

Tense, Events, and Extraction from Adjuncts*

Robert Truswell
University College London

1 Introduction

A common assumption in syntactic theory is that extraction from adjuncts is impossible. This position has its roots in the definition of *island* in Ross' seminal (1967) dissertation, and has remained largely unchallenged in subsequent stages in the development of Chomskyan syntactic theory, to such an extent that a recent paper, Stepanov (2007), makes the claim that 'extraction out of adjuncts seems to be banned universally' (Stepanov 2007:81).

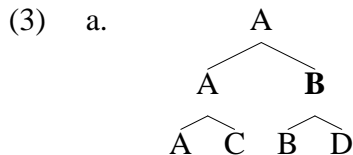
Over time, this assumption has come to be increasingly integrated into the syntactic architecture. For example, the Condition on Extraction Domain of Huang (1982), repeated in (1), represents a milestone in the development of locality theory, by giving us a unified account of the islandhood of subjects and adjuncts. As neither subjects (at least, VP-external subjects) nor adjuncts are properly governed, the CED predicts extraction from both to be impossible, and for the same reason.

- (1) A phrase A may be extracted out of a domain B only if B is properly governed. (Huang 1982:505)

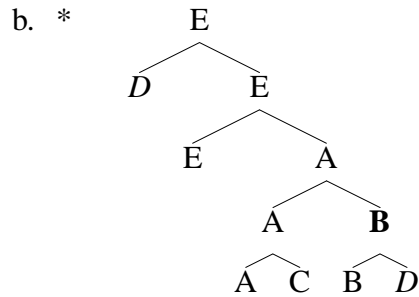
In recent years, the group of theories which I will collectively call the *Multiple Spell-Out* approach (see in particular Uriagereka 1999 and Johnson 2002) have taken this a step further, and integrated CED effects into a theory of the basic minimalist structure building operation, Merge. Although each theory implements the basic idea in a subtly different way, these approaches all derive the following theorem, where both subjects and adjuncts are assumed not to project and to have phrasal sisters.

- (2) A non-projecting phrase is a strong island if its sister is a phrase.

Pictorially, this means that in the tree in (3a) (where linear order should be taken to be irrelevant), the higher node labelled B is an island, and so a dependency formed across it, such as movement of D in (3b), is ungrammatical. This, however, is precisely the configuration which extraction out of any adjunct would instantiate.



*Thanks to Ad Neeleman for helping me develop these ideas, and to all at CLS 43, particularly Philip Hofmeister, Robert Kluender, Jeff Runner, and Ivan Sag, for very helpful discussion. This work was supported by a Wingate Scholarship.



Of course, the further we move along this path of integration, the more important it is that the basic empirical assumptions are correct. If, according to the Multiple Spell-Out approach, the impossibility of extracting out of an adjunct falls out as a theorem from properties of the basic structure-building operation, then it is imperative that we really can't extract out of an adjunct.

It seems, though, that in certain circumstances, extraction out of an adjunct is quite possible. Such examples were first noted, to my knowledge, in Chomsky (1982), who gives examples such as those in (4), and they have since been discussed, notably, in Cinque (1990) and Levine & Sag (2003).

- (4) a. The man that I went to England [without speaking to *e*]. . .
 b. Here is the influential professor that John went to college [in order to impress *e*]. (Chomsky 1982:72)

The existence of such examples is sufficient to make us reconsider the universality of the claims lying behind the CED and the Multiple Spell-Out approach. However, simply removing such conditions from our theory of grammar is somewhat unsatisfying. For the majority of speakers, most examples involving extraction out of adjuncts are unacceptable, and any theory must include something which explains this fact. A typical sample is given in (5).

- (5) a. *The man that I went to England [after I spoke to *e*]. . .
 b. *Here is the influential professor that John went to college [because he impressed *e*].

A pressing empirical question therefore concerns the distribution of legitimate extractions out of adjuncts. What exactly distinguishes the acceptable examples from the unacceptable examples? In fact, one apparently empirically robust generalisation has already been proposed in this regard, by Cinque (1990) (see also Szabolcsi 2006). This is that, while extraction is possible from certain untensed adjuncts, it is universally impossible from tensed adjuncts. This, then, would derive the difference between the acceptable examples in (4) and the unacceptable examples in (5): only the cases in (5) involve extraction out of a tensed adjunct and are automatically prohibited as a consequence.¹

¹An alternative hypothesis would be that the distinction between (4) and (5) is due to the presence of a subject in the latter cases only. Although this is a plausible alternative, I find it less likely overall because of the existence of a few relatively acceptable cases of extraction out of an untensed adjunct which includes a subject, as in (i).

This paper will take Cinque's generalisation to be correct, and explore a further question which emerges in its light, namely *why* tensed adjuncts should have this property. At a very broad level, there are two major approaches which may be taken. On the one hand, we could stipulate that there is something syntactically special about a T head or about TP, which somehow interacts with adjunction to induce the island effects seen in (5). Alternatively, we might claim that the semantic impact of tense somehow induces those effects.

At first sight, neither of these approaches seem very appealing. If we take the syntactic tack, then we must explain why T, or TP, is singled out for special treatment. However, there is a long tradition of saying that T(P) is somehow different syntactically: the Extended Projection Principle of Chomsky (1981) applied exclusively to [Spec,T] (in more modern terms), for example, and TP was a stipulated exception to the general algorithm for determining barriers to movement in Chomsky (1986), so a syntactic approach to the ban on extraction from tensed adjuncts, although perhaps a little suspicious, would not be entirely without precedent in this respect.

Moreover, both approaches face the challenge of explaining why tense, or T(P), interacts only with adjuncts to produce these island effects, as extraction from a tensed complement is, of course, often possible.

(6) What did John say [that Bill did *e*]?

In fact, space constraints will prevent me from sketching more than the barest of analyses of this latter discrepancy here (see Truswell 2007b for a fuller account). However, with respect to the former problem, we will see that there is good reason to favour a semantic account over a purely syntactic account. The argument will be based on the assumption that, all else being equal, we would prefer the condition ruling out (5) to do more empirical work as well. Section 2 lays out a further generalisation concerning extraction from *untensed* adjuncts. A particular class of untensed adjuncts are shown to only allow extraction if the matrix VP has certain aspectual characteristics. Section 3 summarises an event-structural account of this phenomenon initially proposed in Truswell (2007a), while remaining agnostic about any possible syntactic implementation thereof. Finally, in section 4, we return to Cinque's generalisation concerning the lack of extraction from tensed adjuncts. We see that the semantic account developed in section 3 automatically derives this generalisation, offering a unified theory of two distinct facts, while any syntactic account along the same lines will retain an essentially disjunctive character. As argued in the conclusion, this suggests that the semantic account is to be preferred.

(i) ?This is the problem that we went on holiday [without anyone mentioning *e*].

Although such examples are admittedly far from perfect, some native speakers judge them to be more acceptable than the cases in (5), which suggests that the presence of a subject is not the major factor behind the degradation of extraction out of tensed adjuncts.

2 Aspectual Class and Extraction from Bare Present Participial Adjuncts

It is at least conceivable that either a syntactic or a semantic story could capture the contrast between extraction out of an untensed adjunct, which is sometimes possible, as in (4), and extraction out of a tensed adjunct, as in (5), which is universally impossible. However, if we look at patterns of extractability from untensed adjuncts alone, we find cases of legitimate extraction which look like they are distributed along primarily semantic lines. I give one example in this section, concerning extraction from a class of constructions that I will call *Bare Present Participial Adjuncts* (BPPAs), where the adjunct is built around a present participial verb, and is not introduced by a preposition or other phrase such as *in order*.

Acceptable examples of extraction from such adjuncts follow a curious distribution. (7) gives a few unacceptable examples.

- (7) a. *What does John work [whistling *e*]?
b. *What did John write the cheque [complaining about *e*]?
c. *What did John see the spectacle [looking through *e*]?

These contrast minimally with examples such as (8), where extraction from the adjunct is readily possible, despite the fact that these adjuncts are identical in internal structure to those in (7a–b).

- (8) a. What did John arrive [whistling *e*]?
b. What did John drive Mary crazy [complaining about *e*]?

In fact, it seems that the extraction possibilities in these cases are determined by aspectual properties of the matrix predicate. Following Vendler (1957), we may divide the set of verb phrases in English into four broad classes, the *accomplishments*, *achievements*, *activities* and *states*, on the basis of two intersecting families of diagnostic tests. For example, accomplishments and activities readily allow use of the progressive, as shown in (9a–b), whereas this is at least degraded for achievements and states, as in (9c–d).

- (9) a. I am running a mile (drawing a circle, building a house, . . .).
b. I am running (writing, working, . . .).
c. *I am spotting the plane (appearing, blinking, . . .).
d. *I am knowing the answer (loving you, understanding antisymmetry, . . .).

On the other hand, accomplishments and achievements reject *wh*-questions of the form *for how long*. . . , as in (10a) and (10c), while these are quite acceptable with activities (10b) and states (10d).

- (10) a. #For how long did he run a mile (draw a circle, build a house, . . .)?
b. For how long did he run (write, work, . . .)?
c. #For how long did you spot the plane (appear, blink, . . .)?
d. For how long did you know the answer (love me, understand antisymmetry, . . .)?

We can make sense of these results by postulating a decompositional model of event structure, in the manner first formalised by Dowty (1979), and with its ultimate roots in works in the generative semantics tradition such as Lakoff (1970). I will assume that a verb phrase can describe an event which consists of two *subevents*: a *preparatory process*, which takes place over a temporally extended period and during which no linguistically significant change of state occurs, and which immediately precedes and directly causes a *culmination*, a pointlike instant of linguistically significant change. We can then take the ability to form a progressive as diagnostic of the presence of a preparatory process, and the ability to form a *for how long* question as diagnostic of the absence of a culmination. This leads to a first-pass decompositional model of the aspectual classes as follows.

- (11) Internal structure of the aspectual classes (first pass):
- a. **Accomplishment:** ⟨Process, culmination⟩
 - b. **Achievement:** Culmination
 - c. **Activity:** Process
 - d. **State:** \emptyset

However, things are more complicated than this, most pertinently in the case of achievements. Contrary to a simple-minded interpretation of Vendler's tests, many predicates which have come to be classed as canonical achievements are quite acceptable in the progressive, as in (12), while many others are much more resistant, as in (13). I will refer to the former as *true achievements* and the latter as *points* in what follows.

- (12) a. I'm reaching the summit as we speak.
 b. John's dying.
- (13) a. #I'm noticing the problem as we speak.
 b. #John's blinking.

The examples in (13) are rescuable, but only on particular interpretations of the event described by the verb. (13a) is perhaps most felicitous if the noticing event is taking place in slow motion, such that *notice* means something like *begin to see*. Equally, (13b) is quite acceptable on an iterated interpretation of blinking. However, these interpretations are quite distinct from the natural interpretations of (12), which have a prospective quality to them, such that (12b), for example, is very similar in meaning to *John will soon die*.

This latter interpretation is very close to the interpretation assigned to a regular progressive of an accomplishment, as in (9a).² *I am running a mile*, for example,

²The two are, however, far from identical. One major difference is that use of a canonical accomplishment such as *run a mile* entails, or at least strongly implies, that the subject is currently performing an activity, while no such inference is possible in the case of an achievement such as *die*. A second difference is that the culmination described by an achievement in the progressive must be much more imminent than is the case with an accomplishment. So it is legitimate to say any of the examples in (9a) as soon as you have started to engage in any activity leading to the culmination, but you must be very close to the top of the mountain before you can claim to be *reaching the summit*. Truswell (2007b) provides an analysis of these contrasts based on a difference in agentivity, but the details take up too much space to repeat them here.

means something like *I am engaged in an activity which will lead to me having run a mile*. The progressive forms of both the true achievements and the accomplishments therefore have a similar “inertial” quality, in the sense of the *inertia worlds* of Dowty (1979), which is absent from the progressive forms of points, which rather rely on coercion into activity-like interpretations: progressives of both accomplishments and true achievements imply that, all else being equal, a specified culmination will be reached in the future.

Returning to the examples in (7–8), then, we see that the five cases fall into four different classes. The unacceptable cases in (7) consist of an adjunct predicate describing an activity, and modifying an activity, an accomplishment, and a point, respectively. Meanwhile, the acceptable cases in (8) consist of an activity-describing adjunct modifying a true achievement, and an accomplishment, respectively.³ This hints at a connection between the internal structure of an event-describing constituent and the subextraction possibilities, but several questions spring immediately to mind. For instance, we would like to know why the aspectual classes distribute themselves in the way that they do with respect to the extraction possibilities. Secondly, the split in the behaviour of BPPAs modifying accomplishments remains puzzling: (7b) shows a degraded case of extraction from such a construction, while (8b) shows an apparently parallel example, which is nonetheless acceptable. We need a way of distinguishing these two cases. Finally, and in the light of the question raised in the introduction, we need to know whether, when we talk about “internal structure”, we are talking in purely semantic terms, or whether it is possible, or desirable, to formulate the relevant condition in terms of a syntacticised decompositional theory of event structure (Lakoff 1970; Hale & Keyser 1993). These questions will be addressed in the following sections.

3 Locality and Event Structure

The hints described in the previous section lead one naturally to consider the possibility of an event-based locality condition regulating the distribution of movement relations. I will now give one formulation of such a condition, which I will call the Single Event Condition, and then show how, in conjunction with the decompositional theory of event structure sketched in section 2, it can account for the data in

³Although space considerations force me to limit myself to these data in this paper, it should be pointed out that extraction from a BPPA modifying an activity VP is occasionally possible, as in (i). See Truswell (2007b) for an modification of this theory which covers this limited class of exceptions.

- (i) What did John lie in bed [reading *e*] all day?

Another minor qualification: I have omitted any discussion of the behaviour of BPPAs modifying states. This is because such constructions are generally quite degraded even in declaratives — (ii), due to Annabel Cormack (p.c.), is the most natural example of which I am aware. As (iii) shows, if anything, extraction from the adjunct degrades it even further.

- (ii) ?I only know Georgian [wearing THIS magic hat].
(iii) *Which magic hat do you know Georgian [wearing *e*]?

(7–8), before returning in section 4 to the issues surrounding tensed adjuncts and the division of labour between syntax and semantics.

The condition is the following.

(14) **The Single Event Condition: Events form locality domains for *wh*-movement.**

A *wh*-chain is legitimate only if the minimal constituent containing the head and foot of the chain asserts the existence (in the actual world) of a single event.

The interest of this comes from the fact that a decompositional theory of event structure opens up the possibility that multiple subevents can jointly compose a single event (which I will call a *macroevent*). We saw in the previous section, for example, that an accomplishment predicate can be decomposed into process and culmination subevents. However, an accomplishment such as running a mile continues to feel like a single event. An event, then, can have proper subparts which are also events, as first explicitly proposed by Bach (1986). All that (14) cares about is that the constituent in question describes a single (macro)event: whether that event is composed of subevents or not is irrelevant to that condition.

Of course, the Single Event Condition would not be much use if the only constituents which could describe single events were single verbal projections or clauses: extraction is possible in many more environments than these. The rest of this paper will sketch some ways in which (14) may apply beyond this baseline case. There will be a few comments in section 4 about the interaction of this condition with clausal complementation constructions, but before that, I want to sketch an approach, initially proposed in Truswell (2007a), to the untensed adjunct subextraction data from section 2.

In all the cases of extraction from BPPAs documented in section 2, the adjunct has described an activity, such as whistling or complaining. As we have classed activities as processes without culminations, we do not expect them to have any subevent structure: an activity predicate is a simple predicate of an event variable, as in (15).

$$(15) \quad \llbracket \text{whistle} \rrbracket = \lambda x \lambda e \lambda y. \text{whistle}(y, x, e)$$

However, there is substantially more variation to be found in the subevent structure of the matrix predicates. I will firstly describe the range of cases in (7), where extraction from an adjunct was shown to be degraded, before contrasting them with the cases in (8).

Work, from (7a), is also an activity predicate, and consequently has a similar subevent structure to that of *whistle*. In contrast, *see the spectacle*, a point predicate, consists of a culmination without a process. However, from the perspective of the Single Event Condition, this nonetheless means that it has a single event variable with no relevant subevent structure. From the point of view of that condition, then, the two predicates behave identically.

- (16) a. $\llbracket \text{work} \rrbracket = \lambda e \lambda x. \text{work}(x, e)$
b. $\llbracket \text{see} \rrbracket = \lambda x \lambda e \lambda y. \text{see}(x, y, e)$

The case of *write the cheque* (7b) is slightly more complex: here we have a typical accomplishment predicate, and therefore we must represent the decomposition of the macroevent into process and culmination subevents.⁴ Moreover, there is descriptive content attached to both the process and culmination subevents in this case: the predicate consists of a culmination of the cheque being filled in which results from a writing process.

$$(17) \quad \llbracket \text{write the cheque} \rrbracket = \lambda e_1 \lambda E \lambda x \exists e_2. \text{write}(x, e_1) \wedge \text{filled_in}(\text{cheque}, e_2) \\ \wedge E = \langle e_1, e_2 \rangle$$

Now, compare this with the representations of the matrix predicates in the two examples in (8), where extraction from the adjunct *is* possible. Although these are similar to (17) in that, as an accomplishment (*drive Mary crazy*) and a true achievement (*arrive*), they contain two distinct subevents, they differ from any of the examples in (7) in that there is no descriptive content attached to the process subevent. In other words, the truth of *John drives Mary crazy* does not depend on what John does to drive Mary crazy, so long as Mary ends up crazy as a result of John's behaviour. Equally, the truth of *John arrives* is independent of the manner of John's arrival. This means that the representation of such VPs will have a λ -abstracted subevent variable, e_1 , which is not the argument of any descriptive predicate, as in (18).

$$(18) \quad \text{a. } \llbracket \text{drive Mary crazy} \rrbracket = \lambda e_1 \lambda E \lambda x \exists e_2. \text{crazy}(m, e_2) \wedge E = \langle e_1, e_2 \rangle \\ \text{b. } \llbracket \text{arrive at Y} \rrbracket = \lambda e_1 \lambda E \lambda x \exists e_2. \text{at}(x, Y, e_2) \wedge E = \langle e_1, e_2 \rangle$$

We now come to the integration of semantics of the matrix VP and the BPPA. I will assume that there are two options in this respect, both in principle freely available but constrained by an interaction of factors to be described below. On the one hand, as is commonly assumed in the literature, the relationship between modifier and modifiee could correspond roughly to conjunction. If this option is chosen, we end up with a semantic representation of a VP modified by a BPPA along the following lines.⁵

$$(19) \quad \llbracket \text{John works whistling tangos} \rrbracket = \lambda x. (\exists e_1. \text{work}(x, e_1) \wedge \exists e_2. \text{whistle}(x, \text{tangos}, e_2))(j)$$

Note that extraction from a constituent with such a denotation falls foul of the Single Event Condition (14), as this constituent describes two separate events, one

⁴Exactly how this is best done is not a matter I will address in this paper. I use the notation $e_1 = \langle e_2, e_3 \rangle$ to indicate, firstly, that e_1 is composed of the proper subparts e_2 and e_3 ; and secondly, that e_2 immediately precedes e_3 in temporal and causal terms. This is, though, a largely arbitrary choice.

I also ignore many issues concerning compositionality in what follows, for example the well-known issue of the compositional derivation of telic *write the cheque* from the addition of a bounded direct object *the cheque* to atelic *write*, as the issue is orthogonal to my present purposes. More seriously, however, space prevents me from attempting to justify the alternations between λ -abstracted and existentially quantified event variables in the examples that follow.

⁵Once again, I have to duck serious questions about the compositional derivation of such a representation because of lack of space. Many of the pertinent details are spelt out in Truswell (2007b).

of working and one of whistling. The corresponding interrogative is, accordingly, degraded, as in (7a). However, there is a second option, which involves identifying the λ -abstracted event variable within the adjunct denotation with a λ -abstracted event variable in the matrix VP denotation. If the previous process corresponds roughly to Predicate Modification in the terms of Heim & Kratzer (1998), then, this second option is closer (though not quite identical) to Event Identification. This will produce representations like the following, where the descriptive content within the adjunct predicate comes to be part of the description of the matrix event.

$$(20) \quad \llbracket \text{John drove Mary crazy whistling tangos} \rrbracket = \exists e_1, e_2, E. \text{whistle}(j, \text{tangos}, e_1) \wedge \text{crazy}(m, e_2) \wedge E = \langle e_1, e_2 \rangle$$

Unlike the case of (19), extraction from a constituent with a denotation like (20) does not violate the Single Event Condition: although there are distinct subevents within this representation, they jointly form a single macroevent. The Single Event Condition predicts extraction from a BPPA to be possible just in case it contributes to such a denotation. As we saw in (8b), this is indeed the case.

On this line of reasoning, then, we can describe the distribution of extraction out of BPPAs in terms of the availability of a process of identification of event variables. Whenever the BPPA's λ -abstracted event variable can be identified with a variable within the matrix VP's denotation, then, all else being equal, extraction from the BPPA should be possible. And a natural claim to make is that such identification is possible when there is no clash between the predicates in the matrix VP and in the adjunct. In other words, a representation like the one in (21) will be illegitimate because a working event is, intuitively, not a tango-whistling event, but the predicates *work* and *whistle* are predicated of the same event variable.

$$(21) \quad * \exists e. \text{work}(j, e) \wedge \text{whistle}(j, \text{tangos}, e)$$

Of course, a major case when there will be no clash between the descriptive content coming from the matrix VP and from the adjunct will be when there is simply no descriptive content attached to the event variable from the matrix VP. This gives a way of explaining the pattern noted above with respect to (16–17) on the one hand, and (18) on the other. The only cases which allowed extraction were those with no descriptive content attached to the process subevent variable in the matrix VP denotation, and we can now say that this is because there is nothing to block identification of the event variable in the BPPA denotation with a subevent variable within the matrix VP denotation in just such cases, and so satisfaction of the Single Event Condition becomes possible. In the other cases, no such identification can take place, and so the Single Event Condition will inevitably be violated and the extraction is predicted to be degraded.

4 ... And Back to Tense

The previous section spelled out one way of accounting for the distribution of extraction out of BPPAs, based on a decompositional model of event structure, plus the Single Event Condition (14). To be sure, though, alternatives can be envisaged. For instance, there are numerous ways in which a more purely syntactic approach

might proceed. Although I will not develop a full syntactic alternative here, a couple of comments about the likely shape of such an account are in order.

Most importantly, we can see from comparison of (7a) and (8a) that extraction from one and the same BPPA is sometimes possible and sometimes impossible, depending on the content of the matrix VP. It seems unlikely, then, that a syntactic account of these patterns would be based on factors relating to the internal structure of the BPPAs in question. Instead, it would be more natural to locate the distinction in some factor pertaining to the relationship between the BPPA and the matrix VP. For example, we may suggest that those BPPAs which allow extraction are attached lower than those which don't, and that the asymmetries noted in section 2 are due to a syntactic height effect. Alternatively, we might suggest that some syntactic relation holds between the matrix verb (phrase) and the BPPA in all and only those cases where extraction from the BPPA is possible. This would assimilate these cases to the mode of explanation lying behind the CED, where the relation of proper government distinguishes those domains which allow subextraction from those which prohibit it. I will not attempt to choose between these options here, or even properly develop either of them. The only important point for the discussion to follow is that both these approaches rely on factors external to the BPPA.

In order to choose between the semantic account developed in section 3 and this family of possible syntactic approaches, I will examine the possibility of extending them to cover the original generalisation discussed in the introduction, concerning the impossibility of extraction from tensed adjuncts. As ever, a more parsimonious theory in which fewer elements do more work is preferable, so if either of these approaches can be put to work in deriving the ban on extraction from tensed adjuncts, then it would be preferred.

In actual fact, the possibilities for extending a syntactic account of the BPPA subextraction patterns to cover the ban on extraction from tensed adjuncts seem rather remote. As we just saw, any syntactic account of the distribution of extraction from BPPAs is likely to be based on factors external to the structure of the BPPA itself. In contrast, in the case of the ban on extraction from tensed adjuncts, it seems like factors *internal* to the adjunct itself are responsible for the difference in extractability. Of course, it is in principle possible to claim that tensed adjuncts attach at a different height from untensed adjuncts, and so keep alive the syntactic modes of explanation alluded to above. However, such a move is unsupported by any independent evidence that I am aware of, and indeed, standard constituency tests, such as ellipsis and coordination, suggest that tensed and untensed counterparts such as those in (22a–b) are attached at roughly the same height despite the difference in extractability, as the interested reader will be able to verify.

- (22) a. I went to England [after speaking to him].
b. I went to England [after I spoke to him].
c. The man that I went to England [after speaking to *e*]. . .
d. *The man that I went to England [after I spoke to *e*]. . .

As suggested in the introduction, then, the natural place to look for a syntactic explanation of this difference would be in factors relating to the *internal* structure of the adjunct: there is something special about T(P) that interacts with an adjunction

structure to create a strong island. Once again, exactly what this “something special” would be is not relevant to the current argument. What *is* important, though, is that it is a factor internal to the adjunct in this case, whereas it was a factor external to the adjunct in the BPPA subextraction case discussed in the previous section. It seems, therefore, that any attempt at a syntactic explanation of these two phenomena must either end up as a disjunctive, rather than a unified, account, or resort to stipulations about the syntactic structure of the examples in question which lack any independent motivation.

On the other hand, it is possible to apply the Single Event Condition in a straightforward way to explain the ban on extraction from tensed adjuncts, as well as the patterns of extraction from BPPAs. The core of the explanation relates to the claim that, of the two ways described above of integrating the adjunct and the matrix VP semantically, only one, based on identification of event variables, can satisfy the Single Event Condition, whereas the other, closer to regular conjunction, always violates that condition. All we need is a reason why identification of event variables would be impossible when the adjunct is tensed, and the ungrammaticality of extracting out of tensed adjuncts will fall out automatically, as a consequence of the Single Event Condition.

One reason to presume that tense will render identification of event variables impossible comes from a simple-minded theory of the function of tense, as proposed, for example, in Higginbotham (1985). Higginbotham hypothesises that one function of tense is to existentially quantify the event variable introduced by a verb. While it is at least possible that an event variable within an untensed adjunct’s denotation remains λ -abstracted at the point where the adjunct is attached to the matrix VP, this cannot happen with a tensed adjunct, as the existential quantifier will bind the event variable at all subsequent stages in the derivation.

Now, any straightforward implementation of the operation of identification of event variables (such as that in Kratzer 1996, probably the most well-known example) will require that the event variables to be identified remain λ -abstracted at the point of identification — in intuitive terms, binding by any other operator such as an existential quantifier will “close off” the variable in question and render it inaccessible to further manipulation. Tensed adjuncts are doomed to be integrated semantically without recourse to identification operations, which, as we saw above, means that they will lead to representations, roughly as in (23), which violate the Single Event Condition and therefore disallow extraction.

$$(23) \quad \llbracket \text{I went to England after I spoke to John} \rrbracket = \exists e_1, e_2. \text{go}(\text{I}, \text{England}, e_1) \wedge \text{speak}(\text{I}, \text{j}, e_2) \wedge \text{after}(e_1, e_2)$$

In contrast, untensed adjuncts, whose event variables may remain abstracted at the point of integration with the matrix VP, are able to go down the identification route, and thereby satisfy the Single Event Condition. Contrasts like that in (22c–d) therefore emerge exactly as predicted.

We see, therefore, that with a bare minimum of natural additional assumptions, the event-structural account of patterns of extraction from BPPAs generalises quite naturally to cover the ban on extraction from tensed adjuncts. In contrast, syntactic accounts of the two phenomena, while not implausible in isolation, resist unification, and so inevitably retain either a disjunctive or a stipulative flavour.

Moreover, the semantic approach described in section 3 may offer us more chance of an explanation of a further puzzle raised in the introduction, namely the question of why it is specifically tensed *adjuncts*, as opposed to, say, tensed complements, which exhibit this property. Once again, syntactic theories will struggle to find any reason for T(P) and adjunction, being apparently unrelated properties, to interact, and so it is hard to see how to relate the prohibition on extraction from specifically tensed adjuncts directly to independently motivated properties of tense or of adjunction. However, there is a semantic difference between the behaviour of tensed adjuncts and many tensed complements.

A salient characteristic of the class of *bridge verbs* (verbs taking a clausal complement from which extraction is possible) in English is that the truth of the proposition expressed by the clausal complement is neither asserted, presupposed, nor denied by the speaker, as shown by the patterns of inference in (24).

- (24) a. John said [that Mary kissed Bill] \nrightarrow Mary kissed Bill.
 b. John didn't say [that Mary kissed Bill] \nrightarrow Mary kissed Bill.

In contrast, it is commonly assumed that tensed adjuncts are in fact presupposed, as shown in (25).

- (25) a. John talked to me [after Mary kissed Bill] \rightarrow Mary kissed Bill.
 b. John didn't talk to me [after Mary kissed Bill] \rightarrow Mary kissed Bill.

However, given that a typical matrix clause denotes a proposition asserting the existence of an event, what this suggests is that we can infer that *two* events (the talking or the not talking, and the kissing) have been asserted to exist in the actual world, in the tensed adjunct cases in (25), while only one event (the saying) has been asserted to exist in the actual world in the tensed complement cases in (24), the kissing event only being attributable, *modulo* considerations relating to the flouting of Gricean maxims, to John's belief world. The similarity of this pattern to that described by the Single Event Condition should be clear: where extraction from a tensed clause is possible, as in (26a), only a single event in the actual world is described. And where it is impossible, as in (26b), two distinct events are described. We therefore predict that extraction from bridge verb cases as in (24), but not from tensed adjunct cases as in (25), is possible.

- (26) a. Who did John say [that Mary kissed *e*]?
 b. *Who did John talk to me [after Mary kissed *e*]?

As a final twist in this story, it has been known since at least Erteschik-Shir (1973) that, among verbs taking tensed clausal complements, there is a relation between verbs displaying inference patterns like (24) and acceptability of extraction. Although I cannot go into full detail here, Erteschik-Shir discusses the class of *factive* verbs, which presuppose the truth of their complement, as in (27). In the terms of the above discussion, then, these are also cases in which we can infer the existence of two distinct events (the regretting or otherwise and the kissing) in the actual world.

- (27) a. John regrets [that Mary kissed Bill] \rightarrow Mary kissed Bill

- b. John doesn't regret [that Mary kissed Bill] → Mary kissed Bill

The Single Event Condition therefore predicts that extraction from complements of factive verbs should be degraded. Indeed, although opinions vary as to the strength of this degradation, there has been widespread agreement in the literature since Erteschik-Shir's dissertation that this is what we find.

- (28) ??Who does John regret [that Mary kissed *e*]?

Although this has only given a very brief sketch of these issues, then, it seems fair to say that the event-structural account of these extraction phenomena has the potential, not only to provide a principled theory of the interaction of tense and adjunction, but also to extend this theory to factive island phenomena, normally considered to be quite unrelated. I see no way to achieve parallel results on a purely syntactic approach.

5 Conclusion

This paper proposed an argument in favour of a partially semantic, rather than purely syntactic, account of the distribution of grammatical cases of extraction out of adjuncts. To be sure, the argument is a long way from a proof that no syntactic account of these data is possible. In fact, given the current lack of consensus about which elements are strictly syntactic, rather than located at or beyond the interfaces, I doubt that such a proof is presently possible. Rather, the argument presented here has hopefully shown that we can make fewer elements do more work in locality theory if we allow some of those elements to be semantic in nature, rather than attempting to shoehorn all of locality theory into the syntax, as we traditionally have. All else being equal, then, the partially semantic theory is to be preferred.

In and of itself, this is a pretty fragile argument, depending rather heavily on whether all else really is equal. In this conclusion, I want to address two issues which hopefully strengthen the argument somewhat. Firstly, it should be clear that, although this paper has concentrated on extraction from BPPAs, these do not represent the only legitimate case of extraction from adjuncts. Other classes which sometimes allow extraction include *in order* clauses and other purpose clauses, as discussed in Jones (1991) and illustrated in (29a); prepositional participial adjuncts (29b) and related constructions in Italian discussed by Cinque (1990); and, of course, the better-known case of adjunct P-stranding under \bar{A} -movement (29c). Truswell (2007b) shows that a slightly more complex version of the Single Event Condition can handle all of these cases in a natural way.

- (29) a. What did you come here [to talk about *e*]?
b. Who did you go home [without talking to *e*]?
c. Which play did you fall asleep [during *e*]?

On the other hand, it is likely that each extra class of extractions will only add to the disjunctive nature of any putative syntactic account of extraction from adjuncts. For example, the acceptability of (29c) cannot be due purely to factors relating to the internal structure of the adjunct, as extraction from the same PP is impossible in other cases.

(30) *Which meal did you listen to the radio [during *e*]?

However, the obvious explanation in terms of external syntactic factors, namely height effects, is also unavailable, as *during* PPs are often taken (for example by Hornstein & Weinberg 1981) to be relatively high, and so in theory unavailable for P-stranding even in the legitimate case (29c). Some other factor would have to be appealed to in this case, and the disjunctive nature of the account would become more noticeable.

The other issue I want to raise here concerns the overall architecture of a grammar containing a condition such as the Single Event Condition. It is important to be clear that such a condition can be implemented perfectly well in any relatively sparse grammatical architecture, such as Chomsky's Y-model, as well as in more elaborate models such as the parallel architecture of Jackendoff (2002). All this story requires is the assumption that certain words can impose conditions on the semantic environments in which they are used, and that these environments are defined in terms that do not necessarily map directly onto syntactic factors. Such a position is implicit in almost any work in formal semantics (NPI-licensing springs to mind as an obvious example: there is no purely syntactic counterpart of the notion of *downward-entailing environment*). There is therefore nothing intrinsically more costly about using a semantically-based condition to motivate locality effects than using a syntactically-based condition, and so the comparison between the two seems to me to remain valid. The arguments given above therefore seem to me to be a reason to prefer the semantic account outlined in this paper to any syntactic alternative that I can conceive of.

References

- Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9.5–16.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- . 1982. *Some Concepts and Consequences of the Theory of Government and Binding*. Cambridge, MA: MIT Press.
- . 1986. *Barriers*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 1990. *Types of \bar{A} -dependencies*. Cambridge, MA: MIT Press.
- Dowty, David. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and Montague's PTQ*. Dordrecht: Reidel.
- Erteschik-Shir, Nomi, 1973. *On the Nature of Island Constraints*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Hale, Kenneth, & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In *The View from Building 20: Essays in Honor of Sylvain Bromberger*, ed. by Kenneth Hale & Samuel Jay Keyser, 53–109. Cambridge, MA: MIT Press.
- Heim, Irene, & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.
- Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16.547–593.
- Hornstein, Norbert, & Amy Weinberg. 1981. Case theory and preposition stranding. *Linguistic Inquiry* 12.55–91.

- Huang, C.-T. James, 1982. *Logical Relations in Chinese and the Theory of Grammar*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Jackendoff, Ray. 2002. *Foundations of Language*. Oxford: Oxford University Press.
- Johnson, Kyle, 2002. Towards an etiology of adjunct islands. Ms., University of Massachusetts, Amherst, MA.
- Jones, Charles. 1991. *Purpose Clauses: Syntax, Thematics, and Semantics of English Purpose Constructions*. Dordrecht: Kluwer.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase Structure and the Lexicon*, ed. by Johan Rooryck & Laurie Zaring, 109–137. Dordrecht: Kluwer.
- Lakoff, George. 1970. *Irregularity in Syntax*. New York: Holt, Rinehart and Winston.
- Levine, Robert, & Ivan Sag. 2003. Some empirical issues in the grammar of extraction. In *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar*, ed. by Stefan Müller, 236–256. Stanford, CA: CSLI.
- Ross, John Robert, 1967. *Constraints on Variables in Syntax*. Cambridge, MA: Massachusetts Institute of Technology dissertation.
- Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10.80–126.
- Szabolcsi, Anna. 2006. Strong vs. weak islands. In *The Blackwell Companion to Syntax, Volume IV*, ed. by Martin Everaert & Henk van Riemsdijk, 479–531. Oxford: Blackwell.
- Truswell, Robert. 2007a. Extraction from adjuncts and the structure of events. *Lingua* 117.1355–1377.
- , 2007b. *Locality of Wh-movement and the Individuation of Events*. University College London dissertation.
- Uriagereka, Juan. 1999. Multiple spell-out. In *Working Minimalism*, ed. by Samuel Epstein & Norbert Hornstein, 251–282. Cambridge, MA: MIT Press.
- Vendler, Zeno. 1957. Verbs and times. *Philosophical Review* 66.143–160.