Resumptive Pronouns, Wh-island Violations, and Sentence Production

Cassandre Creswell

University of Pennsylvania

1. Introduction

In spontaneous speech, English speakers produce relative clauses whose structures violate wh-island constraints but where resumptive pronouns appear in place of the gap/trace normally found in a relative clause. It is difficult to explain these island-violation remedying resumptive pronouns in a model of sentence production where only grammatically-licensed trees are available to the speaker's production system. This paper argues that the only explanation for these island-violating relative clauses in such a model of sentence production is that they are fully grammatical constructions, whose marginal status reflects non-syntactic factors.

The paper is structured as follows. The next section outlines the problem that resumptives in island contexts present for a model of sentence production using TAG. Section 3 argues that the problem can be resolved given the existence of resumptive pronouns in English in non-island contexts. Section 4 discusses two potential diffiulties for the analysis. Section 5 outlines some possible implications of the analysis for the typology of relative clauses in general. Finally, Section 6 concludes with suggestions for further investigation.

2. The problem of "remedying" island violations
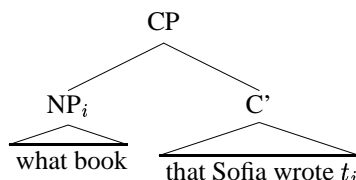
2.1. Accounting for island violations in TAG

Frank (2002) presents a conception of grammar applying the machinery of tree adjoining grammar within an independent theory of well-formedness of grammatical structures. The domains over which any structural dependencies are expressed are the elementary trees in the system. The TAG operations of substitution and adjunction can then apply to any of the set of well-formed trees. No movement transformations outside the domain of the elementary tree are possible in this system.

In order to account for the well-known phenomena of island violations, rather than stating principles constraining wh-movement (or the location of traces of this movement), Frank's theory instead rules out the elementary and auxiliary trees needed to derive structures with island-violations primarily based on two independent principles: 1) the Theta Criterion: only nodes that are selected for can appear in an elementary tree; and 2) in order to obey the requirement that derivation trees should be context-free, a well-formed derived auxiliary tree rooted in X cannot be derived via substitution of a tree with an X foot node into a tree with an X root node.

For example, extraction from an NP complement, as in (1), is ruled out because in addition to the legitimate elementary tree rooted in *wrote* in (2), it would require the impossible auxiliary tree in (3).
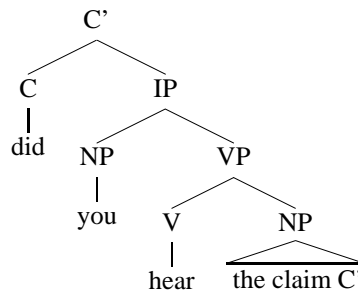
(1)   *What book$_j$ did you hear [the claim that Sofia wrote t$_j$]?

(2)

---

(3)

```
                              C'
                         ____/\____
                        C          IP
                        |        __/\__
                       did     NP      VP
                               |      _/\_
                              you    V    NP
                                     |   /\
                                   hear  the claim C'
```

The latter is ill-formed because neither of the trees that make it up—an elementary tree rooted in C' and anchored by *hear* and an elementary DP tree anchored by *claim* taking a C' complement—are themselves a C' auxiliary tree. Extraction from wh-islands, relative clauses, adverbial adjuncts, and sentential or DP subjects are ruled out in similar ways.

2.2. Remedying island violations with resumptive pronouns

In unplanned speech speakers of English sometimes produce utterances with a resumptive pronoun in place of the gap or trace which would be an island-violation if a wh-word had been extracted from that position, as shown in the naturally-occuring examples of extraction from a wh-island, (4a), a relative clause, (4b), an adverbial clause, (4c), and a subject in (4d).

(4)   a. There are always guests who I am curious about what **they** are going to say. (Prince (1990)'s 3a)

   b. That asshole X, who I loathe and despise the ground **he** walks on, pointed out that...(Prince (1990)'s 5a)

   c. Apparently, there are such things as bees in the area which if you're stung by **them**, you die. (Prince (1990)'s 5b)

   d. You have the top 20% that are just doing incredible service, and then you have the group in the middle that a high percentage of **those** are giving you a good day's work... (http://www.ssa.gov/history/WEIKEL.html)

Kroch (1981) argues that such resumptive pronouns are the result of a processing effect. Although a dependency between a wh-word and a gap in these positions is ungrammatical, as an artifact of how sentence production proceeds, forms like those above are uttered. Kroch explicitly rejects any possibility of a formal solution within the grammar to account for these forms. He uses an incremental model of sentence production in which a wh-element is produced in a fronted position before the entire sentence has been planned out, allowing the utterance of fluent speech. Rather than having to wait to create the entire sentence in which the distance between the wh-element and its gap is potentially unbounded, the speaker can begin speaking while the remainder of the structure is being constructed. This wh-element is adjoined to a full clause where it appears again in its base-generated position. When the repeat occurrence is reached in the course of production, it is not pronounced. If the sentence produced has an island-violation, then a pronoun or NP is inserted in the base-generated position. In this model, resumptive pronouns are expressly inserted as a "last resort" to avoid violating the ECP by leaving an ungoverned trace. The resumptive pronoun remedies the violation because an empty category will no longer be present.

2.3. Sentence production using TAG

In a TAG-based model of incremental sentence production (Ferreira, 2000; Frank and Badecker, 2001), elementary trees are the basic "building blocks" of generation. Selecting a lexical item corresponds to selecting an elementary tree. The grammatical properties of that tree's structure are a natural domain for explaining the planning commitments a speaker makes while producing a sentence. The grammar provides a set of trees to work with, and as sentence planning proceeds, the speaker attempts to select the correct trees to encode the meaning she wishes to express and to put them together in a particular configuration. The only source of production errors is incorrectly combining trees; ill-formed elementary (and auxiliary) trees never arise. To produce a structure with a wh-dependency, the speaker selects the elementary tree headed by the main verb with the appropriate wh-element

fronted. In a structure with a long-distance dependency, additional material can be adjoined in between the wh-element and the remainder of the clause it originated in, but the commitment to the main clause structure has already been made.[1]

In this model of production where we assume that a speaker only has grammatical resources with which to work, we can not use Kroch (1981)'s explanation of the appearance of resumptive pronouns in island-violation contexts. The resources needed to produce island-violating structures are not available in the grammar that licenses the set of tree building blocks. On the face of it then, it seems that the existence of resumptive pronouns in island violating contexts would prove devastating for this model of sentence production. Based on the assumptions that 1) the processing system has only grammatically-licensed trees with which to create larger structures and 2) the structures needed to extract from island-violation contexts are not grammatically-licensed, speakers could not be remedying violations that should not even be created given their underlying grammars. As we will argue in the following section, however, the underlying grammar of English speakers independently requires the resources needed to produce these forms. Hence, we can preserve both the assumptions above and still have a grammar that characterizes all the structures that English speakers use.

3. Resumptive pronouns as grammatical resource
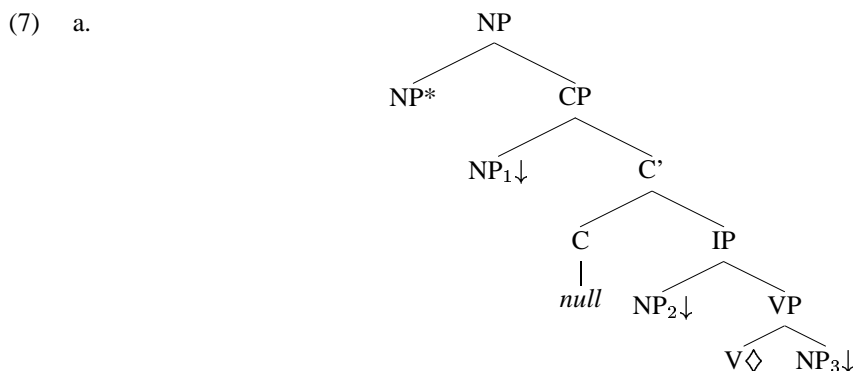
3.1. Resumptives in non-island contexts

Resumptive pronouns appear in relative clauses in English in *non*-island violation contexts, as in (5), from Prince (1990).

(5)   a. My son, God bless him, he married this girl which I like **her**. (Prince's 28a)

      b. If there's any message that she can forward **it** to us, then...(Prince's 15b)

      c. You get a rack that the bike will sit on **it**. (Prince's 15c)

      d. I have a friend who **she** does all the platters. (Prince's 4c)

Prince presents an analysis of the function of this type of resumptive pronoun, claiming that they serve as normal discourse pronouns rather than as a type of bound pronoun or trace. These pronouns appear in contexts where the relative clause is being used simply to predicate additional information about a discourse entity evoked by the head noun, not to assist the hearer's identification of the referent evoked by the head noun. For example, they are far more common with non-restrictive relatives and relatives modifying indefinite NPs. Additional evidence she presents for the "discourse pronoun" analysis are cases where a co-referential demonstrative or full NP or even a non-coreferential (but related) NP appears instead of a resumptive pronoun, as in (6) (Prince's 34(a-d)).

(6)   a. I had a handout and notes from her talk that **that** was lost too.

      b. He's got this lifelong friend who he takes money from the parish to give to **this lifelong friend.**

      c. I have a manager, Joe Scandolo, who **we**'ve been together over twenty years.

      d. You assigned me to a paper which I don't know anything about **the subject**.

In order to produce relative clauses like these, speakers must be using structures like those in (7).

(7)   a.

---

b.

```
                                    NP
                          ┌─────────┴─────────┐
                        NP*                   CP
                                        ┌──────┴──────┐
                                      NP₁            C'
                                       │        ┌─────┴─────┐
                                       e        C          IP
                                                │      ┌────┴────┐
                                              that   NP₂↓        VP
                                                            ┌─────┴─────┐
                                                           V◊         NP₃↓
```
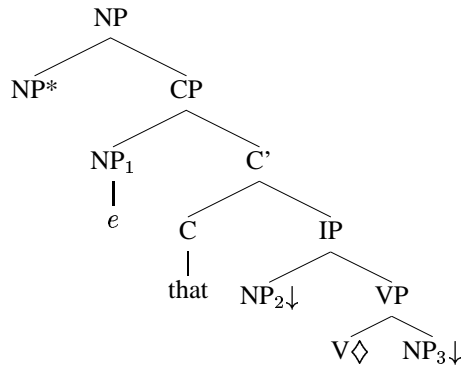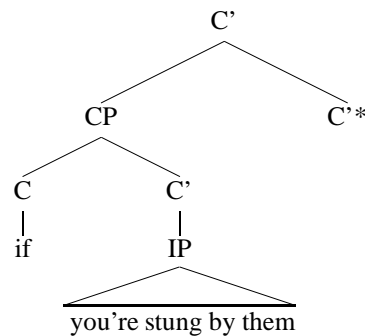
Here NP₁ would have features requiring the substituting NP to be a wh-word. But NP₂ and NP₃ could have any NP substituted within them, including a pronoun coreferential with the NP to which this tree adjoins.

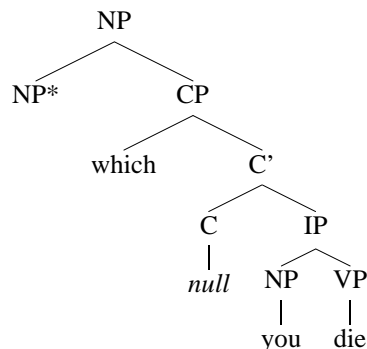### 3.2. Generating resumptives in island contexts

Because the production system necessarily allows relative clause auxiliary trees like those in (7), we can now explain where speakers find the grammatical resources to produce relative clauses with island-violating resumptive pronouns. The trees in (7) are projections of a head verb. As such any finite verb can project this (or a related) structure. Any legitimate auxiliary tree can be adjoined into it. There is no syntactic dependency between the relative pronoun and any of the lower NPs. Therefore, there needs to be no tree at any point in the derivation that reflects such a local dependency. For example, in order to derive (8a), we now only need a relative clause auxiliary tree with *die* as its head, into which we substitute *which* and *you*, as in (9b).

(8)   a. bees which if you're stung by them, you die

      b. * bees which$_j$ if you're stung by t$_j$, you die

(9)   a.

```
                                    C'
                          ┌─────────┴─────────┐
                         CP                   C'*
                   ┌──────┴──────┐
                   C            C'
                   │             │
                   if           IP
                           ┌─────┴─────┐
                        you're stung by them
```

b.

```
                                    NP
                          ┌─────────┴─────────┐
                        NP*                   CP
                                        ┌──────┴──────┐
                                      which           C'
                                                ┌──────┴──────┐
                                                C            IP
                                                │       ┌─────┴─────┐
                                              null     NP          VP
                                                        │           │
                                                       you         die
```
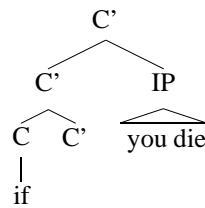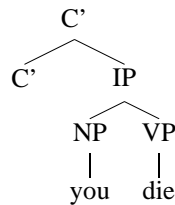
An initial tree headed by *stung* can then be substituted into a tree headed by *if*; the resulting tree, (9a), can be adjoined into the *die* tree, (9b). Neither of the suspect trees which would be required in some derivation of the unacceptable NP (8b), e.g. those in (10), are needed to generate the resumptive pronoun version in (8a).

(10)   a.

```
                    C'
                  /    \
               C'        IP
             /   \      /  \
            C    C'    you die
            |
            if
```

       b.

```
                    C'
                  /    \
               C'        IP
                        /  \
                      NP    VP
                      |      |
                     you    die
```

As a result, the impossibility of generating island-violations without resumptive pronouns is preserved in this analysis.

4. Potential problems

In the previous sections we saw that given the elementary trees needed to generate resumptive pronouns in relative clauses without island-violations, the constraints preventing the generation of island violations need not be altered in order to correctly allow a TAG of English that generates resumptive pronouns in otherwise-island-violating contexts.

This section will discuss two potential difficulties for the syntactic analysis of resumptives in island contexts given above. The first is to explain the apparent overgeneration of such an analysis. The second is to explain why resumptives in island and non-island contexts in English are often characterized as marginally acceptable if in fact the grammar generates them.

4.1. Overgeneration

The relative clause trees given in (7) would overgenerate by allowing non-gap-containing relative clauses, like (11–12):

(11)   the police officer who John prefers spinach

(12)   the smell that my mom is baking bread

Relatives like (11) could be ruled out with the pragmatic requirement that relative clauses must make a statement about their head NP (Kuno, 1976), in that there is no easily apparent relation between the head and the relative clause. The oddness in English of example (12) is more difficult to explain.

These forms are far more common in topic-oriented languages like Korean and Japanese, as illustrated by the perfectly acceptable relative clause in (13).

(13)   emeni-ka   ppang-ul   kwup-nun naymsay
       mom-Nom bread-Acc bake-Rel   smell

       'smell that my mom is baking bread' (Na-Rae Han, p.c.)

The pattern here could be explained by a cross-linguistic difference in the pragmatics of structures where a NP is adjoined to a clause, main or relative, without a gap. In English, these structures are possible, as shown in (6d) above, but the relation between the NP and clause is subject to strict pragmatic restrictions. In a language like Korean, however, the discourse conditions such constructions are subject to are less restrictive.[2]

---

2.   This difference in the pragmatic factors conditioning identical or similar syntactic forms is not unknown in the literature. See (Prince, 1999) for a discussion of the different restrictions on the conditions allowing the formally identical constructions of Focus Movement in Standard English and Yiddish-Movement in Yiddish English.

4.2. Marginal acceptability

One issue that remains to be explained in this analysis is the following: if these sentences are created using legitimate parts of the grammar, why then are they regarded as being only marginally acceptable? I will suggest three reasons here; two of which apply to both island- and non-island resumptive pronoun uses.

First, resumptive pronouns in both island and non-island contexts are primarily confined to informal, spoken registers in English. In addition, for reasons to be discussed below, they are infrequent forms. Both of these factors are reasons linguists might be inclined to classify them as only "marginally acceptable." [3]

Secondly, relative clauses are supposed to make a statement "about" their head noun (Kuno, 1976). This functional requirement might make processing island-violating relative clauses difficult for hearers because the resumptive NP appears in a position unlikely to be the theme of the clause; that is, in general, sentences are unlikely to be "about" an NP embedded in an adjunct clause, a relative clause, or other island context.

One final possible reason for the marginal acceptability of resumptives in relative clauses applies to both island- and non-island violating contexts. Given the predominance of gap-containing relative clauses in English, hearers expect to find a gap in a relative clause. This gap allows the semantic relation between the extracted NP and the relative clause to be computed directly from the syntax. When a resumptive pronoun (or other NP) appears in such a slot instead, additional non-syntactic processing is required to compute this relation. This explains why native speakers often judge resumptive pronouns in non-island violating contexts as unacceptable as those in island-violating contexts (although they *produce* both.)

Given the statistical preference in English for gap-containing relatives, for English speakers constructions with resumptive pronouns in both island and non-island contexts could be explained as a result of poor planning commitments during the production process. Under this line of explanation, we would hypothesize that at the point at which they begin selecting trees to encode meaning, speakers have already committed to combining more information into a single utterance than they, under ideal circumstances, would wish to express in just a main clause and a relative clause. Rather than using two separate main clauses to make two separate predications, they are "stuck" with using one main clause and a relative clause tree that is essentially a main clause adjoined in as a modifier of an NP.

5. Implications for the typology of relative clauses

As discussed extensively in the literature, the syntactic strategies for relativizing within a given language are closely tied to the Accessibility Hierarchy (AH). In general, the lower on the AH an NP's syntactic role is, the less likely a language is to use a gap rather than a resumptive when relativizing that NP (Kuno, 1976; Keenan and Comrie, 1977; Prince, 1990). In addition, for any position on the AH, if it can be relativized using a gap, all positions higher than it can too.

Three questions that this analysis might be expected to shed light on with respect to the strategies for relativization will be examined:

- Why are NPs low on the Accessibility Hierarchy harder to relativize with gaps?
- Are gaps always easier to process than resumptives?
- Are resumptives always interpreted as discourse pronouns (and gaps as bound)?

5.1. Relativization strategy and the Accessibility Hierarchy

The syntactic analysis given here predicts only that in non-island contexts either gaps or resumptives should be possible, and in island contexts only the latter are possible. However, in English, relativization of NPs lower on the AH may be more likely to appear as either gaps or resumptives in relatives, as in the datives in (14) [=Prince (1990)'s (7)a-c].

(14)   a. ...the man who this made **him** feel sad...

     b. Some of the same judges who we told **them** that if you mess with John Africa...

     c. He looks like one of those guys you got to be careful throwing **them** fastballs

---

3.   Although it is an important question, whether such a classification is meaningful will not be further discussed here. Instead, we will simply take this sub-acceptability as a given and explore further reasons for it.

Besides the thematic claim of Kuno (1976) mentioned above, we do not know why the AH is related to relativization. We can only speculate why producing a relative clause where the relativized NP is "unthematic" and gapped is difficult. With respect to production models based in TAG, taking the AH into account in relativization would possibly require the activation of multiple trees in cases where either resumptives or gaps are possible.[4] The lower the NP role is on the AH, the more equally the trees might be weighted probabilistically allowing either to be produced with equal likelihood.

5.2. Relative processing difficulty of gaps and resumptives

English is unusual in its ability to relativize from even rather deeply embedded positions (Keenan, 1985). As discussed above in Section 4.2 the comprehension system of English is likely affected by the relative frequency of gaps vs. resumptive pronouns. Because gap-containing forms are the more frequent, and hence less "marked", form, they should be easier to comprehend. In addition, they should be the preferred form to produce. In other languages where the break between gaps and resumptives falls at a different position on the Accessibility Hierarchy, and so the frequency of resumptive forms is much greater, the relative ease of processing resumptive relative clauses in both production and comprehension may very well differ from that in English.

5.3. Interpretations of gaps and resumptives

In English, resumptive pronouns in non-island contexts have a specific discourse function, in that they serve as discourse pronouns. This appears to correspond to the syntactic analysis given here, because in resumptive relative clauses, any NPs within the relative clause are unbound by the relative pronoun, while in the case of a gap/trace, there is syntactic binding. However, zero pronouns can function as discourse pronouns in numerous languages, and phonologically explicit pronouns can function as bound pronouns in many languages, including English. Therefore, the syntactic analysis here does not necessarily have implications for the semantic/pragmatic interpretation of the two strategies of relativization examined here. Any typological claims would need to be based on corpora of naturally-occurring examples, in order to take into account the properties of their contexts.

6. Conclusion and areas for further investigation

This paper has presented evidence to resolve the paradox that resumptive pronouns in island-violating contexts would otherwise present for TAG-based model of sentence production. I have argued that given the grammatical resources of English, specifically resumptive pronouns in non-island violating relative clauses, resumptive pronouns in island-violating contexts are part of the grammatical competence of speakers of English. Their marginal acceptability is most likely due to frequency factors affecting their processing.

In order to confirm whether these forms are actually the result of "poor planning," psycholinguistic experimentation is needed. A focus in such experimentation on how and why speakers decide to split up information they encode into units (a main clause with attached subordinate vs. two main clauses) would be particularly useful.

An additional linguistic area of exploration would be whether the syntactic analysis here of resumptives in island vs. non-island relatives could be extended to their role in other related structures, like wh-questions and topicalization. Preliminary evidence seems to support an affirmative answer for topicalization structures, as shown by the distribution in (15-18).

(15)   [Most of those people]$_i$ I never met $\emptyset_i$. (SSA)[5]

(16)   And [those that hadn't]$_i$, I assumed **they**$_i$ were interested in basketball or football, so it was not difficult to figure out a way to get to see them. (SSA)

(17)   But [the field office claims of the Claims Manual]$_i$, you tended to learn pretty much all that was in **it**$_i$ as a Claims Rep...(SSA)

(18)   (Being Unix, it comes with Gnu Emacs out of the box,)
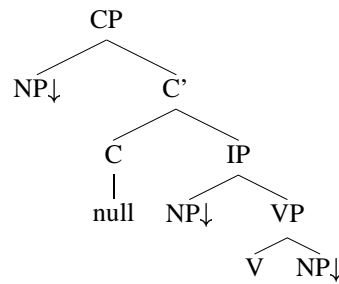   Latex there are also a few options. (Email: T.M.)

Here, we have a coreferential gap in (15), a coreferential pronoun in (16), a coreferential pronoun in an NP-complement island in (17), and finally one with neither a coreferential NP nor a gap in (18). On first glance then, it

---

4.   See Ferreira (2000) for a discussion of the simultaneous activation of multiple trees in sentence production.
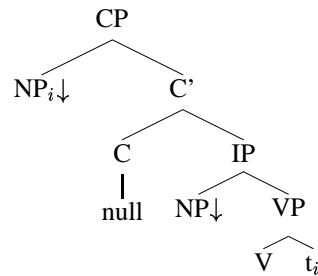5.   The SSA marked examples are from transcribed oral histories at http://www.ssa.gov/history/orallist.html.

appears that speakers do have the option of both of the following trees for extraction from full clauses on analogy with the options they have for relatives clauses:

(19)

```
                    CP
                  /    \
              NP↓       C'
                      /    \
                    C       IP
                    |      /   \
                  null  NP↓    VP
                             /    \
                            V     NP↓
```

(20)

```
                    CP
                  /    \
              NPᵢ↓      C'
                      /    \
                    C       IP
                    |      /   \
                  null  NP↓    VP
                             /    \
                            V     tᵢ
```

References

Ferreira, Fernanda. 2000. Syntax in Language Production: An Approach Using Tree-adjoining Grammars. In L. Wheeldon, editor, *Aspects of Language Production*. MIT Press, Cambridge, MA.

Frank, Robert. 2002. *Phrase structure composition and syntactic dependencies.* Cambridge, MA: MIT Press.

Frank, Robert and William Badecker. 2001. Modeling syntactic encoding with incremental tree-adjoining grammar. In *Proceedings of the 14th Annual CUNY Conference on Human Sentence Processing*, University of Pennsylvania.

Keenan, Edward. 1985. Relative Clauses. In Timothy Shopen, editor, *Language Typology and Syntactic Description, Vol. II.* Cambridge University Press.

Keenan, Edward and Bernard Comrie. 1977. Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry*, 8(1):63–99.

Kroch, Anthony. 1981. On the role of resumptive pronouns in amnestying island constraint violations. In *Papers from the 17th regional meeting of the Chicago Linguistic Society*, pages 125–35, Chicago. Chicago Linguistic Society.

Kuno, Susumo. 1976. Subject, theme and speaker's empathy: A reexamination of relativization phenomena. In Charles Li, editor, *Subject and topic*. Academic Press, New York, pages 417–44.

Prince, E. F. 1999. How not to mark topics: 'Topicalization' in English and Yiddish. University of Pennsylvania.

Prince, Ellen. 1990. Syntax and discourse: a look at resumptive pronouns. In Kira Hall, Jean-Pierre Koenig, Michael Meacham, Sondra Reinman and Laurel Sutton, editors, *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley Linguistics Society, Berkeley, CA, pages 482–97.