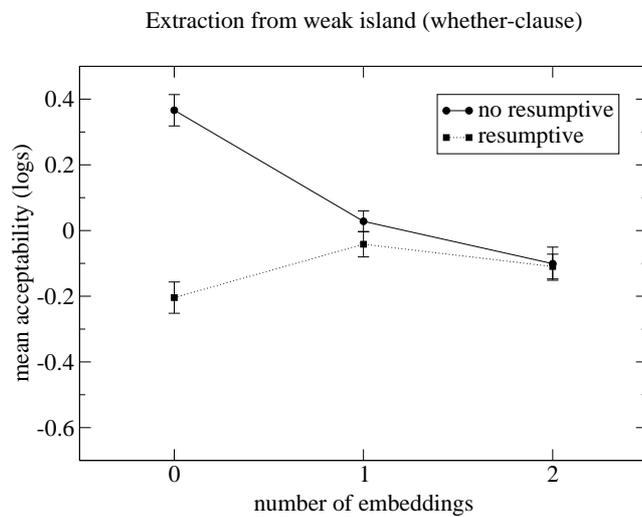

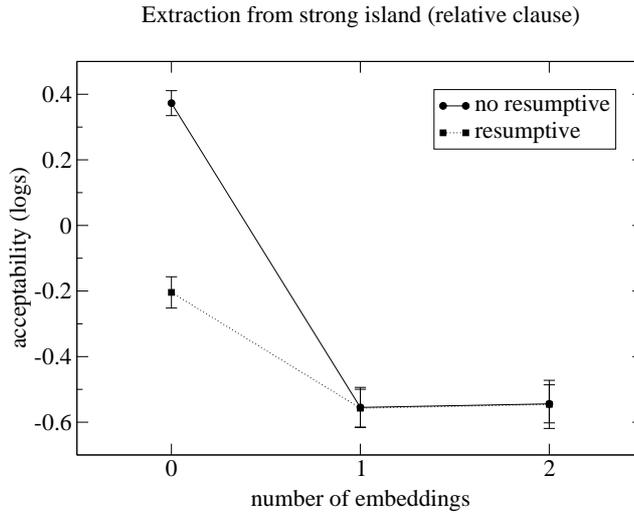Figure 6: Effect of embedding and resumption on extraction in Greek (*whether*-clause condition)

11

Extraction from strong island (relative clause)

Figure 7: Effect of embedding and resumption on extraction in Greek (relative clause condition)

3.3 Results

The data were normalized and log-transformed as in Experiment 1. Figures 5–7 graph the mean judgments for all three island conditions.

An ANOVA yielded significant main effects of *Embedding* ($F_1(1,58) = 26.509$, $p < .0005$; $F_2(1,6) = 19.933$, $p = .004$), *Island* ($F_1(2,116) = 82.828$, $p < .0005$; $F_2(2,12) = 137.211$, $p < .0005$), and *Resumption* ($F_1(1,58) = 22.875$, $p < .0005$; $F_2(1,6) = 12.006$, $p = .013$). We are mainly interested in an interacting of *Island* and *Resumption*, as this indicates that the acceptability of resumptives is sensitive to island violations. This interaction was significant ($F_1(2,116) = 10.005$, $p < .0005$; $F_2(2,12) = 4.016$, $p = .046$). All the other interactions were only significant by subjects: *Island/Embedding* ($F_1(1,116) = 15.072$, $p < .0005$; $F_2(2,12) = 3.409$, $p = .067$), *Embedding/Resumption* ($F_1(1,58) = 7.705$, $p = .007$; $F_2(1,6) = 4.494$, $p = .078$), and *Island/Embedding/Resumption* ($F_1(2,116) = 5.888$, $p = .004$; $F_2(2,12) = 3.872$, $p = .050$).

As in Experiment 1, we conducted a Tukey post-hoc test to further investigate the interaction of *Island* and *Resumption*. In the no island condition (complement clause with *that*), the gap was more acceptable than the resumptive ($\alpha < .01$). There was no significant difference between the gap and the resumptive in either the weak island condition (complement clause with *whether*) or in the strong island condition (relative clause).

Dunnett's test was again used to compare the embedded conditions to the control conditions. We will first report the results for the gapped stimuli. For the no island condition, both the single and the double embedding condition were less ac-

ceptable than the control ($t_{d_1}(59,7) = 5.641, p < .01; t_{d_2}(6,7) = 4.472, p < .05$ and $t_{d_1}(59,7) = 8.695, p < .01; t_{d_2}(6,7) = 9.562, p < .01$). Also in the weak island condition, both levels of embedding were significantly less acceptable than the control ($t_{d_1}(59,7) = 8.619, p < .01; t_{d_2}(6,7) = 14.428, p < .01$ and $t_{d_1}(59,7) = 7.532, p < .01; t_{d_2}(6,7) = 6.005, p < .01$). The same picture emerged in the strong island condition, again both levels of embedding were worse than the control ($t_{d_1}(59,7) = 12.323, p < .01; t_{d_2}(6,7) = 12.017, p < .01$ and $t_{d_1}(59,7) = 12.470, p < .01; t_{d_2}(6,7) = 17.066, p < .01$).

A separate test compared the resumptive stimuli to the resumptive controls. In the no island condition, neither the single nor the double embedding were significantly different from the control. In the weak island condition, the single embedding condition was significantly more acceptable than the control, by subjects only ($t_{d_1}(59,7) = 3.034, p < .05; t_{d_2}(6,7) = 2.930, p > .05$). There was no difference between the double embedding condition and the control. In the strong island condition, both the single and the double embedding condition were significantly less acceptable than the control ($t_{d_1}(59,7) = 4.955, p < .01; t_{d_2}(6,7) = 7.058, p < .01$ and $t_{d_1}(59,7) = 4.284, p < .01; t_{d_2}(6,7) = 6.107, p < .01$).

3.4 Discussion

Our results partially confirm the picture we obtained from the English experiment, but there are some important differences.

Prediction (a). As in English and unlike what was expected, embedding induces a reduction in the acceptability of gaps, which in the case of Greek was significant for both levels of embedding (see Figure 5). Embedding was expected to improve the acceptability of resumptives and resumptives were expected to be as acceptable as gaps after the first level of embedding. However, resumptives were equally unacceptable in the control and in the single and double embedding conditions. More importantly, they were significantly worse than gaps in the first level of embedding, though as acceptable as gaps in the second level of embedding.

Prediction (b). Under the assumption that weak islands do not restrict movement in Greek *whether*-clauses were expected to yield the same results as *that*-clauses and crucially embedding was not expected to affect the acceptability of gaps. However, singly and doubly embedded gaps were significantly less acceptable than the control condition (see Figure 6). Here weak islands do behave on a par with the no island condition. It was also discovered that resumptives were more acceptable in single embedding than in the control, a fact which lends partial support to the claim that resumptives can save islands. Further support for this claim is provided by the fact that resumptives and gaps were equally acceptable in the single and double embedding conditions.

Prediction (c). Here the results are in line with our predictions and on a par with English except for one point. While in English the resumptive in the control was as bad as in the embedded conditions, in Greek both singly and doubly embedded resumptives were less acceptable than the control. This indicates that non-embedded

resumptives (control condition) are more acceptable in Greek than in English, assuming that strong island constraints are equally unacceptable in both languages.

4   Conclusions

Our studies failed to validate the "saving" effect of resumption in weak islands, since at best resumptives were as good as gaps. However, it appears that resumption does reverse the effect weak islands have on gap sentences in Greek, where we found a significant improvement of the embedded resumptive compared to a non-embedded control.

An important crosslinguistic difference is that in English, resumptives are generally significantly worse than gaps (in all levels of embedding). A possible explanation for this is the fact that non-embedded resumptives appear to be more acceptable in Greek than in English: in English they are as bad as strong island violations, while in Greek, they are better than strong island violations. This difference could be viewed as a consequence of the availability of productive resumption in Greek in contexts other than *wh*-questions.

It is further worth noting that the interactions gap/embedding and resumptive/embedding were of the same nature in both *that* and *whether* clauses. This result is important since only the case of weak island is taken to involve a grammatical violation. The effects in *that*-clauses could be explained in terms of processing constraints. But the similarity of the interactions in the two cases indicates that, a processing explanation is needed also for the case of weak islands, possibly incorporating the grammatical constraints that weak islands are subject to (relevant processing models have been proposed by Dickey (1996) and Erteschick-Shir (1992)).

Acknowledgments

References

Alexopoulou, Theodora. 1999. *The Syntax of Discourse Functions in Greek; a non-configurational approach.* Ph.D. thesis, University of Edinburgh.

Bard, Ellen Gurman, Dan Robertson, and Antonella Sorace. 1996. Magnitude estimation of linguistic acceptability. *Language* 72/1:32–68.

Corley, Martin, and Christoph Scheepers. 2002. Syntactic priming in English sentence production: Categorical and latency evidence from an internet-based study. *Psychonomic Bulletin and Review* 9/1.

Cowart, Wayne. 1997. *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. Thousand Oaks, CA: Sage Publications.

Dickey, Michael Walsh. 1996. Constraints on the sentence processor and the distribution of resumptive pronouns. In Michael Walsh Dickey and Susanne Tunstall, eds., *Linguistics in the Laboratory*, no. 19 in Occasional Papers in Linguistics, 157–192. Amherst, MA: University of Massachusetts.

Erteschick-Shir, Nomi. 1992. Resumptive pronouns in islands. In Helen Goodluck and Michael Rochemont, eds., *Island Constraints: Theory, Acquisition and Processing*, vol. 15 of *Studies in theoretical psycholinguistics*, 89–108. Kluwer Academic Publishers.

Haegeman, Lilian. 1991. *Introduction to Government and Binding Theory*. Blackwell.

Keller, Frank, and Theodora Alexopoulou. 2001. Phonology competes with syntax: Experimental evidence for the interaction of word order and accent placement in the realization of information structure. *Cognition* 79/3:301–372.

Tsimpli, Ianthi-Maria. 1999. Null operators, clitics and identification: a comparison between Greek and English. In Artemis Alexiadou, Geoffrey Horrocks, and Melita Stavrou, eds., *Studies in Greek syntax*, 241–262. Kluwer Academic Publishers.

Tsimpli, Maria Ianthi. 1995. Focusing in Modern Greek. In K. E. Kiss, ed., *Discourse Configurational Languages*, Oxford Studies in Comparative Syntax, 176–206. Oxford University Press.