



WILLIAM D. DAVIES and STANLEY DUBINSKY

## ON EXTRACTION FROM NPS \*

**ABSTRACT.** This paper sheds new light on the conditions governing extraction from NPs. A close examination of *wh*-extraction out of object NPs reveals that previously unnoticed semantic factors play a greater role than has been recognized. In particular, we find that NPs lacking 'participant' structure do not permit *wh*-extraction at all, and that certain NPs permit extraction even when they are definite. At the same time, the prohibition on *wh*-extraction from subject NPs is shown to be a purely syntactic phenomenon which arises from the particular way in which the Extended Projection Principle is satisfied in English.

### 1. INTRODUCTION

The contrast shown in (1) has long been taken as an illustration of subject-object asymmetry *vis-à-vis wh*-extraction:

- (1)a. Who is the class reading a book about?
- b. \*Who is a book about being read by the class?

In (1a), *who* is extracted from the object NP *book about*. In (1b), the same extraction, with *book about* in subject position, is illicit. According to Chomsky's (1973, p. 250) Subject Condition, (1b) is bad because *wh*-movement crosses out of a subject phrase. Chomsky formalized this into a general principle constraining transformations, but retained in this principle the basic observation that extraction out of complement NPs is permitted (in opposition to extraction out of subject NPs).

It is clear, however, that nouns differ with respect to whether or not their complements may be extracted. *Concrete* nouns do not allow extraction at all. Consider example (2):

- (2) \*Which church did Ashley prefer/like/describe the café near?

---

\* We thank Alice Davison, Elena Gavruseva, Frederick Newmeyer, Hyeson Park, Roumyana Slabakova, Matthew Traxler, Gert Webelhuth, three anonymous reviewers for *Natural Language and Linguistic Theory*, and audiences at the 2001 Annual Meeting of the Linguistic Society of America, the 31st Annual Meeting of the North East Linguistic Society, the University of Iowa, and the University of South Carolina for helpful discussion of some of the data and issues presented here. Naturally, the authors bear sole responsibility for any errors, misapprehensions, or wrongheadedness in this work, but we sincerely hope they are at a minimum.



In contrast, *representational* nouns permit extraction only under certain circumstances, as shown in (3):

- (3)a. Who did Ashley read/edit/buy/\*destroy/\*stack books about?
- b. \*Who did Ashley read the/that/her book about?
- c. Who did Ashley write her/that/\*your book about?

Extraction is possible when representational nouns are complements of some predicates, but not others, as seen in (3a). Extraction out of representational nouns appears not to be possible when they co-occur with definite determiners (as in (3b)). Some interesting exceptions, such as (3c), beg (and will receive) an explanation.

Finally, at the other extreme from flat nouns lie *process* nouns. For them, extraction is always possible irrespective of verbal context or definiteness:<sup>1</sup>

- (4) Who did Ashley participate in/watch/protest the coronation of?

In the analysis that follows, we will show that a fuller understanding of these facts (and contrasts) can be achieved through reference to Grimshaw's (1990) notions of 'argument structure' and 'participants'. Briefly, we will seek to demonstrate that extraction out of complement NPs is restricted to elements which can be linked to an argument or participant in the lexical conceptual structure of the head noun. Flat nouns, having no argument structure and not having participants linked to their lexical conceptual structure, do not permit extraction at all. Process nouns, coming with a self-contained argument structure, generally do permit extraction.

---

<sup>1</sup> Godard (1992) provides a classification of nouns that is somewhat different from the one presented here. Claiming that no syntactic account of NP extraction is possible, Godard tries to show that extraction out of NPs is, in principle, limited to "the first argument in the NP in which they occur (p. 235). While clearly on the right track, insofar as the analysis recognizes the saliency of argument structure in determining the possibilities for extraction, this account fails to explicitly predict the type of contrast seen in (ia), in which extraction possibilities are determined (in part) by the choice of matrix predicate, or in (ib), where definiteness plays a role.

- (i)a. Who did they read/\*burn a book about?
- b. Who did Joan prefer a/\*that picture of?

That is, if the phrase *about who* is the first argument in *book about who*, then extraction ought to be possible with either *write* or *burn*. This said, it should be acknowledged that Godard's account goes a long way toward demonstrating the shortcomings of any *purely* syntactic account of NP extraction.

Representational nouns, as we shall soon see, have a variable relationship with their associated verbs of creation and use, such that extraction is permitted in certain contexts and banned in others. Before presenting our analysis, it will be useful to review prior analyses of extraction out of complement (representational) NPs, in order to understand where previous studies leave us empirically.

In section 2, we briefly examine previous observations about extraction from NPs, generalizations that help form the basis of our study. Following this, we present our analysis of extraction from NPs, showing that there are both important semantic determinants (section 3) and syntactic determinants (section 4). In section 5 we return to the subject/object asymmetry illustrated in (1), proposing that the Subject Condition in English owes to the fact that DP is an absolute barrier to movement and the requirement that *all* English subjects project a DP.<sup>2</sup>

## 2. BACKGROUND

There is a rich literature on extraction from NPs, and the insights of many previous analyses inform the present study, laying the theoretical and empirical foundation on which our proposal is built. Bach and Horn (1976) argued that the contrast in (1) is not the result of a subject-object asymmetry. They claimed that extraction out of NPs is generally prohibited, and that (1a) is grammatical only because it does not involve extraction out of an NP. In making this claim, they point to contrasts such as (5):

(5)a. Who did they write a book about?

b. \*Who did they destroy a book about?

According to Bach and Horn, the PP *about whom* is contained within the object NP in (5b), but not in (5a). This structural difference also explains

---

<sup>2</sup> In our discussion in some instances we refer to NP and in others to DP. The reasons are both historical and principled. The literature on this topic spans roughly 35 years and many theoretical changes. With the advent of the DP analysis (e.g., Abney 1987), some constituents analyzed as NPs in ‘classical’ theory are analyzed as DPs in current theory. However, the principal semantic constraint on extraction proposed in section 3 is, in fact, a constraint on extraction from NPs and not DPs. DP projections become crucial in sections 4 and 5, where we consider syntactic factors relevant to extraction and the EPP in English. Where these factors are not essential to the analysis or discussion, we refer to NP and not DP.

the contrast in (6), where pronominalization is assumed to affect an entire NP:

- (6)a. They wrote [NP it] [PP about Nixon]  
 b. \*They destroyed [NP it [PP about Nixon]]

On this view, the contrast between (1a) and (1b) tells nothing about subject-object asymmetries. Similar to Bach and Horn's constraint on extraction from NP, our analysis takes extraction out of DP to be prohibited in the general case.

Erteschik-Shir (1981) offers a critique of Bach and Horn's account, claiming that sentences such as (5a) have two available structures: the one proposed for it by Bach and Horn, shown in (7a), and the one that Bach and Horn claim it cannot have, shown in (7b):

- (7)a. who<sub>1</sub> did they write [NP a book] [PP about t<sub>1</sub>]  
 b. who<sub>1</sub> did they write [NP a book [PP about t<sub>1</sub>]]

As evidence for her claim, she notes that some verbs that allow extraction in the manner of (5a) do not permit the complement structure that would be necessary for Bach and Horn's analysis (as in (7a)) to work. The data in (8) illustrate:

- (8)a. Who did they read/finish a book about?  
 b. \*They read/finished it about Nixon.

In (8a), we see that extraction is possible from the complement of the verb *read* or *finish* even though the verb does not sanction the NP-PP complement structure (as (8b) shows). Notice further that *finish* in (8a) can be taken to mean 'finish writing' or 'finish reading', and that (8b) is still unacceptable even when it means the former. Erteschik-Shir's (1973, 1981) own analysis focuses on the discourse role that different verbs play in licensing extraction out of their complements (e.g., objects of verbs of creation license extraction in contrast with others, cf. (5a) and (5b)). Although only part of the story, in the analysis we propose here, specific predicates play an integral role in licensing extraction out of certain NPs (section 4).

Alongside these observations, it is also well known that extraction from object NPs is more widely acceptable out of indefinite NPs than out of definite ones. Observe example (9):

- (9)a. Who did you read some/many books about?  
 b. \*Who did you read the/tha book about?

To account for this, Fiengo and Higginbotham (1981) invoke a ‘Specificity Condition’ that blocks extraction from NPs having ‘some definite reference’. On the syntactic side, Bowers’ (1987) solution to this problem proceeds from a division of quantifiers into weak and strong classes. For him, the weak (indefinite) ones are taken to be adjectival, while the strong (definite) ones are determiners (with the latter blocking extraction). While we will see below (sections 3 and 4) that definite DPs are not uniformly islands for extraction, any adequate analysis, including our own, must account for any ‘definiteness effect’.

Perhaps the most comprehensive proposal regarding extraction from NPs is provided in Diesing (1992). There she ties together many of these observations, proposing that extraction out of NPs is governed by whether or not the NPs in question have undergone quantifier raising (QR). On her account, NPs that have strong determiners (such as *each*) are presuppositional and undergo obligatory QR before extraction can take place. At the same time, NPs that are objects of verbs that presuppose their existence (such as *destroy*) are also taken to undergo obligatory QR before extraction can take place. Diesing’s solution is to permit extraction only out of NPs that are not required to undergo QR.

While Diesing’s account comes closest to accommodating all the syntactic and semantic observations that antecede it, it still runs into problems. Given the scope of her account, we examine three key cases. First, if definiteness correlates with QR, as Diesing claims, then all definite NPs should prohibit *wh*-extraction. Since NPs having genitive pronouns are a subclass of definite NPs, we would therefore expect extraction to be impossible from them. Indeed, this is usually the case, as (10) shows:

- (10) \*Who did Tom read his/their/my book about?

However, notice first that genitive determiners can have either possessive or agentive meanings, as illustrated in (11):

- (11)a. Roger owns my pictures of Jane.  
 b. Roger was the photographer who took my pictures of Jane.

Accordingly, in (11a) *my* denotes the ‘taker of the pictures’ and not their possessor, and in (11b) *my* denotes the possessor. With this in mind, examine (12):

- (12) Who did Tom write his story about?

Example (12) is perfectly grammatical when *his* both takes *Tom* as an antecedent and carries an agentive, rather than a possessive, meaning. The point is made even more salient in (13):

- (13)a. Tom finished his/my book about Nixon.

- b. Who did Tom finish his/\*my book about?

In (13a), the genitive pronoun can have a possessive or agentive sense, depending on whether we interpret *finish [X’s] book* to mean ‘finish reading [X’s] book’ or ‘finish writing [X’s] book’. Of course, in the latter instance, only *his* is appropriate. The extraction in (13b) is possible only with the genitive pronoun *his* and only with the meaning ‘finish writing his [Tom’s] book’. Under Diesing’s account, both (12) and (13b) should be uniformly ungrammatical, since both involve extraction out of NPs that are definite and that undergo QR. Thus, Diesing’s analysis appears to be too strong, in that it would rule out a whole class of grammatical cases.

The second case involves Diesing’s analysis being too weak. According to her account, indefinite NPs that are not in presuppositional contexts should allow *wh*-extraction, since they do not undergo obligatory QR. However, consider the contrast between (14) and (15):

- (14)a. What did Jake take a picture of? (answer: The Alamo)  
 b. What did Tannu Tuva start a war over? (answer: trade)  
 (15)a. \*What did Jake write a letter of? (answer: apology)  
 b. \*What did Sherman commit a crime of? (answer: passion)  
 c. \*What did Joseph make himself a coat of? (answer: many colors)  
 cf. What color coat did Joseph make himself? (answer: one of many colors)

Given that the NPs in (14) and (15) are all indefinite and are complements of verbs of creation, they should not (according to Diesing) undergo

obligatory QR and should allow *wh*-extraction. The ungrammaticality of (15a,b,c) is therefore unexpected under Diesing's account.<sup>3</sup>

Another shortcoming of the QR-based account is that it makes wrong predictions concerning extraction from NPs whose quantifiers are negative polarity items (NPIs). NPIs are 'dependent' quantifiers, and dependent (existential) quantifiers cannot assert or imply existence (den Dikken and Giannakidou 2001, p. 167). Thus, the following contrast obtains:

- (16)a. ?He didn't burn some books about Truman, because there were none.
- b. He didn't burn any books about Truman, because there were none.

In (16a), the quantifier *some* creates an implicature of existence regarding *books about Truman*, even though this implicature is defeasible. Example (16a) is somewhat odd for this reason. In (16b), on the other hand, the existence of *books about Truman* is neither entailed nor implicated, and the sentence sounds perfectly fine. Returning to Diesing's account, we would assume that NPs introduced by existential NPI quantifiers should not have to undergo QR because they are not presuppositional. Accordingly, *wh*-extraction should always be permissible out of an indefinite NP with an NPI quantifier such as *any*. Example (17) shows that this is not the case:

- (17)a. Who didn't he read/write any books about?
- b. \*Who didn't he burn/shelve any books about?

The extraction in (17b) ought to be as good as that in (17a) under Diesing's analysis, since the NP *any books about* is not presuppositional and need not undergo QR. In contrast, the ungrammaticality of (17b) is easily explained under our account, since predicates such as *burn* and *shelve* do not involve the participant structure associated with the noun *book*.

<sup>3</sup> Alongside the extractions out of NPs under consideration here, one finds grammatical cases such as in (i).

- (i)a. Which church did he see the towers of?
- b. Which ingredient did you use two cups of?

However, extractions involving partitive and part-whole constructions display different properties, and suggest the possibility that the sentences in (i) do not involve extractions out of NP at all.

What we find is that the prohibition of *wh*-extractions out of NP is in fact determined more by nature of the NP's denotation. Extraction turns out to be sensitive to differences between metaphysical concept denotations and physical token denotations, as illustrated in (18):

- (18)a. When writing papers, which presidents do children usually use books about?
- b. \*When propping open their desks, which presidents do children usually use books about?

Example (18a) is significantly better than (18b). The difference between them resides not in whether the book is existentially presupposed or not, but rather in whether *book* denotes a physical or metaphysical object. Using a book to write a paper entails utilizing the conceptual entity denoted by *book*. In using a book to prop open a desk, on the other hand, one utilizes the physical entity denoted by the noun. It is only the conceptual (or metaphysical) denotation of *book* that involves argument structure (i.e., one writes and reads instances of the former, not the latter). Thus, while contextual factors (such as definiteness) do indeed play a role in determining extraction possibilities out of NPs, nominal argument structure is just as crucially involved.

The analyses reviewed here, with the exception of that of Bach and Horn, share the general feature of explaining the extraction out of NPs on the basis of the context into which the NP is inserted. Only Bach and Horn's account (with its general prohibition on extraction from NPs) is context free. The others attribute licensing (or prohibition) of extraction out of NPs to: (i) whether the NP is a subject or complement (Chomsky); (ii) whether the NP is the focus of discourse (Erteschik-Shir); (iii) whether the NP is contained within a DP (Bowers); or (iv) whether the NP has undergone QR (Diesing). We note that all of the above approaches are on the right track. The subject-object asymmetry central to Chomsky's account is real. At the same time, extraction out of DPs is generally disallowed, along the general lines of Bach and Horn's analysis. The choice of verb does, as Erteschik-Shir observed, affect whether a *wh*-operator can be moved out of it. And finally, definiteness does determine that extraction is blocked in certain cases. This all said, we will see that the key to understanding extraction possibilities is to be found in the semantics of the NPs themselves. In the following two sections, we propose an analysis that captures the following: (i) extraction out of NPs is only licensed for elements that count as participants in the lexical conceptual structure of the head noun



(section 3), and (ii) extraction out of definites is permitted when agentive structure triggers abstract N-V incorporation (section 4).

### 3. SEMANTIC DETERMINANTS OF EXTRACTION POSSIBILITIES

With the foregoing in mind, we embark on our proposal, namely to develop a classification of nominals that keys on the important semantic criterion of those elements that can be extracted from NPs, that is whether or not they are participants (in the sense of Grimshaw 1990). We first motivate the typology on the basis of complex event nominals, result nominals, and concrete nouns; we then situate representational nouns in the typology. It turns out that representational nouns, like some other nominals, belong to more than one class, and it is this ambiguity that accounts for why extraction from NPs headed by representational nouns is sometimes permitted and sometimes prohibited.

#### 3.1. *A Typology of Nominals*

Grimshaw's (1990) typology of nouns is most useful for categorizing the behavior of nouns with respect to extraction. Roughly speaking, Grimshaw examines nouns in terms of argument structure and their relation to lexical conceptual structure. She identifies a continuum of noun types. At one end are process nouns that always take A-structure. At the other end are concrete nouns that never do and are never linked to arguments in the lexical conceptual structure.

Nominals in *-ing* are at the A-structure end of the continuum since they always have a process reading (as pointed out by Lebeaux 1986). In Grimshaw's analysis, *-ing* nominals are complex event nominals, the only type of nominals that can take an A-structure and obligatorily occur with their arguments, as in (19) (Grimshaw 1990, p. 50):

- (19)a. the felling \*(of the trees)
- b. the destroying \*(of the city)

Grimshaw also analyzes certain process nominals as complex event nominals; these also can take A-structure, as in (20):

- (20)a. the examination/\*exam of the patients
- b. the expression of one's feelings

With these nominals it appears that the occurrence of arguments is optional, given the grammaticality of (21):

- (21)a. The exam(ination) was long.
- b. The expression is desirable.

However, within Grimshaw's system the occurrence of arguments is not optional; thus, the optionality of the argument with these nominals is illusory. In fact, nominals such as *examination* and *expression* are ambiguous between an *event* reading, which must take A-structure, and a *result* reading, which cannot take A-structure. As Grimshaw (1990, p. 59) puts it,

Complex event nominals and corresponding simple event and result nominals have related lexical conceptual structures, or lexical meanings, but only complex event nominals have an event structure and a syntactic argument structure like verbs. The argument structure of complex event nominals licenses (and indeed requires) arguments.

Grimshaw provides extensive argumentation for this distinction between complex event nominals and simple event and result nominals. One way of disambiguating the two is through the use of certain modifiers, such as *frequent*, which can only occur with the event reading. When the *modifier* occurs with a singular noun, the *arguments* of the nominal are *obligatory*:

- (22)a. The examination was annoying.
- b. \*The *frequent* exam(ination) was annoying.
- c. The frequent examination of the patients was annoying.

On the *result* reading, these nominals share with concrete nouns such as *dog* and *stone* the *inability* to take arguments or to occur with modifiers such as *frequent*:

- (23)a. the (\**frequent*) dog
- b. the (\**frequent*) stone

There is an important difference between result nominals and concrete nouns that should not be overlooked, however. As Grimshaw points out, "even result nominals imply the existence of certain *participants* [our emphasis] in the situation they are used in" (p. 54). Thus, the existence of an examination implies that someone created it. Likewise, the existence of an expression implies that someone expressed it. Grimshaw proposes

to account for these implications in terms of a notion ‘participant’, part of the information contained in the lexical conceptual structure (LCS) of verbs and certain nouns. Grimshaw claims that the difference between verbs and complex event nominals on the one hand and other nominals that have participants on the other hand is that the former project their participants into A-structure, making them grammatical arguments, while the latter have participants that are not grammatical arguments. Of course still other nominals, such as *dog*, do not have any participants as part of their LCS.

Returning to our emerging typology, we find that there are also nominals that always imply participants but can never take A-structure. Included in this group are result nominals such as *victory*. Note that *victory* can take overt participants, but that in Grimshaw’s terms these are not arguments, since *victory* is not a complex event nominal, as shown by the ungrammaticality of the modifier *frequent* with the singular nouns *victory* and *triumph* (compare 24a and 24b):

- (24)a. The Yankees’ (\*frequent) victory/triumph over Seattle delighted their New York fans.
- b. The Yankees’ frequent subordination of Seattle delighted their New York fans.

Thus, we have identified four types of nominals: (a) nominals that always take A-structure, e.g., *examining*; (b) nominals that sometimes take A-structure but always imply participants, e.g., *examination*; (c) nominals that never take A-structure but always imply participants, e.g., *victory*; and (d) nominals that never take A-structure and never imply participants, e.g., *dog*. These four nominal types can be categorized into three nominal classifications, as shown in (25), on the basis of their argument/participant structure. These nominal classes are (I) complex event nominals which have argument participants, (II) result nominals which have non-argument participants, and (III) concrete nominals which have no participants. Type (a) belongs to class I; type (c) belongs to class II; and type (d) belongs to class III. Type (b), exemplified by *examination*, belongs to class I or class II, depending on its interpretation and context:

(25)

<i>I. complex event nominals</i>	<i>II. result nominals</i>	<i>III. concrete nominals</i>
(have argument participants)	(have non-argument participants)	(have no participants)
-----	-----	-----
<i>examining</i> (i.e., -ing nominals)	<i>victory</i>	<i>dog</i>
<i>examination</i> (process reading)	<i>examination</i> (result reading)	

A key fact about this typology is the existence of a noun type (the ambiguous process/result nominals such as *examination*) that belong to two classes, as noted in Grimshaw.

3.2. *The Classification of Representational Nouns*

We will now situate representational nouns in this classification (25). We propose that representational nouns are yet a different type of ambiguous noun, those that sometimes imply participants but never take A-structure. As noted above with respect to the discussion of (18), representational nouns such as *book* and *picture* carry two basic meanings. They can denote the concrete object to which they refer or the information represented by or contained in the concrete object that they denote. When we say that one uses a book to prop open a desk, we refer to the physical or concrete entity denoted by the noun *book*. On the other hand, when we say that one uses a book to write a paper, we refer directly to the information contained in the book, not to the physical object itself. The same is true with predicates such as *write*. When an author writes a book, it is not the individual physical object with printed pages, a binding, and (potentially) illustrations that the author creates; rather the author creates the information that is contained in one of the physical entities denoted by the noun *book*. There are certain idioms using representational nouns that make this distinction very salient; there are times when it is impossible to actually refer to a concrete object denoted by the noun. This is illustrated in the following dialogue:

(26)A: Do you have confidence in this consultant’s advice?

B: Oh, Terry wrote the book on effective reorganization.

In B’s answer, there is no reference to any actual book, but to the information that would be contained in such a book, should it really exist. One

finds a similar situation with the use of *picture* in the idiom *paint me a picture*, as in (27):

(27)A: Do you understand precisely what the ramifications are?

B: No, **paint me a picture.**

Again, in B's response there is no physical object that will result; only the information that would be represented by such a picture should one actually be created.

Thus, representational nouns such as *book* have both an informational and a concrete sense, and we would argue that the particular senses are accessed separately. The informational sense is accessed when the focus is on the information represented by the noun. So, as above, the informational sense of *book* is accessed when the content of the book is the focus, which is true in cases of reading events, writing events, and some others. In its informational sense *book* is linked to participants in their LCS. A book implies that someone wrote the book and that that book was written about someone or something. We saw this above when we noted that genitives can be interpreted either as possessors or agents. The agentive reading is forced in a sentence such as (28):<sup>4</sup>

(28) Tom<sub>i</sub> wrote his<sub>i</sub> book about Nixon.

In (28) *his* is linked to the agent or author participant in the LCS of *book* and *Nixon* is linked to the theme or topic of writing participant.

However, only the concrete sense of *book* is accessed when the physical object is the focus. When a book is burned, it is the physical object that is burned. When a book is shelved, again it is the printed pages, binding, and so on that is shelved. In these events, it is clearly the physical object and not the information contained in the book that is relevant. We would argue that in these instances *book* behaves like a concrete noun such as *dog* and therefore involves no participants.<sup>5</sup>

<sup>4</sup> Of course, it is possible to interpret *his book* in (28) to refer to a physical object that Tom is currently holding in his hand, but this is clearly a secondary interpretation that requires a bit of context to be possible.

<sup>5</sup> It should be noted that the concrete and informational interpretations of representational nouns are best observed with those that involve non-visual information. Accordingly it is easier to distinguish the concrete vs. informational sense of nouns such as *book*, *exam*, *report*, *compact disk*, and *cassette* than nouns such as *picture*, *photo*, and *painting*. It is further the case that these senses are even more difficult to distinguish for representational nouns that denote individual entities such as *painting* (as opposed to nouns whose denotation might involve multiple copies, such as *photograph*). It is thus very difficult

Therefore, **representational nouns** such as *book*, *picture*, *essay*, and *photograph*, like Janus, have **two faces**. They have a *victory*-face, that is, a side of them that is **linked to participants** in their LCS, and they have a *dog*-face, a side that like other concrete nouns is **not linked to participants** in their LCS. And as argued above, these two different faces are accessed in distinct events, and the context in which they occur determines which LCS is accessed. Thus, representational nouns fit into **our classification system** as in (29):

(29)

<i>I. complex event nominals</i> (have argument participants)	<i>II. result nominals</i> (have non-argument participants)	<i>III. concrete nominals</i> (have no participants)
-----	-----	-----
<i>examining</i> (-ing nominals)	<i>victory</i>	<i>dog</i>
<i>examination</i> (process reading)	<i>examination</i> (result reading)	
	<i>book</i> (informational reading)	<i>book</i> (physical reading)

We will see in what follows that this Janus-like quality of representational nouns provides a means for explaining their behavior with respect to the possibility or impossibility of extraction from noun phrases that they head.

3.3. *Participants and Extraction from NPs*

We are now in a position to motivate the key semantic determinant of extraction from NPs. We propose that only participants may extract. This generalization together with the dual nature of representational nouns explains why extraction from representational NPs is possible in some cases and impossible in others.

---

to distinguish the concrete (physical) from the informational (metaphysical) senses of *Cezanne's portrait of his father*, since there is only one concrete entity that embodies this representation. In contrast, *Melville's novel about the white whale* denotes a single informational entity which exists in the world as thousands of concrete tokens.

We observe first that complex event process nominals freely allow extraction of their arguments, as illustrated in (30):<sup>6</sup>

- (30)a. What did they observe/hear about/remember/decry the production of?  
 b. Which patient did the med students participate in/observe/miss the operation on?

Notice that extraction in these cases fails to exhibit two effects that we have observed with regard to extraction out of representational NPs. First, extraction out of the complex event NP is unaffected by the verb that selects it. Second, extraction appears to be licensed across a definite determiner.<sup>7</sup>

<sup>6</sup> Note, however, that extraction out of certain *-ing* constructions is prohibited. Compare (i) and (ii)/(iii) below:

- (i) Which patients did you watch the examination of \_\_ ?  
 (ii) \*Which patients did you watch the/his examining of \_\_ ?  
 (iii) \*Which patients did you watch his examining \_\_ ?

This prohibition on extraction affects only *-ing* constructions that have a definite determiner *the* or a genitive subject. In contrast, *-ing* constructions that have an accusative subject freely permit extraction. Compare (ii)/(iii) and (iv):

- (iv) Which patients did you watch him examining \_\_ ?

We would attribute these differences in extraction possibilities to certain differences among the four cases in question. The expression in (i), *the examination of*, involves the simple projection of a simplex derived head noun. The expression *him examining*, in (iv), involves something on the order of a small clause and has no nominal structure. But the analogous expressions in (ii) and (iii) involves a nominal projection of the affix *-ing*, a VP projected by *examine*, and head movement. The differences among these would then rest on (i) not having the sort of inner clausal structure often attributed to (ii)/(iii), and on (iv) not having the outer nominal structure exhibited by (ii)/(iii).

<sup>7</sup> The status of the definite determiner *the* with process (complex event) nouns is constrained by the fact that other determiners cannot co-occur with them. Thus, while *the production of X* can have either a result or process interpretation, *a production of X* denotes only a result meaning (Grimshaw 1990, pp. 54–56). Accordingly, the noun *revolution* in (ia) refers either to the process of revolving or to the result of this process, while in (ib) it only denotes the result:

- (i)a. Scientists observed the revolution of a satellite around Mars. (process/result)  
 b. Scientists observed a revolution of a satellite around Mars. (result/\*process)

In contrast with complex event nominals, concrete nouns do not allow extraction at all, either of possessive/descriptive complements or of modifiers. This is shown in (31):

- (31)a. \*Which neighbor did Shelly chain [some dogs of \_\_] to a tree?  
 b. \*Which wood was Gene hoping to find [a table of \_\_] for the kitchen?  
 c. \*What sort of fur was George looking for [a dog with \_\_]?

In (31), the apparent complements of *dogs* and *table* are unable to extract, despite the fact that they are both indefinite and in (31b,c), the existence of the object is not presupposed.

In the middle space between complex event and concrete nominals, are result nominals, such as *loss*, *victory*, and *triumph*:

- (32)a. Who were the Phillies hoping for a victory/some victories over \_\_?  
 b. \*Who were the Phillies hoping for the/tha victory over \_\_?

Extraction of participants out of result NPs is generally allowed, as (32a) shows. However, definiteness blocks extraction, as in (32b).<sup>8</sup>

This is an important consideration when evaluating definiteness effects, since definite determiners do not function in quite the same way with process nouns as they do with result nouns. Pace Grimshaw (1990, p. 55), *the* does not appear to encode definiteness in the process interpretation of (ia). In the first place, the definite determiner adds no meaning to the NP and is optional in the process interpretation, as (ii) shows. Secondly, the use of *the* with process nominals does not require presupposed existence, as the intensional context in (iii) shows. Finally, as (iv) shows, apparently definite process nouns occur in *there* constructions, from which definite NPs are normally precluded:

- (ii) Scientists observed revolution of a satellite around Mars.  
 (iii) To account for the observed solar wobbles, some scientists have posited (the) revolution of the sun around some unknown gravity source.  
 (iv) And then, there was the revolution of a satellite around Mars on the video transmission, which the scientists had not expected.

<sup>8</sup> When the definite determiner is a genitive pronoun denoting an agent, extraction is again possible, as we see in (i).

- (i) Who did the Phillies<sub>1</sub> relish their<sub>1,\*2</sub> victory over \_\_?



From the above, we can distill the following generalization regarding extraction out of NPs: *only participants may extract*. Note that arguments are necessarily participants, but not vice versa:

(33)

<i>I. complex event nominals</i> (have argument participants)	<i>II. result nominals</i> (have non-argument participants)	<i>III. concrete nominals</i> (have no participants)
extraction of participants permitted	extraction of participants permitted definiteness effects observed	extraction not permitted

On the basis of this generalization and classification, we can now examine the behavior of nouns that belong to more than one class.<sup>9</sup>

As Grimshaw notes, nouns such as *examination* have both a process and a result reading, and behave differently with respect to extraction, depending on whether they have a process or a result interpretation:

(34) Which patient did the surgeon forget to observe the examination of \_\_ ?

(35) What subject was the student hoping to pass an/some/\*the/\*that examination on \_\_ ?

As a process nominal in (34), extraction out of an NP headed by *examination* is licensed across the apparent definite determiner *the* (however, see note 7). In (35), *examination* is a result nominal, and while extraction is licensed, it is clearly subject to definiteness effects. Here, extraction is licensed across *a* and *some*, but not over *the* or *that*. This behavior accords well with the claim that *examination* belongs to both class I and class II.

As we have seen, representational nouns can also belong to either of two classes, depending on their denotation. In their informational sense, they belong to the class of result nouns, such as *victory*. In their physical sense, they belong to the class of concrete nominals, such as *dog*. Often, a determination of the informational or physical readings of these nouns is

In (i), *their* can only refer to *the Phillies*. We will introduce an account for ‘agentive’ exceptions to the definiteness restriction on extraction in section 4.

<sup>9</sup> In Fu et al. (2001), process nominal NPs are claimed to contain VP projections. While such a proposal explains a number of syntactic properties of such NPs, it is of little help in explaining the licensing of extractions out of NPs, for the simple reason that the licensing of these involves both argument participants (of process nouns) and non-argument participants (of result nouns). The latter of these clearly does not involve any NP-internal VP.

dependent on the nature of the verb which selects it as a complement, as is demonstrated in (36):

- (36) Which presidents do children usually  
read/hate/buy/\*shelve/\*soil books about?

In (36), *which presidents* is extracted from the indefinite NP *books about*. The extraction is licensed when the main verb determines an informational reading for *book*, but not when it determines a concrete (physical object) sense for the noun. Note that it is not the verb itself that licenses extraction, but rather the sense of the object noun induced in the verbal context. The noun *books* in *buy books* can denote either the physical or metaphysical sense of the word. This distinction becomes quite clear in certain discourse circumstances. If one points to a book that you own and says “I would like to buy that book”, it is ambiguous in that one might wish to buy the actual physical object that you possess or might simply wish to buy their own copy of the metaphysical object that your book is a token of. It is in the latter case that extractions such as in (36) with the verb *buy* are grammatical.

That extraction is dependent on noun meaning, rather than on the choice of the selecting verb, is clear when aspectual verbs, such as *finish*, are considered. Verbs such as *finish* require that their complement denote an event or activity:

- (37) Joan finished a book about the war.

If, as in (37), the noun *book* is inserted as the complement of *finish* without additional context, we construe Joan to have finished reading, writing, or perhaps editing the book, thereby referring to the informational denotation of the noun. Not surprisingly, extraction is permitted, but subject to definiteness effects (as is expected of result nouns in class II):

- (38) What did Joan finish a book about?

While it is significantly harder to coerce (37) to mean that she finished dusting, shelving, or repairing the book (which requires reference to the noun’s physical denotation), it is indeed possible. Imagine, for instance, that Joan was asked to repair the bindings of several books. She carefully tended to a three-volume collection of Shakespeare’s works, and then tackled a large volume about World War II. After she finished the book about the war, she came in and asked what else needed to be done. However, even in this rich context, we could not ask (38) with *finished*

meaning *finished repairing*. Modifiers can also provide context that affects a noun's denotation. For instance, the adjective *scandalous* leads to an informational denotation for the noun *article*, while the adjective *coffee-stained* induces a concrete noun denotation. In (39), we see how this affects extraction:

(39)a. Who are you reading a scandalous article about?

b. \*Who are you reading a coffee-stained article about?

In our account, then, *wh*-extraction out of an NP headed by a representational noun, such as *book*, is licensed only when the noun carries its informational reading. This distribution is illustrated in (40):

(40)

<i>I. complex event nominals</i> (have argument participants)	<i>II. result nominals</i> (have non-argument participants)	<i>III. concrete nominals</i> (have no participants)
----- <i>examination</i> (process reading)	----- <i>examination</i> (result reading) <i>book</i> (informational reading)	----- <i>book</i> (physical reading)
extraction of participants permitted	extraction of participants permitted definiteness effects observed	extraction not permitted

### 3.4. *The Place of Participants in Lexical Representation*

A formal account of these extraction facts can be stated with reference to Grimshaw's notion of 'nominal argument structure'. According to Grimshaw, complex event nominals such as *examination* (in its 'process' meaning) have a complex argument structure that is derived by embedding the argument structure of the related verb (*examine* in this case) in an Ev (event) argument of the nominal affix, as in (41). In this way, the argument participants of the verbal base directly map to argument participants of the derived nominal:

(41) examine *V*, (*x*(*y*)) + -ation *N*, (Ev) =  
examination *N*, (Ev (*x*(*y*)))

Result nominals, on the other hand, involve an external argument 'R', which binds an argument of the related verb, as in (42). Here, the argument

participants of the verbal base remain as participants in the LCS of the derived nominal (and may thereby be bound by *R*):

- (42) examine *V*, (*x*(*y*))  
 examination *N*, (*R* = *x*) such that *y* examines *x*

However, although the argument structure of the verbal base is related to the argument structure of the noun, it is not incorporated into it, as it was in (41). Concrete nominals have the same external argument '*R*', but have no related verbal participant structure and no participants:

- (43) dog *N*,(*R*)

As we have already noted, the distinction between nouns that involve participants in their LCS and those that do not turns out to be a critical one in accounting for the extraction possibilities. Simply put (setting aside definiteness effects), only participants may be extracted.<sup>10</sup>

Having said this, it must be pointed out that it is not clear what counts as a participant in the case of result nominals. Sometimes, the extractable constituent is clearly a participant related to the A-structure of the verbal base as in (44). The nominals in (44) are unambiguously result nominals, since (as Grimshaw 1990, p. 54 notes) indefinite determiners (e.g., *a*(*n*)), numerals (e.g., *one*), and demonstratives (e.g., *that*) are compatible only

<sup>10</sup> Precisely why only participants can extract from NPs remains somewhat mysterious. However, the restriction recalls the argument/adjunct asymmetry in subadjacency violations. In (i), both *wh*-traces violate locality restrictions (i.e., the ECP). The contrast in grammaticality is due to the fact that the trace of *what* is thematically governed (and the interpretation is lexically recoverable):

- (i)a ??What<sub>1</sub> do you wonder whether Joseph bought *t*<sub>1</sub>?  
 b. \*Why<sub>1</sub> do you wonder whether Joseph bought a new car *t*<sub>1</sub>?

Thus, in (i), a general syntactic constraint on antecedent-trace relations is overridden by the lexical properties of the verb. In (39), the situation is similar. The *about* clause is no more central to the physical denotation of the noun *article* in (39b) than is any other modifier (e.g., *in a red binder* in the *article in a red binder*). In (39a), however, the *about* clause is centrally linked to the informational denotation of *article*, and the syntactic restriction on extraction out of NP is weakened. Together, this evidence suggests that the lexical conceptual structure of words can play a part in overriding syntactic restrictions even when this structure is not syntactically projected (in a formal sense). Of course, teasing out all the differences between projected and unprojected participant structure lies well beyond the scope of this article.

with result nominals, not complex event nominals, and “complex event nominals do not pluralize, while event nominals do”.

- (44)a. Which patients did you schedule examinations of \_\_ ?
- b. Which moon did NASA film one rotation of \_\_ ?

The LCS of the result nouns in (44) can reasonably be given as in (45), and the extracted element identified with a argument of the verbal base:

- (45)a. examination  $N$ , ( $R = x$ ) such that  $y$  examines  $x$   
*which patients* in (44a) corresponds to participant “ $x$ ”
- b. rotation  $N$ , ( $R = x$ ) such that  $x$  rotates  
*which moon* in (44b) corresponds to participant “ $x$ ”

However, it is not always the case that result nouns can be readily identified with a verbal base. Consider (46):

- (46)a. Which team did they hope to achieve a victory over \_\_ ?
- b. What topic do most of the students write essays on \_\_ ?

Clearly, the extracted elements in (46) do not correspond to participants that can be identified with the arguments of a verbal base. In (46a), the result noun *victory* has no verbal base, although the extracted element might be identified with the patient of the related expression *win over* (that is, the ‘ $x$ ’ in *victory over  $x$*  is identifiable with the ‘ $x$ ’ in ‘ $y$  wins over  $x$ ’). In (46b), the result noun *essay* has no verbal base, and the extracted element is identified with the content of the essay. As (47) shows, the extracted element can also be the agent of the related verb:

- (47)a. Which team were they hoping for victories by \_\_ ?
- b. Which writer did the students read lots of essays by \_\_ ?

In (47a), the extracted element could be construed as the agent of the related verb *win*, and in (47b), the extracted element is understood as the agent of the related verb *write*.

At the same time, (48) and (49) illustrate that extractions are not always licensed for PP modifiers of result nouns:

(48)a. A victory on land was what they hoped for, having lost the support of their navy.

b. \*What location were they hoping for a victory on \_\_ ?

(49)a. The students read many essays in Spanish on the Mexican Revolution.

b. \*What language did the students read many essays in \_\_ ?

In this regard, it seems reasonable to adopt a broader notion of participant than that articulated in Grimshaw's original formulation. If we understand result nominals to denote an entity or state that arises as the outcome of some event, then we can restrict participants of result nominals to the necessary participants of the related event. In this view, victory and essay would have the following participant structure in their LCS:<sup>11</sup>

(50)a. victory  $N$ , ( $R = x$ ) such that  $x$  results from  $y$  winning over  $z$

b. essay  $N$ , ( $R = x$ ) such that  $x$  results from  $y$  writing about  $z$

In the case of representational nouns, participant structure is associated with one meaning (i.e., its informational LCS) and not with the other (i.e.,

<sup>11</sup> The modifications to LCS suggested here bear more than a passing resemblance to what Pustejovsky (1991) terms the 'qualia structure' of a lexical item. These include: (1) the relation between it and its constituent parts; (2) that which distinguishes it within a larger domain; (3) its purpose and function (telic); and (4) whatever brings it about (agentive). The last two are, as noted here, its telic and agentive aspects. In its informational sense, a *book* is created through an act of writing/composition (the agentive part of its qualia structure), and experienced through a process of reading (the telic part of its qualia structure). In its physical sense, a *book* is an object having pages, a cover, and perhaps a dust jacket. In this sense, the telic and agentive aspects of qualia structure are not as salient. De Kuthy (2000) incorporates Pustejovsky's notions of qualia structure into an analysis of NP-PP split constructions in German.

As useful as Pustejovsky's notions are generally, they do not provide (for us) a precise enough basis on which to make predictions. Consider, for example, the notion of 'telic' structure. While both result and concrete nominals may have a purpose or function (a telic aspect), only result nominals permit extraction. To illustrate, the telic (purpose and function) aspect is clearly referenced in the complements or adjuncts of the nouns in (i):

(i)a. an office of admissions

b. a table for computers

c. a comb for fine hair

its physical LCS). Thus, in (50b), the ‘*x*’ which results from an event of writing is an informational result, not a physical one. And, as we have seen, the object of *about* only counts as a participant when the informational sense is indicated:

- (51) Which subject did the students read/\*shred lots of essays about \_\_\_?

This reference to participant structure accounts for the grammaticality differences observed in (51), in which the complement of the verb *read* is informational, but not the complement of the verb *shred*. As long as the informational LCS is determined by a selecting verb, extraction is licensed.

#### 4. SYNTACTIC DETERMINANTS OF EXTRACTION POSSIBILITIES

Having established the semantic determinants for extraction out of NPs, we now turn our attention to syntactic considerations. We have seen that participant structure is necessary to license any extraction at all out of NPs, and that such extraction is subject to definiteness effects. However, definiteness effects can in fact be overridden for result nominals when they are complements of verbs of creation (e.g., *write* for *book*, *tell* for *joke*, *paint* for *portrait*, etc.). Consider the contrasts in (52)–(54):

- (52)a. Who did you write/??read those essays about?  
       b. Who did you write/read essays about?
- (53)a. Who did you tell/\*hear those jokes about?  
       b. Who did you tell/hear jokes about?
- (54)a. Who did you paint/??see that portrait of?  
       b. Who did you paint/see a portrait of?

Regardless of the reference to ‘telic’ structure, none of these NPs allows extraction, as seen in (ii):

- (ii)a. \*What did he visit the office of \_\_\_?  
       b. \*What did he move a table for \_\_\_ into the office?  
       c. \*Which kind of hair did she buy a comb for \_\_\_?

These facts suggest to us that a refined notion of ‘participant’ structure comes closer to the mark in explaining extraction out of NPs.

In each case, when a verb of creation such as *write*, *tell*, or *paint* selects a result nominal as its complement, *wh*-extraction becomes possible even out of a definite DP. This effect was noted previously in (12), repeated here, where *wh*-extraction is possible across a genitive pronoun denoting the agent role, but not across one denoting possession:

- (12) Who did Tom<sub>1</sub> write his<sub>1,\*2</sub> story about?

Examples like (12) are perfectly grammatical when the genitive (*his*) both takes the subject (*Tom*) as its antecedent and carries an agentive, rather than a possessive, meaning. Similarly, in (55), *finish* can mean only *finish writing* and not *finish reading*:

- (55) Who did Tom<sub>1</sub> finish his<sub>1,\*2</sub> article about?

While NP-internal agentivity is necessary to license *wh*-extraction out of a definite DP, it is not sufficient, as (56) shows:

- (56)a. \*Who did Tom<sub>1</sub> revise his<sub>1,2</sub> article about?  
 b. \*Who did Tom<sub>1</sub> send them his<sub>1,2</sub> article about?

In other words, both NP-internal agentivity and a selecting verb of creation are needed to license *wh*-extraction out of a definite DP.

In the account that follows, we will assume that definiteness does indeed block *wh*-extraction generally, and that this effect is syntactic. Adopting a version of the strong/weak determiner hypothesis of Bowers (1987), we assume that DP (as opposed to NP) blocks *wh*-movement categorically, and that definite (strong) determiners are D-heads, while indefinite (weak) determiners are not, as shown in (57):<sup>12</sup>

- (57)a. [<sub>DP</sub> those [<sub>NP</sub> books about Nixon]]  
 b. [<sub>QP</sub> some [<sub>NP</sub> books about Nixon]]

---

<sup>12</sup> We part with Bowers on the status of universal quantifiers (such as *every*), which he includes in the class of strong determiners:

- (i)a. Who did Sharon read a book about?  
 b. ?Who did Sharon read every book about?  
 c. \*Who did Sharon read that/the/my book about?

We do not find (ib) to be much worse than (ia), while we do think that both (ia) and (ib) contrast markedly with (ic). Based on facts such as these, we would suggest that *every* heads a QP (just like *some* in (57b)). Any difference between (ia) and (ib) may therefore be a function of quantifier meanings, rather than the blocking of extraction.



Operating with this assumption, we must now explain how agentivity within the NP and a verb of creation without, combine to license the otherwise illicit extraction out of DP seen in (12). Our analysis will proceed as follows: First, we will show that NP internal agentivity exhibits binding effects, even when the agentive element is covert (i.e., PRO) (section 4.1); second, we will then show how the *wh*-extraction in (12) can be licensed by abstract (LF) incorporation (Baker 1988, p. 202) of the result nominal into the verb of creation (section 4.2); and third we will show that LF incorporation can explain similar violations of complex NP islandhood (section 4.3).

#### 4.1. NP-Internal Agentivity in Result Nominals

Fiengo and Higginbotham (1981) show that under some circumstances indefinites are transparent while definites create opaque domains. Thus the clitic pronoun *'im* cannot be bound in its own clause (58a), but can be in (58b) (= Fiengo and Higginbotham 1981, (32) and (34) respectively):

(58)a. \*John<sub>i</sub> read books about'im<sub>i</sub>.

b. John<sub>i</sub> read that book about'im<sub>i</sub>

In (58b), the definite determiner *that* creates an opaque domain that shields the clitic pronoun from incurring a 'condition B'-type violation. However, when the same result NP is selected by its verb of creation, entailing that John is the writer of *that book*, then coindexation is again prohibited, as shown in (59):

(59) \*John<sub>i</sub> wrote that book about'im<sub>i</sub>.

We claim that the presence of a covert agentive PRO in (59) results in a binding theory violation. To get a clearer understanding of the contrast between (58b) and (59), examine the sentences in (60):

(60)a. Why did you send me that letter about yourself/you?

b. Why did you write me that letter about yourself/\*you?

In (60a), the definite deictic determiner creates an opaque domain and the subject of the sentence does not locally bind the lower instance of *you*. However, note that both the pronoun and the reflexive are possible, suggesting that more is going on here than Fiengo and Higginbotham initially suggested. Crucially, we find that the interpretation of (60a) changes in

accordance with the choice of either *yourself* or *you*. If *yourself* is selected, we understand *you* to be the *author* of *that letter*. If *you* is selected, then we understand *that letter* to have been written by someone else. These interpretations and their association with the choice of reflexive anaphor or pronoun correspond to the overt alternation shown in (61):

- (61)a. Why did you send me your letter about yourself/\*you?
- b. Why did you send me his letter about \*yourself/you?

Chomsky (1986b) makes much the same observation with respect to indefinites. He shows that in the following context both the anaphor and the pronoun may be coindexed outside the NP:

- (62)a. The children heard stories about each other.
- b. The children<sub>i</sub> heard stories about them<sub>i</sub>.

In (62a) *each other* is bound by *the children* as expected, given the standard formulation of the binding theory. To account for the coindexation of *them* with *the children* in (62b), Chomsky posits a PRO in the determiner which is disjoint in reference from *the children*. Chomsky examines cases that are somewhat more analogous to the situation in (60):

- (63)a. We felt that any criticisms of each other would be inappropriate.
- b. We felt that any criticisms of us would be inappropriate.

Chomsky proposes that in (63a) a PRO is coindexed with *we*, with the interpretation that *we* is the agent of *criticisms*, while in (63b) the PRO is disjoint from *we*, with the interpretation that some third party or parties is the agent of *criticisms*.

Returning to the example in (60), we propose then that the DP in (60a) does indeed create an opaque domain for binding and, following Chomsky, that the choice of reflexive anaphor or pronominal depends on the index of an agentive PRO within the NP. The structure for (60a) is given here in (64), where PRO corresponds to the agent of *letter* (i.e., its creator). When PRO is coindexed with the subject of *send*, then *yourself* is selected, as in (64a). When PRO is not coindexed with *you* in the higher clause, then *you* is selected within the NP, as in (64b). On this analysis,

the ungrammaticality of (60b) with *you* is also explained. Its structure is shown in (65):

- (64)a. Why did you<sub>1</sub> send me [DP that [NP PRO<sub>1</sub> letter about yourself<sub>1</sub>/\*you<sub>1</sub>]]?
- b. Why did you<sub>1</sub> send me [DP that [NP PRO<sub>2</sub> letter about \*yourself<sub>1</sub>/you<sub>1</sub>]]?
- (65) Why did you<sub>1</sub> write me [DP that [NP PRO<sub>1,#2</sub> letter about yourself<sub>1</sub>/\*you<sub>1</sub>]]?

In (60b), the subject of *write* and the agent of *letter* must be identical. Not coindexing the subject of *write* and PRO is anomalous (i.e., # PRO<sub>2</sub>), and only the reflexive *yourself* is possible within the NP. Thus, the ungrammaticality of (59) arises as a result of a DP-internal condition B violation. As we shall see below, a syntactic account of *wh*-extraction out of definites will rely to some extent on the presence of a controlled agentive PRO within NP.

#### 4.2. *Abstract Noun Incorporation and Extraction across DP*

Baker (1988) points to cases of what he calls ‘Noun reanalysis’ in Chichewa and Chamorro. In these languages, under certain conditions, the head of an object NP undergoes abstract (i.e., LF) noun incorporation (NI) with the verb. According to Baker (1988, p. 269), “N Reanalysis cases ... [are] instances of Noun Incorporation, but without the morphological incorporation”.<sup>13</sup> Baker applies NI to what have been otherwise analyzed

<sup>13</sup> The mechanism of ‘abstract incorporation’ has antecedents and analogues in other theories of grammar. Aissen and Perlmutter (1976) introduced the notion that morphologically autonomous heads might merge at an abstract level of grammar, calling this merger ‘clause reduction’ (or ‘union’). Aissen and Perlmutter take ‘clitic climbing’ as evidence for clause reduction in Spanish. In (i) the cliticized argument of the complement infinitive, *las*, can appear either affixed to the complement verb or in the matrix clause:

- (i) Pedro quiere comer*las*/Pedro *las* quiere comer.  
*Pedro wants to eat them.*

Aissen and Perlmutter show that clause reduction is triggered only by a subset of complement-taking verbs in Spanish. What is important about the second sentence in (i) is that *Pedro las quiere comer* behaves in many ways as a single clause, even though the two verbs are not necessarily merged at PF. Davies and Rosen (1988) further formalized the RG account of this construction as “predicate union”.

as ‘possessor raising’ constructions (see Kimenyi 1980). An example of this alternation from Chichewa is shown in (66):

- (66)a. Fisi a-na-dy-a nsomba z-a kalulu  
 hyena *SP-PAST-eat-ASP fish AGR-of hare*  
 (=Baker 1988, p. 271 (106a))

The hyena ate the hare’s fish.

- b. Fisi a-na-dy-er-a kalulu nsomba  
 hyena *SP-PAST-eat-APPL-ASP hare fish*  
 (=Baker 1988, p. 271 (106a))

The hyena ate the hare’s fish.

The head noun *nsomba* ‘fish’ in (66b) undergoes abstract NI, which involves the coindexation of the verb ‘eat’ and the head noun ‘fish’ as shown in (67):

- (67) [<sub>S</sub> hyena [<sub>VP</sub> [<sub>V</sub> eat]<sub>1</sub> [<sub>NP\*</sub> [<sub>NP</sub> hare] [<sub>N</sub> fish]<sub>1</sub> ]]]  
 (= Baker 1988, p. 276 (114b))

Abstract NI is claimed to have the same properties as overt NI with respect to government, movement, and Case properties. The coindexation of ‘eat’ and ‘fish’ precludes the need for the object NP\* to be “linked to the verb’s Case feature”. The coindexation further makes the possessor NP ‘hare’ visible to the verb ‘eat’, allowing it to govern and Case mark this NP.

Considering Baker’s account, we propose that exceptions to the definiteness restriction on *wh*-extraction in English involve abstract Noun Incorporation. This operation is licensed when the following conditions are met: (i) the head noun is a result nominal; (ii) the result nominal is the complement of a causative verb semantically linked to the denoted result (e.g., *write-book*, *paint-portrait*, etc.); and (iii) the subject of the

It is worth noting, therefore, that the construction *write a book* bears more than a passing similarity to a case of ‘causative union’ (Davies and Rosen 1988, pp. 54–59). The verb *write* is a “causative” verb with respect to its complement, and the noun *book* is a result. The most salient difference between these cases and that of Romance clause-reduction causatives is that in the English cases the complement is a noun instead of a verb. However, as in the Romance case and other examples of predicate union, the abstract incorporation that we have proposed for English is triggered by a small number of verbs. Of importance here is that our analysis does not crucially depend on the specific implementation of abstract incorporation we have pursued. It is compatible with any theory of grammar that can successfully handle the array of facts accounted for in predicate union analyses.

verb controls the agentive subject of the result noun. Under this analysis, the grammatical interpretation of (12) would have the structure given in (68):<sup>14</sup>

- (68) Who<sub>3</sub> did Tom<sub>1</sub> [<sub>VP</sub> write<sub>2</sub> [<sub>DP\*</sub> his<sub>1</sub> [<sub>NP</sub> PRO<sub>1</sub> story<sub>2</sub> about t<sub>3</sub>]]]

Here, the coindexation of *write* and *story* indicates the abstract incorporation of the latter with the former. Just as the presence of a DP projection creates an opaque domain for government generally (as we saw in section 4.1), the abstract incorporation of the noun *story* makes this domain again transparent (in accordance with the Government Transparency Corollary<sup>15</sup>). Thus, DP\* in (68) is no longer a blocking category for *wh*-extraction, once incorporation has applied. Abstract incorporation is restricted by the conditions set forth above, as the following data will show. First, only result nominals may incorporate. In (69), we see that certain verbs of ‘creation and transformation’, such as *develop*, can take as their object either a result nominal (69a) or a nominal denoting the ‘raw material’ (69b) (Levin 1993, pp. 172–173):

- (69)a. Tom carefully developed his story about the bombing (from his intense feelings)
- b. Tom carefully developed his feelings about the bombing (into an intense story)

Notice that extraction is possible only out of the result complement of the verb headed by *story* (70a), and not out of the ‘raw material’ object headed by *feeling* (70b):

- (70)a. What did Tom develop his story about?

<sup>14</sup> To account for so-called ‘split NP-PP’ data in German, Müller (1991) also proposes an abstract N-V incorporation analysis, making the assumption that NP is an absolute barrier to extraction and this abstract incorporation circumvents this restriction. But Müller’s analysis comes under attack from Fanselow (1991), de Kuthy (2000), and others as being unable to account for all of the data and being unable to account for lexical idiosyncrasies and the separation of PPs from subject NPs. While examination of all the German data are beyond the scope of the present work, it appears that taking DP rather than NP as being the strict barrier to extraction and recognizing the role of the notion participant might account for the majority of these structures. Bauer (2000) represents a promising move in this direction.

<sup>15</sup> The Government Transparency Corollary: A lexical category which has an item incorporated into it governs everything which the incorporated item governed in its original structural position (Baker 1988, p. 64).

- b. \*What did Tom develop his feelings about?

Second, the main verb must be a causative verb linked to the creation of the result nominal object. In (71), when *Sharon* and *her* are coindexed, we may understand Sharon to be the identified as writer/creator of the article:

- (71) When did Sharon<sub>1</sub> write/copy-edit/sell her<sub>1</sub> article about herself<sub>1</sub>?

However, notice that *wh*-extraction is licensed only with the first of the three verbs in (71). We see this in (72):

- (72)a. What did Sharon write her article about?

- b. \*What did Sharon sell/copy-edit her article about?

The contrast in (72) shows that abstract NI is restricted to cases in which the result nominal is selected by its associated verb of creation; *article* is a result nominal created through an event of ‘writing’. Neither *sell* nor *copy-edit* have anything to do with bringing an article into existence.

Finally, the subject of the relevant verb must control the agentive PRO in order for noun incorporation to be licensed. We can illustrate this with the data in (73):

- (73)a. Tom<sub>1</sub> composed/sang/performed his<sub>1</sub> song about mussels.

- b. Tom composed/sang/performed my song about mussels.

Note first that ‘singing a song’ necessarily involves performing it and also may (but need not) involve composing it. Thus, *sing* may (optionally) be a verb of creation for the result nominal *song*. In (73a), if Tom *sang his song about mussels*, he certainly performed it and may also have composed it. In (73b), on the other hand, if Tom *sang my song about mussels*, he performed it but did not compose it. Now observe (74):

- (74)a. What did Tom sing his/\*my song about?

- b. What did Tom compose/sing/\*perform his song about?

The *wh*-extraction in (74) is licensed only when Tom is the composer and creator of the song. That is, he must be both the subject of an appropriate verb of creation and control the agentive PRO in the NP headed by *song*. In

(74a), extraction is impossible with *my*, because the subject of *sing* does not control the agentive PRO in the NP headed by *song*:

- (75) What<sub>2</sub> did Tom [<sub>VP</sub> sing [<sub>DP</sub> my<sub>1</sub> [<sub>NP</sub> PRO<sub>1</sub> song about *t*<sub>2</sub>]]]

The structure in (75) is ill-formed even though the complement is a result nominal and the selecting verb is an appropriate verb of creation. The failure of *Tom* to control PRO in (75) prevents the application of abstract NI, leaving the DP opaque to government from outside and preventing licit *wh*-movement.

#### 4.3. Other Cases of Abstract NI (CNPC Violations)

The part played by abstract NI in licensing extraction is further illustrated by some striking data first noticed by Ross (1967). The examples in (76) (inspired by Ross but adjusted for inflation) illustrate the surprising fact that even the Complex NP Constraint can be violated:

- (76)a. \*The money which I am discussing the claim that the company squandered amounts to \$2,050,694.11.  
 b. ?The money which I am making the claim that the company squandered amounts to \$2,050,694.11.

What we find in (76a) is what we expect, namely, that extraction out of the complex NP is prohibited. (76b), which is grammatical (though degraded for some speakers), differs from (76a) only in containing the verb *make* instead of *discuss*. Notice, however, that (76b) contains all of the conditions outlined above as necessary to license abstract noun incorporation: (i) *claim* is a result nominal; (ii) the result nominal is the complement of a causative verb semantically linked to the denoted result, *make*; and (iii) the subject of the verb controls the agentive subject of the result noun, here a PRO coindexed with the subject *I*. The relative clause in (76b) would then have the structure in (77):

- (77) the money which<sub>3</sub> I<sub>1</sub> am making<sub>2</sub> [<sub>DP</sub>\* the [<sub>NP</sub> PRO<sub>1</sub> claim<sub>2</sub> that the company squandered *t*<sub>3</sub>]]

Just as in the cases in the previous section, abstract noun incorporation renders the DP a transparent domain and DP\* is no longer a blocking category for *wh*-movement. From our current perspective then, (76b) should actually be expected and not surprising. Data paralleling Ross' observation exist for other complex NPs:

- (78)a. Who did Kerry start the rumor that Kelsey is fond of?

- b. \*Who did Kerry hear the rumor that Kelsey is fond of?

In (78a), *start* is a causative verb that results in the noun *rumor* and the subject *Kerry* controls the PRO agent of *rumor*; thus abstract incorporation can take place and the *wh*-phrase who can move out of the complex NP. Conversely, the conditions for incorporation are not met with the verb *hear* (78b) and movement results in ungrammaticality.

The fact that these complex NPs are rendered transparent for movement under the incorporation analysis leads to a further prediction. Given the control of PRO by the subject of the sentence, we should find the same binding facts within complex NPs that we found with result nominals in section 4.1. That is, it should be possible for an anaphor within the complex NP to be bound by an antecedent outside of it. This is indeed the case, as (79) shows:

- (79) They<sub>1</sub> started [<sub>DP</sub> the [<sub>NP</sub> PRO<sub>1</sub> rumor that pictures of themselves<sub>1</sub>/them<sub>\*1/2</sub> were hanging in the post office]].

When the verb is the causative verb *start*, *themselves* appears to be bound by the subject, outside of the complex NP. This is, of course, because *they* controls the agentive PRO of *rumor*, which binds the reflexive within the DP. The impossibility of coindexing *them* with the subject is a Principle B violation; *them* is bound by PRO within the NP.

## 5. WHY THERE STILL IS A SUBJECT CONDITION

Up to this point, we have presented a fairly comprehensive account of the conditions involved in the extraction shown in (1a); we turn now to the ungrammaticality of (1b). Asking the obvious question of whether the same mechanism that amnesties extraction from objects can apply to subjects in English, we find that the answer is clearly ‘no’, given the subject-object asymmetry so long noted in the literature. This is illustrated again in (80), where the same semantic environment allows extraction from objects but not subjects.

- (80)a. Who is the class reading a book about?  
b. \*Who is a book about being read by the class?

It is the uniform ungrammaticality of extraction from subjects which led to the formulation of the Subject Condition and its descendants. But is there still a Subject Condition? As we will see, the answer is ‘yes’.



Chomsky (1986a) attempted to do away with such a device, proposing that subjects and adjuncts are islands for the same reason, that is, they are not L-marked. However, Culicover (1997) has shown that adjuncts are not uniformly islands, as we can see in the following data:

- (81)a. \*who<sub>1</sub> did [she go to Harvard [because she wanted to work with t<sub>1</sub>]]
- b. who<sub>1</sub> did [she go to Harvard [in order to work with t<sub>1</sub>]]  
(Culicover 1997, p. 253)
- c. ?what<sub>1</sub> did [he finish his thesis [without checking t<sub>1</sub>]]
  
- (82)a. \*what<sub>1</sub> is [[that she wanted to learn t<sub>1</sub>] unfortunate]
- b. \*what<sub>1</sub> is [[to learn t<sub>1</sub>] very hard]
- c. \*what<sub>1</sub> is [[checking t<sub>1</sub>] absolutely necessary before you turn in your thesis]

(81) illustrates three varieties of adjuncts, while in (82) the same phrases are subjects. As (81b,c) show, extraction is possible from some adjuncts while extraction is uniformly ruled out from subjects. Given this robust difference, some of the motivation for the *Barriers*-style account of the Subject Condition is lost.

We suggest that the key to explaining subject islandhood resides in earlier structural accounts. Specifically, we argue that the analysis of sentential subjects adopted in Chomsky (1973) is essentially correct. Under this analysis, sentential subjects are islands through violation of Subadjacency.

While Chomsky's (1973) bounding nodes solution worked for sentential subjects, a purely structural account of the contrast in (80) must face the fact that the arguments in question are NPs already. However, we have shown above that DP can be assumed to be an absolute barrier to extraction in the absence of abstract noun incorporation, while NP cannot be. In fact, a strong case can be made that in English *all* subjects must be DPs. That is, while we have claimed that the object *stories about t<sub>i</sub>* in (80a) is an NP, the subject *stories about t<sub>i</sub>* in (80b) is a DP. This DP requirement is a purely structural effect that emanates from the manner in which the EPP is satisfied in English, and it is inviolable. Subject islandhood arises, on our account, from the requirement that subjects check a (D) category feature on T.

In a series of papers (Davies and Dubinsky 1999, 2000, 2001) we have argued that all English subjects, whether nominal, clausal, verbal, adjectival, or prepositional, are dominated by a DP node. The evidence for this comes from the fact that non-NP subjects, as in (83), have an identifiable set of properties that can all be reasonably classed as DP properties, including undergoing obligatory raising, triggering subject agreement, and licensing emphatic reflexives:

- (83)a. [<sub>CP</sub> that Shelby lost it] is true
- b. [<sub>PP</sub> under the bed] is a good place to hide
- c. [<sub>AP</sub> very tall] is just how he likes his bodyguards

As we argue in Davies and Dubinsky (2001), this constellation of facts provides evidence that non-NP subjects are, in fact, DPs, so that the subject in (83a) has the structure in (84):

- (84) [<sub>TP</sub> [<sub>DP</sub> D [<sub>CP</sub> that Shelby lost it]] ...]

The structure in (84) provides an explanation for the well-known fact that subjects in English are islands: movement out of a DP subject violates Subjacency. This DP-shell analysis is able to account for a variety of syntactic properties of English subjects. That is, Tense contains a D-feature which must be checked prior to Spell-out, which forces all tensed clauses to contain a DP subject. This analysis is given schematically in (85):

- (85) [<sub>ArgP</sub> Spec [<sub>Agr'</sub> Agr [<sub>TP</sub> T ... ]]]
- ↑  
└─ [+D] ─┘

As we argue, especially in Davies and Dubinsky (1999), this DP-node accounts for sentential subject and subject island facts simply as Subjacency violations. In (86a), the DP node dominating the CP subject results in movement violating Subjacency. Conversely, the sentential complement in (86b) is a simple CP; thus, movement in this environment does not violate Subjacency:

- (86)a. \*what<sub>i</sub> do you think [<sub>DP</sub> [<sub>CP</sub> t'<sub>i</sub> that John lost t<sub>i</sub>]] is a tragedy?
- b. what<sub>i</sub> do you think Judy regrets [<sub>CP</sub> t'<sub>i</sub> that John lost t<sub>i</sub>]?

## 6. CONCLUSION

In the foregoing discussion, we have demonstrated that extraction from NPs is determined both by semantic and syntactic considerations.

Extraction from object NPs turns out to depend on the extracted element being linked to an argument in the LCS of the noun. This means that complex event nouns and result nouns both permit extraction, albeit to a different extent (as contrasted with the class of concrete nouns, which do not). Complicating this picture is the fact that many nouns have membership in two classes, and that context plays a role in determining the denotation of a noun in a particular instance. Thus, alongside nouns such as *examination*, which can serve as a complex event noun or a result noun, we find that nouns such as *book* can have an informational or physical sense, which correlates with its dual membership in the result and concrete classes, respectively. It is nouns such as *book* for which class membership determines the possibility of extraction. When used as a result noun, *book* allows for the extraction of participants. When used as a concrete noun, it has no participants (in the formal sense) and does not permit extraction at all.

On the syntactic side, we find that the prohibition on extraction out of definites can be overridden in a very narrow class of cases. When the result noun is complement of a causative verb semantically linked to the denoted result, then extraction of a participant is licensed. Taking DP (but not NP) to be an absolute barrier to extraction, we proposed that N-V incorporation is triggered by these special conditions and that N-V incorporation voids the barrierhood of DP in these cases. Taking DP to categorically block extraction out of definites turns out to have important ramifications. In particular, we find that it provides corroboration for the syntactic nature of the subject island effect. The nature of the EPP in English is such that subject position must be filled by a DP constituent, regardless of the semantic nature of the subject phrase. Accordingly, subject islandhood is just another instance of the ‘definiteness effect’ observed for complement positions. The absence of any exceptions to the ban on extraction (as are found for objects) is a simple consequence of the impossibility of N-V incorporation from this position.

## REFERENCES

- Abney, Steven. 1987. *The English Noun Phrase in its Sentential Aspect*, unpublished Ph.D. dissertation, MIT.

- Aissen, Judith and David Perlmutter. 1976. 'Clause Reduction in Spanish', *Berkeley Linguistic Society* 2, 1–30. Revised version in David Perlmutter (ed.), *Studies in Relational Grammar*, Vol. 1, University of Chicago Press, Chicago, pp. 360–403.
- Bach, Emmon and George Horn. 1976. 'Remarks on "Conditions on Transformations"', *Linguistic Inquiry* 7, 265–361.
- Bauer, Susi. 2000. 'By This Philosopher, I Would Destroy a Book!: What Licenses PP Extraction Out of German Noun Phrases', unpublished manuscript, University of Iowa.
- Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing*, University of Chicago Press, Chicago.
- Bowers, John. 1987. 'Extended X-Bar Theory, the ECP, and the Left Branch Condition', *Proceedings of the Sixth West Coast Conference on Formal Linguistics* (WCCFL 6), The Stanford Linguistics Association, Stanford, pp. 47–62.
- Chomsky, Noam. 1973. 'Conditions on Transformations', in Stephen Anderson and Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, Holt, Rinehart, and Winston, New York, pp. 232–286.
- Chomsky, Noam. 1986a. *Barriers*, MIT Press, Cambridge, MA.
- Chomsky, Noam. 1986b. *Knowledge of Language*, Praeger, New York.
- Culicover, Peter. 1997. *Principles and Parameters: An Introduction to Syntactic Theory*, Oxford University Press, Oxford.
- Davies, William and Stanley Dubinsky. 1999. 'Sentential Subjects as Complex NPs: New Reasons for an Old Account of Subjacency', Papers from the 34th Annual Meeting of the Chicago Linguistic Society (CLS 34), Chicago Linguistic Society, Chicago, pp. 83–94.
- Davies, William and Stanley Dubinsky. 2000. 'Functional Structure and a Parametrized Account of Subject Properties', *Proceedings of the Fifteenth Eastern States Conference on Linguistics* (ESCOL '99), CLC Publications, Ithaca, New York, pp. 48–59.
- Davies, William and Stanley Dubinsky. 2001. 'Functional Structure and a Parameterization of Subject Properties', in W. Davies and S. Dubinsky (eds.), *Objects and Other Subjects: Grammatical Functions, Functional Categories, and Configurationality*, Kluwer Academic Press, Dordrecht, pp. 247–279.
- Davies, William and Carol Rosen. 1988. 'Unions as Multi-Predicate Clauses', *Language* 64, 52–88.
- de Kuthy, Kordula. 2000. *Discontinuous NPs in German: A Case Study of the Interaction of Syntax, Semantics, and Pragmatics*, unpublished Ph.D. dissertation, Universität des Saarlandes.
- den Dikken, Marcel and Anastasia Giannakidou. 2001. 'What the hell?!', *Proceedings of the North East Linguistic Society* 31, 163–182.
- Diesing, Molly. 1992. *Indefinites*, Linguistic Inquiry Monograph, Vol. 20, MIT Press, Cambridge, MA.
- Erteschik-Shir, Nomi. 1973. *On the Nature of Island Constraints*, unpublished Ph.D. dissertation, MIT.
- Erteschik-Shir, Nomi. 1981. 'On Extraction from Noun Phrases (Picture Noun Phrases)', in A. Belletti, L. Brandi, and L. Rizzi (eds.), *Theory of Markedness in Generative Grammar: Proceedings of the 1979 GLOW Conference*, Scuola Normale Superiore di Pisa, Pisa, pp. 147–169.
- Fanselow, Gisbert. 1991. *Minimale Syntax Untersuchungen zur Sprachfähigkeit*, Habilitationsschrift, Universität Passau, Passau, published as Number 32 in *Groninger Arbeiten zur Germanistischen Linguistik*, Rijksuniversiteit Groningen, Groningen.
- Fiengo, Robert and James Higginbotham. 1981. 'Opacity in NP', *Linguistic Analysis* 7, 395–422.

- Fu, Jingqi, Thomas Roeper and Hagit Borer. 2001. 'The VP within Process Nominals: Evidence from Adverbs and the VP Anaphor do-so', *Natural Language and Linguistic Theory* **19**, 549–582.
- Godard, Danièle. 1992. 'Extraction Out of NP in French', *Natural Language and Linguistic Theory* **10**, 233–277.
- Grimshaw, Jane. 1990. 'Argument Structure', *Linguistic Inquiry Monograph*, Vol. 18, MIT Press, Cambridge, MA.
- Kimenyi, Alexandre. 1980. *A Relational Grammar of Kinyarwanda*, University of California Press, Berkeley.
- Lebeaux, David. 1986. 'The Interpretation of Derived Nominals', papers from the General Session of the 22nd Regional Meeting of the Chicago Linguistic Society (CLS 22), pp. 231–247.
- Levin, Beth. 1993. *English Verb Classes and Alternations*, University of Chicago Press, Chicago.
- Müller, Gereon. 1991. 'Abstrakte Inkorporation', in S. Olsen and G. Fanselow (eds.), *DET, COMP and INFL*, Max Niemeyer Verlag, Tübingen, pp. 155–202.
- Pustejovsky, James. 1991. 'The Syntax of Event Structure', *Cognition* **41**, 47–81.
- Ross, John Robert. 1967. *Constraints on Variables in Syntax*, unpublished Ph.D. dissertation, MIT; published in 1986 as *Infinite Syntax!*, Ablex Publishing, Norwood, NJ.

Received 26 January 2001

Revised 13 January 2002

William D. Davies  
Linguistics Department  
University of Iowa  
Iowa City, IA 52242  
USA  
<william-davies@uiowa.edu>

Stanley Dubinsky  
Linguistics Program  
University of South Carolina  
Columbia, SC 29208  
USA  
<dubinsky@sc.edu>

