Resumption still does not rescue islands[1]

Dustin Heestand[a], Ming Xiang[b], and Maria Polinsky[a]

(*Harvard University[a] and University of Chicago[b]*)

1. Introduction.

Since Ross 1967, island constraints have been a major topic in syntactic research; however, the status of different types of islands and their psychological reality to this day remain subjects of hot debate. Environments where island constraints appear to be violated are particularly important to our understanding of the constraints. Resumptive pronouns (RPs) have traditionally been cited as an island-rescuing device in English and other languages (Ross 1967, Kroch 1981, Erteschik-Shir 1992). And indeed, in some languages, RPs show immunity from island constraints (Aoun et al. 2001, McCloskey 2006), as in the following Lebanese Arabic example:

(1) ħdrna   l-masraħiyye yalli   tʕarrafna   ʕala l-muχriʒ   yalli   ʔaχraʒ-*(a)

    saw.1PL the-play   that   met.1PL on   the-director that   directed.3SM-*(it)

'We saw the play that we met the director that directed it.' (Aoun and Choueiri 1996, ex. (12))

In English, RPs have been reported to appear in island and non-island contexts, as attested in spontaneous speech (Prince 1990) and in laboratory production studies (Zukowski and Larsen 2004, Ferreira and Swets 2005):

(2) a. I have this friend who *she* does all the platters. (Prince 1990)

    b. This is the donkey that I don't know where *it* lives. (Ferreira and Swets 2005)

    c. The man who the spider is falling on *his* head… (Zukowski and Larsen 2004)

This productive use of RPs might suggest that resumption is a strategy available to the grammar of English just as it is available to the grammar of languages like Irish or Lebanese Arabic. On this view, the main difference may be in how acceptable languages find this strategy, and how often they use it. In English, for instance, resumption is extremely rare (Creswell 2002; Hermann 2005; Jaeger 2006; Manetta 2007; Bennett 2008), so it may have the same "last resort" character that Shlonsky (1992) proposes for Hebrew and Palestinian Arabic (both languages with fully grammatical resumption). When speakers have painted themselves into a syntactic corner, they use RPs to salvage what they can of their intended meaning. Nevertheless, this use of RPs forms part of the grammatical knowledge of native English speakers, and therefore a descriptively adequate syntactic theory should have a means of representing them. Assuming such an account, the relative ungrammaticality of resumption in English might result merely from frequency effects.

Turning now to experimental investigation of RPs, there is a deep discrepancy between production studies and comprehension studies. On the production side, Zukowski and Larsen (2004) and Ferreira and Swets (2005) asked the participants to judge the acceptability of sentences using the same resumptive structure that the participants had readily produced (cf. (2b, c) above). Resumptive structures were consistently rated significantly below the grammatical controls. However, these studies did not compare the RPs with their illicitly gapped counterparts, and therefore were inconclusive as to whether resumption can rescue islands.

Alexopoulou and Keller 2007 (A&K) was the first study to extensively test the rescuing ability of RPs in comprehension. They tested extraction of wh-elements out of relative clauses (strong-island condition), *whether*-clauses (weak-island condition), and *that*-clauses (control condition) in English, Greek and German. For each condition, they tested for multiple levels of embedding, for

instance:

(3) a. Who does Mary wonder [whether we will fire __ /him]? (*single embedding*)

b. Who does Jane think [that Mary wonders [whether we will fire __ /him]]? (*double emb.*)

Across different languages and conditions, A&K's results consistently showed that when extracting from an island, strong or weak, a resumptive structure was never more acceptable than its gapped counterpart. On the other hand, the violation caused by the RP was ameliorated by increased syntactic distance. The latter finding is particularly surprising in light of corpus studies (Bennett 2008, Jaeger 2006: 101), which show English resumption to be more common for highest subjects than for embedded subjects.

If we combine all these findings, a paradox emerges: first, resumption in English seems to be most common where it is least acceptable; and second, the amelioration effect is not reciprocal. That is, more deeply embedded islands can improve the acceptability of RPs, but RPs can never improve the acceptability of island violations.

The current study takes this paradox seriously and attempts to fill two gaps in the previous studies. First, we wish to test the acceptability of more types of islands in both declarative and wh-question contexts. It is well known that syntactic islands are not a homogeneous group, as they show amenability in various degrees to gap creation (Phillips 2006, Sprouse et al. in press, Sprouse et al. 2010), and thus, presumably, to RPs. In testing resumption, A&K did not test declarative statements, where resumption is rampant in production (instead, they tested wh-questions, where no such effect has been observed). In order to strike at the heart of the production-comprehension mismatch, an investigation of RPs in declaratives is necessary.

Second, RP acceptability should be tested in an online, not offline task (A&K used the latter). One reason that RPs occur more often in production might be due to the temporal constraint

people are facing during production. In order to address this possibility, we need studies of RP acceptability in an online comprehension task, which puts people under a similar kind of time pressure. Setting a longer-term agenda, we need to obtain online comprehension results for both visual (reading) and auditory presentation. If we find no rescuing effect of RPs in any of these circumstances, we have stronger evidence that at least for English, resumption has no capacity to amelorate island constraint violations in comprehension. The visual presentation is necessary to better contextualize the results reported by A&K and to remove the possible confounds their study had. The auditory presentation is necessitated by the fact that RPs are very much a spoken register phenomenon (see also Bennett 2008 and Jaeger 2006). Looking back to the many years of introspection into resumption, auditory presentation has always been the dominant method of establishing if something is acceptable or not: linguists would say a sentence to themselves or their friends.

This squib takes the first step to this end, presenting three experiments on reading. The first one tested RPs in complex NP islands, using an offline judgment task (the data which are needed to fill in the gap from A&K's study). The second and third experiments used online judgment tasks, in which we tested relative clauses islands and adjunct islands. In all these experiments, we also compared declarative sentences with wh-questions; the latter allowed us a more direct comparison with A&K's results.

2. EXPERIMENT 1: Resumptive pronouns within complex NP islands.

In this experiment we tested the rescuing function of RPs for complex NP islands, using two types of complex NP islands, one with a factive complement clause and one with a standard relative clause: Experiment 1a examines RPs in factive clauses; Experiment 1b, in relative clauses.

2.1. EXPERIMENT 1a. Factive Complement Clauses

Materials, subjects, procedure. In a 2x3 design, we crossed the following two factors: construction type and gap type. For the first factor, construction type, participants read either a wh-construction or a complex–NP (factive clause) construction; for the second factor, gap type, participants read a sentence with a gap within the island, a sentence with a resumptive pronoun within the island, or a grammatical control sentence with a gap in the main clause. Example sentences are given in (4):

(4)   Factive clauses—Declaratives

a.   This is the man that the news that the police arrested __ shocked the public. (*Gap in an island*)

b.   This is the man that the news that the police arrested **him** shocked the public. (*RP in an island*)

c.   This is the man that Mary thought that the police arrested __ to protect the president. (*Grammatical control*)

    Factive clauses—Wh-questions

d.   Which man did the news that the police arrested __ shock the public? (*Gap in an island*)

e.   Which man did the news that the police arrested **him** shock the public? (*RP in an island*)

f.   Which man did Mary think that the police arrested __ to protect the president? (*Grammatical control*)

There were a total of 30 sets of experimental items, and each had the six conditions described above. All the experimental sentences were distributed into six lists with a Latin-square design. In addition, there were 108 fillers. Each participant was given a printed copy of one of the lists, with a total of 138 sentences in a randomized order. The participants were instructed to score the acceptability of each sentence on a 1 to 7 scale, where 7 indicated perfect acceptability and 1, unacceptability. They were instructed to judge based on their native-speaker intuition rather than any prescriptive rules, and to go with their first instincts rather than spending time pondering their answers. 18 native speakers of English from the Boston area participated in our study.

Results. Average ratings for Experiment 1a are shown in Figure 1. The 2x3 ANOVA reveals a

main effect of gap type both by subjects (F(2, 34)=68.9,p<.001) and by items (F(2, 58)=122.5,p<.001). We did not find any main effect of construction type, nor any interactions. Planned comparisons found that the main effect of gap type was driven by the difference between the grammatical control and the ungrammatical conditions. In the group of the declarative constructions, the control condition was rated significantly higher than the gapped condition (t1(1, 17)=10.6, p<.001; t2(1,29)=11.2, p<.001), and the resumption condition (t1(1,17)=9.3, p<.001; t2(1,29)=12, p<.001). Similarly, for the group of wh-constructions, the control condition was rated higher than the gapped condition (t1(1, 17)=5.9, p<.001; t2(1. 29)=7, p<.001), and the resumption condition (t1(1, 17)=5.9, p<.001; t2(1, 29) =9.7, p<.001). No other significant difference was observed. Critically, there is no difference between the gapped condition and the resumption condition at all, in either the declarative or the wh-constructions.

2.2 EXPERIMENT 1b: Relative Clauses

Materials, procedure and subjects. This experiment focuses on RPs within relative clause islands. The experimental design is the same as Experiment 1a. An example of the stimuli is given in (5):

(5)   Relative clauses—Declaratives

a.   This is the man that the policeman who arrested __ saved the President's life. (*Gap in an island*)

b.   This is the man that the policeman who arrested **him** saved the President's life. (*RP in an island*)

c.   This is the man that the policeman who arrested the thief saved __. (*Grammatical control*)

         Relative clauses—Wh-questions

d.   Which person did Mary think that the policeman who arrested __ saved the President's life? (*Gap in an island*)

e.   Which person did Mary think that the policeman who arrested **him** saved the President's life? (*RP in an island*)

f.   Which man did Mary think that the policeman who arrested the thief saved__? (*Grammatical control*)

There were 30 sets of experimental items and 114 fillers. The procedure was the same as Experiment 1a. 18 native speakers of English from the Boston area participated in the study.

Results. Average ratings for Experiment 1b are shown in Figure 2. The 2x3 ANOVA found a main effect of construction type, which is marginal by subject analysis ($F(1, 17)=4.05, p=.06$), and significant by item analysis ($F(1,29)=9.3$, $p<.01$). There is also a main effect of gap type ($F_1(2,24)=99.8$, $p<.001$), $F_2(2, 58)=160.4$, $p<.001$). In addition, there is a significant interaction ($F_1(2, 34)=3.85$, $p<.05$; $F_2(2, 58)=3.27$, $p<.05$). Planned comparisons found that in the group of declarative constructions, the control condition was rated significantly higher than the gapped conditions ($t_1(1, 17)=11.1$, $p<.001$; $t_2(1,29)=9.6$, $p<.001$), and the resumption condition ($t_1(1,17)=10.1$, $p<.001$; $t_2(1,29)=11.3$, $p<.001$). Similarly, for the group of wh-constructions, the control condition is rated higher than the gapped condition ($t_1(1, 17)=6.3$, $p<.001$; $t_2(1. 29)=7.9$, $p<.001$), and the resumption condition ($t_1(1, 17)=7.4$, $p<.001$; $t_2(1, 29) =11.7$, $p<.001$). In addition, the grammatical wh-construction was rated significantly lower than the grammatical declarative construction ($t_1(1, 17)=-2.5$, $p<.05$; $t_2(1, 29)=-3.0$, $p<.01$). Again, critically there is no difference between the resumption conditions and the gapped conditions in any of the comparisons. <Figures 1, 2 here>

Summary of Experiment 1. The offline acceptability judgment results in Experiment 1 showed that for complex NP islands, including both factive complement clauses and relative clauses, resumption has no rescuing effect. Subjects disfavored resumption as much as sentences with illicit gaps. For the grammatical controls, at least for relative clauses, wh-constructions were rated lower than declarative constructions. This difference could be due to the fact that in our material, the wh-controls are slightly longer than the declaratives, and hence a bit more complex.

With respect to the null-rescuing function of RPs, a number of studies have suggested that RPs

are used as a "last-resort" strategy. If the acceptability of resumptions is related to processing resources, it raises the possibility that people might consider RPs as more acceptable than gaps when they have a limited amount of time to process them. To test if this is the case, in Experiment 2, we repeated Experiment 1b in an online acceptability judgment task. An additional advantage of the online task is that it allows us to collect response time data, in addition to the acceptability judgment data.

3. Experiment 2: Online acceptability judgments of relative clause islands

<u>**Materials, procedure, and subjects.**</u> The materials were the same as in Experiment 1b. The experiment was implemented using Linger (Rohde 2007). All the sentences are automatically randomized. Each sentence was presented word by word, with 400 ms presentation for each word. After the last word of each sentence, participants used the mouse to choose on a 1 to 7 acceptability scale (7: perfectly grammatical, 1: ungrammatical). 24 native speakers of English from the Boston area participated in the study.

<u>**Results.**</u> The average rating results for Experiment 2 are presented in Figure 3. A 2x3 ANOVA found a main effect of construction type ($F_1(1, 23)=9.15, p<.01$, $F_2(1,29)=10.9$, $p<.01$). There was also a main effect of gap type ($F_1(2,46)=105.7$, $p<.001$), $F_2(2, 58)=164.5$, $p<.001$). In addition, there was a significant interaction ($F_1(2, 46)=16.7$, $p<.001$; $F_2(2, 58)=11.9$, $p<.001$). Planned comparisons found that in the group of declarative constructions, the control condition is rated significantly higher than the gapped condition ($t_1(1, 23)=10.8$, $p<.001$; $t_2(1,29)=14.5$, $p<.001$), and than the resumption condition ($t_1(1,23)=10.9$, $p<.001$; $t_2(1,29)=15.0$, $p<.001$). Similarly, for the group of wh-constructions, the control condition was rated higher than the gapped condition ($t_1(1, 23)=6.7$, $p<.001$; $t_2(1. 29)=7.0$, $p<.001$), and the resumption condition ($t_1(1, 23)=7.1$, $p<.001$; $t_2(1, 29) =7.3$, $p<.001$). In addition, the grammatical wh-construction was rated significantly lower than

the grammatical declarative construction (t_1(1, 23)=-4.5, p<.001; t_2(1, 29)=-4.6, p<.001). There is no difference between the resumption conditions and the gapped conditions.

Turning now to the reaction times, RTs longer than 4500ms (2 standard deviations from the mean) were not included for data analysis. The mean RT result is presented in Figure 4. A 2x3 ANOVA found no main effect of construction type (F_1(1, 23)=0.5, p>.5, F_2(1,29)=0.2, p>.5). There is a significant effect for gap type (F_1(2, 46)=3.5, p<.05; F_2(2, 58)=4.6, p<.05.). No interaction was found (F_1(2, 46)=0.8, p>.1; F_2(2, 58)=1.4, p>.1). Planned comparisons found that the significant effect of gap type was mainly driven by the difference between the resumption conditions and other conditions. For the group of wh-constructions, the RTs for the resumption condition were significantly shorter than for the control condition (t_1(1, 23)=-2.4, p<.05; t_2(1. 29)=-2.7, p<.05). There was also a numerical trend that the RT for the resumption condition was also shorter than the gap condition, but this difference did not reach significance. For the declarative group, the RTs for the resumption condition were marginally shorter than the control condition by subject analysis (t_1(1, 23)=-1.8, p=.08; t_2(1. 29)=-1.5, p>.1); and also in this group the RTs for the resumption condition were significantly shorter than the gap condition by item analysis, and this difference is marginally significant by subject analysis (t_1(1, 23)=-1.8, p=.08; t_2(1. 29)=-2.4, p<.05). No other difference was observed.   <Figures 3, 4 here>

Summary of Experiment 2. The online rating results from Experiment 2 replicated the offline results in Experiment 1b. Once again, RPs did not show any rescuing effect, even when subjects were under time pressure. Interestingly, the RT data also showed that people actually judged the resumption condition fastest. In other words, rather than having any rescuing effect, RPs made it easier for people to detect the ungrammaticality of the sentences. In addition, as in Experiment 1b, we also observed that grammatical wh-constructions are again rated lower than grammatical

declarative constructions. The RT difference on the two grammatical conditions, however, did not reach significance.

Overall, Experiments 1 and 2 showed that RPs have no rescuing effect for violations of complex NP islands. Complex NPs are generally considered strong islands (Postal 1998, Szabolcsi 2006). In the last experiment, we tested RPs in adjunct clauses, at least some of which have been considered weak(er) islands (Cinque 1990; Truswell 2007), thus establishing a contrast with the strong islands considered in previous work.

4. Experiment 3: Online acceptability judgments of adjunct islands

<u>Materials, procedure, and subjects.</u> The experimental design is the same as the previous experiments, except that the target stimuli are adjunct islands. In theoretical work, such islands are considered strong; an experimental investigation by Hiramatsu (1999, 2000) shows their status is particularly opaque. An example of the experimental stimuli is given in (6).

(6)   <u>Adjunct Clause—Declaratives</u>

a.   This is the dish that, although the chef overcooked __, the guests were not upset. (*Gap within an island*)

b. This is the dish that, although the chef overcooked ***it***, the guests were not upset. (*RP within an island*)

c. This is the dish that, although the chef overcooked the sauce, the guest enjoyed __. (*Grammatical control*)

        <u>Adjunct Clause—Wh-questions</u>

d. Which dish did Gina think that, although the chef overcooked __, the guests were not upset? (*Gap within an island*)

e. Which dish did Gina think that, although the chef overcooked ***it***, the guests were not upset? (*RP within an island*)

f. Which dish did Gina think that, although the chef overcooked the sauce, the guests enjoyed __?

(*Grammatical control*)

There were a total of 30 sets of experimental sentences and 114 fillers that were used in an online rating task. The procedure is the same as Experiment 2. 24 native speakers of English were recruited from the Boston area to participate in our study.

Results. The average rating results for Experiment 3 are presented in Figure 5. A 2x3 ANOVA found a main effect of construction type ($F_1(1, 23)=13.8$, $p<.01$, $F_2(1,29)=22$, $p<.001$). There was also a main effect of gap type ($F_1(2,46)=4.5$, $p<.05$), $F_2(2, 58)=7.0$, $p<.01$). There was no interaction ($F_1(2, 46)=1.2$, $p>.1$; $F_2(2, 58)=1.6$, $p>.1$). Planned comparisons found that for the group of declarative constructions, the control condition was rated significantly higher than the gapped conditions by items, and the difference was only marginal by subjects ($t_1(1, 23)=1.8$, $p=.09$; $t_2(1,29)=2.7$, $p<.05$). For the same declarative group, there is no difference between the control and the resumption condition ($t_1(1,23)=1.5$, $p>.1$; $t_2(1,29)=1.7$, $p=0.1$). For the group of wh-constructions, the control condition is rated higher than the resumption condition ($t_1(1, 23)=2.9$, $p<.01$; $t_2(1. 29)=3.4$, $p<.01$), but it is only higher than the gap condition by item analysis ($t_1(1, 23)=1.6$, $p>.1$; $t_2(1, 29) =2.2$, $p<.05$). In addition, the grammatical wh-construction is rated significantly lower than the grammatical declarative construction ($t_1(1, 23)=-2.2$, $p<.05$; $t_2(1, 29)=-2.5$, $p<.05$). No other difference was observed.

The RT mean (ms) for Experiment 3 is presented in Figure 6. Again, RTs longer than 4500ms were not included in the data analysis. We did not observe any significant difference in the mean RT data for Experiment 3.   <Figures 5, 6 here>

Summary of Experiment 3. The rating results in Experiment 3 show that adjunct islands, at least the kind presented here, are very weak: sentences without any island violations only had a slight advantage in rating over those with island violations. This result adds to the growing body of evidence that adjunct islands are a heterogeneous group (see also Hiramatsu 2000, Truswell 2007,

Sprouse et al. 2010). The particular adjunct islands we used (*although-, while*-islands) consistently show greater transparency. The understanding of which adjunct islands are weaker and why is still outstanding, and we are not ready to offer any speculations. However, critically for our purposes, sentences with RPs again showed no advantage over sentences with gaps.

5. Conclusions.

This paper took as its starting point the conclusion, reached on the basis of experimental work, that English resumption does not ameliorate island effects (Alexopoulou & Keller 2007). The reasons to question this result had to do with the methodology (offline reading-based judgments) and the nature of islands tested. Our own study included the contexts where resumption in English is more commonly found (declaratives; relative and adjunct clauses) and employed a different methodology. Nonetheless, resumption still failed to rescue the island violations and was judged ungrammatical. While not the final word in the long-standing debate on the role of resumption in English (the next logical step, which we are now planning, is to test the auditory presentation of the offending stimuli), this result suggests that the role English RPs play in production is not that of island rescuers.

If this is the case, what is the function of RPs in English? We have shown that resumption does not help the hearer, or more accurately, the reader (and it has been reasonably well established that it is ungrammatical in written discourse). One possible explanation is that English resumption is, unlike Irish resumption, not a strategy for establishing A'-binding relations, but rather something more similar to cross-sentential anaphora (see also Erteschik-Shir 1992). If this is the case, performance pressures in production could lead to speakers resorting to such anaphora as a way of adding more information without breaking the production chain. If this suggestion is on the right track, the difference between production and comprehension with respect to

resumption falls outside the domain of grammar and pertains to the planning of an utterance. That in turn would account for the paradox we noted earlier: naturally occurring resumption is more common in the subject position of an embedded or relative clause, which is also the context where it is judged most ungrammatical. Subjects (or maybe topics) are privileged with respect to coreference across clauses and in discourse (Keenan and Comrie 1977, Comrie 1987, Lambrecht 1994, Erteschik-Shir 2007, a.o.), and this privileged status with respect to coreference would favor them over other arguments in the use of RPs by the coreference-marking speaker. Thus, resumption in English may be still another instance of phenomena where, contrary to belief, speakers structure an utterance to meet their own needs, in addition to the needs of the hearer (for other instances of speakers following their needs rather than those of the listeners, see Brennan and Clark 1996, Engelhardt et al. 2006, a.o.). Finally, if the use of RPs in English is nothing more than a speaker-centered device for maintaining coreference, we are in a position to better differentiate it from non-intrusive resumption in such languages as Irish (McCloskey 2006 and references therein).

With further experimentation, intrusive resumption in English could be differentiated from yet another type of licit resumption: that found in Italian. In Italian, left-dislocated elements may be doubled by clitic RPs, but the two elements may not be separated by an island boundary, indicating that movement is involved (Cinque 1990). Therefore, we would expect RPs in Italian to surpass the acceptability of gaps under deep embedding in the absence of an island, in contrast to their behavior in English, Greek, or German (Alexopoulou and Keller 2007). However, in the presence of an island, Italian judgments should mirror those in the present experiments.

Finally, we would like to remind the reader that ungrammatical sentences with RPs were judged in our experiments as bad very quickly. In cases where RP judgments were faster than

judgments for sentences with illicit gaps, the gaps seem to be less helpful to the parser, despite being just as unambiguously ungrammatical. This tendency was the clearest in complex NPs with factive complements, which are independently known to be strong islands. Future work on other strong islands will determine if this effect holds. If it does, we can conclude that the extra information available in RPs is useful in parsing difficult dependencies, making its unacceptability all the more puzzling.

References

Alexopoulou, Theodora and Frank Keller. 2007. Locality, cyclicity and resumption: At the interface between the grammar and the human sentence processor. *Language* 83, 110-160.

Aoun, Joseph and Lina Choueiri. 1996. Resumption and last resort. Unpublished manuscript, University of Southern California.

Aoun, Joseph, Lina Choueiri, and Norbert Hornstein. 2001. Resumption, movement, and derivational economy. *Linguistic Inquiry* 32, 371-403.

Benett, Ryan. 2008. English resumptive pronouns and the highest-subject restriction: A corpus study. Paper presented at the Trilateral (TREND) Linguistics Weekend, UC Santa Cruz, May 2009.

Brennan, Susan, and Herbert Clark. 1996. Conceptual pacts and lexical choice in conversation. Journal of Experimental Psychology: Learning, Memory and Cognition 22, 1482-1493.

Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, Mass.: MIT Press.

Comrie, Bernard. 1987. *Language universals and linguistic typology*. Chicago: University of Chicago Press.

Creswell, Cassandre. 2002. Resumptive pronouns, wh-island violations, and sentence production. *Proceedings of the Sixth International Workshop on Tree-Adjoining Grammar and Related Frameworks (TAG+6)*, 101-109. Venice: Università di Venezia.

Engelhardt, Paul, Karl Bailey, and Fernanda Ferreira. 2006. Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language* 54, 554-573.

Erteschik-Shir, Nomi. 1992. Resumptive pronouns in islands. In Helen Goodluck and Michael Rochemont (eds.). *Island constraints: Theory, acquisition, and processing*, 89-108. Dordrecht: Kluwer.

Ferreira, Fernanda and Benjamin Swets. 2005. The production and comprehension of resumptive pronouns in relative clause "island" contexts. In Anne Cutler (ed.), *Twenty-first Century psycholinguistics: Four cornerstones,* 263-278. Mahway, NJ: Lawrence Erlbaum Associates.

Herrmann, Tanja. 2005. Relative clauses in English dialects of the British Isles. In Bernd Kortmann, Tanja Herrmann, Lukas Pietsch, and Susanne Wagner (eds.). *A comparative grammar of British English dialects*, 21-124. Berlin: Mouton.

Hiramatsu, Kazuko. 1999. Subject and adjunct island asymmetries: Evidence from linguistic

satiation. *WCCFL* 18, 183-192.

Hiramatsu, Kazuko. 2000. Accessing linguistic competence: Evidence from children's and adults' acceptability judgments. Ph.D. Diss., University of Connecticut.

Jaeger, Florian. 2006. Redundancy and syntactic reduction in spontaneous speech. Ph.D. Diss., Stanford University.

Keenan, Edward, and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8, 63-99.

Kroch, Anthony. 1981. On the role of resumptive pronouns in amnestying island constraint violations. *Chicago Linguistic Society: Proceedings of the 17th Regional Meeting*, 125–35.

Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.

Manetta, Emily. 2007. Unexpected left dislocation: An English corpus study. *Journal of Pragmatics* 39, 1029-1035.

McCloskey, James. 2006. Resumption. In Martin Everaert and Henk van Riemsdijk (eds.) *The Blackwell Companion to Syntax*, 94–117. Oxford: Blackwell Publishing.

Phillips, Colin. 2006. The real-time status of island phenomena. *Language* 82, 795-823.

Prince, Ellen. 1990. Syntax and discourse: A look at resumptive pronouns. *Berkeley Linguistic Society: Proceedings of the 16th Annual Meeting*, 482-497.

Postal, Paul. 1998. *Three investigations of extraction*. Cambridge, MA: MIT Press.

Ross, John R. 1967. Constraints on variables in syntax. Ph.D. Diss., MIT.

[Published as *Infinite syntax!* Norwood, NJ: Ablex, 1986.]

Sprouse, Jon, Shinichiro Fukuda, Hajime Ono, and Robert Kluender. In press. Grammatical operations, parsing processes, and the nature of wh-dependencies in English and Japanese. *Syntax*.

Sprouse, Jon, Matt Wagers, and Colin Phillips. 2010. A test of the relation between working memory capacity and syntactic island effects. Manuscript, UC Irvine, UC Santa Cruz, and University of Maryland.

Szabolcsi, Anna. 2006. Strong vs. weak islands. In Martin Everaert and Henk van Riemsdijk (eds.) *The Blackwell companion to syntax, Volume IV*, 479–531. Oxford: Blackwell.

Shlonsky, Ur. 1992. Resumptive pronouns as last resort. *Linguistic Inquiry* 23, 443-468.

Truswell, Robert. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117, 1355-1377.
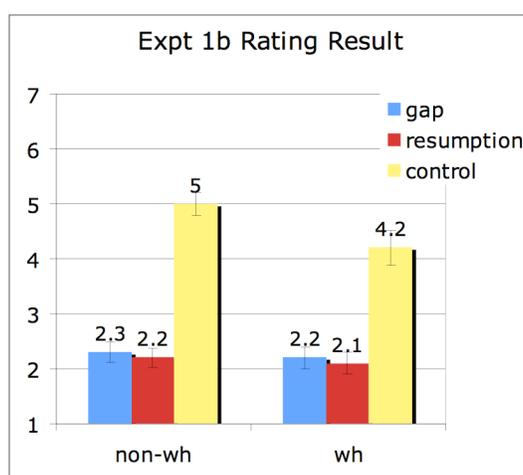
Zukowski, Andrea, and Jenny Larsen. 2004. The production of sentences that we fill their gaps. Poster presented at the CUNY Sentence Processing Conference, University of Maryland.
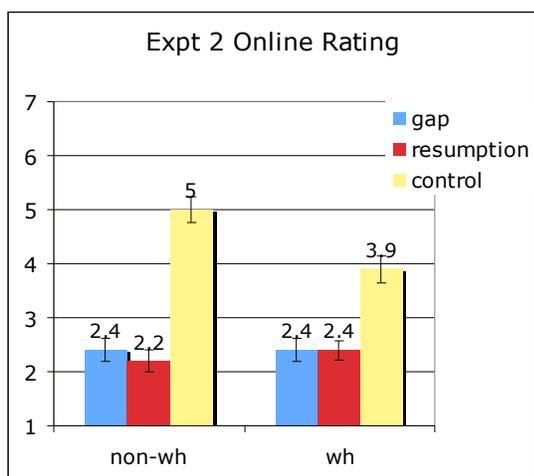
Figures
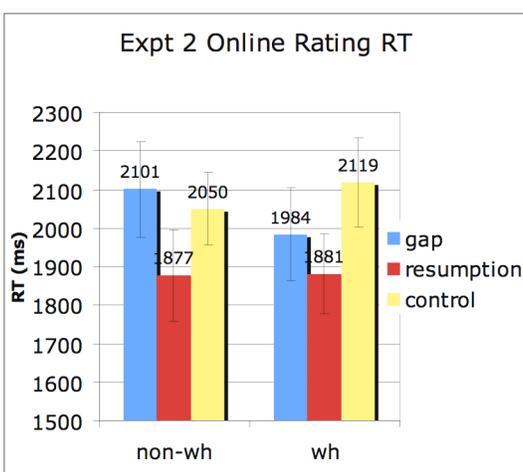
Figure 1. Mean Ratings for Expt 1a
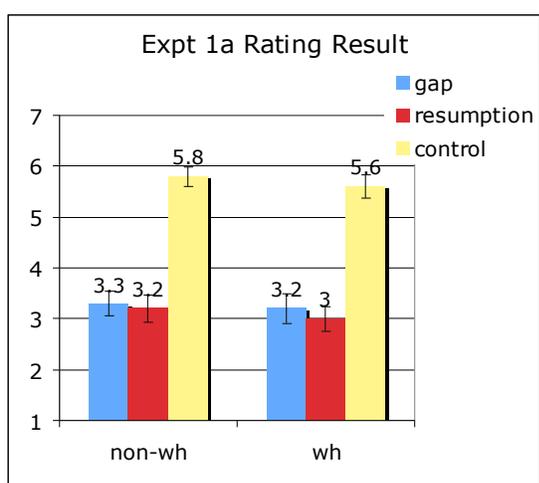
Figure 2. Mean Rating for Expt 1b

Figure 3 Ratings in Expt2

Figure 4 Mean RTs for Expt 2

Figure 5 Mean Rating for Expt 3

Figure 6 Mean RT(ms) for Expt 3

Expt 3 Online Rating

Expt 3 Online Rating RT