Topics in Ellipsis

Edited by
Kyle Johnson
University of Massachusetts, Amherst
Contents

1 Introduction
   Kyle Johnson  page 1

2 VP Ellipsis and constraints on interpretation
   Daniel Hardt  15

3 Direct compositionality and variable-free semantics:
   the case of Antecedent Contained Deletion
   Pauline Jacobson  30

4 The view of QR from ellipsis
   Kyle Johnson  69

5 Argument Contained Ellipsis
   Christopher Kennedy  95

6 Variable island repair under ellipsis
   Jason Merchant  132

7 On binding scope and ellipsis scope
   Winfried Lechner  154

8 The silent content of bound variable pronouns
   Uli Sauerland  183

9 A step-by-step guide to ellipsis resolution
   Satoshi Tomoki  210

10 Shared constituents and Linearization
    Chris Wilder  229

Notes  259
References  285
Index  300
6 Variable island repair under ellipsis

Jason Merchant
University of Chicago

One of the most startling, and hence theoretically challenging, properties of \textit{wh}-movement in Sluicing is that it can move \textit{wh}-phrases out of islands, an important observation which goes back to Ross (1969). Equally challenging is the fact that similar \textit{wh}-movement out of VP Ellipsis sites remains for the most part illicit. Briefly put, it seems that for a wide range of cases, deletion of an IP containing an island voids the effect of that island for \textit{wh}-movement, while deletion of a VP containing an island does not. This chapter investigates one aspect of this puzzling dichotomy with respect to island repair, and attempts to show that an interesting and partly novel range of data follow if island deviations come about due to illicit traces of intermediate movement, working in tandem with a constraint on ellipsis operative in structures that host \textit{wh}-movement. I will argue that a wide range of islands are indeed active at PF, but not in the way that this claim has usually been understood thus far. Instead of the island node itself being responsible for the degradation in acceptability, I will show that the data support the idea that it is the traces of \textit{wh}-movement outside the island itself which trigger a PF-crash.

I begin with some relevant Background on Sluicing, given in section 6.1, before taking up the Sluicing data in section 6.2 and the VP Ellipsis facts in section 6.3.

6.1 Background on Sluicing

Sluicing is ellipsis of the sentential complement to an interrogative complementizer hosting a \textit{wh}-phrase, as in (1):

(1) a. Jack bought something, but I don't know what.
   b. A: Someone called. B: Really? Who?
   c. Beth was there, but you'll never guess who else.
   d. Jack called, but I don't know when/how/why/where from.

These structures have been extensively investigated in the literature (the most detailed studies being Ross 1969; Levin 1982; Riemsdijk 1982; Chao 1987; Chung et al. 1995; Lobeck 1995; Romero 1998; Lasnik 2001; and Merchant 2001), and the most common approach to generating these structures is to assume that the \textit{wh}-phrase has been moved to specCP out of the missing IP by the usual mechanisms of \textit{wh}-movement in questions, and that the IP has undergone deletion (ellipsis). While neither of these two components has been universally accepted, I will follow most recent work in assuming that these are the two operations that generate the sentences in (1) (see Merchant 2001 for extensive justification). The relevant derivation is given schematically in (2).

![Diagram](image)

Given this analysis, one question that immediately arises is the question of what licenses the ellipsis of IP (in the sense of licensing developed in Lobeck 1991). Lobeck (1995) claims that the null IP (a base-generated empty category in her approach) is licensed only by the null [+\textit{wh}, −pred] \textit{C}^0 of interrogatives (following Rizzi's 1990 typology). I will here recast her approach as a featural matching requirement in a head-head (or feature of a head) relation. Assume that PF deletion is triggered by the presence of a feature on a head. Let us call this triggering feature \textit{E}. Ideally, \textit{E} will have exactly those syntactic, phonological, and semantic effects that yield all the attested properties of the elliptical construction at hand, with nothing further needing to be said.

The syntax of \textit{E} must encode the checking requirements, in order to capture Lobeck's licensing. In Framepton and Gutmann's (1999) notation, the syntactic featural makeup of \textit{E} is [−[+\textit{wh}], −[+\textit{Q}]], that is, an element which must be checked by a [+\textit{wh}, +\textit{Q}] head. Such checking heads are limited in languages like English at least to the complementizers that occur in constituent questions, as desired. (Whether \textit{E} is freely available to be merged with any head, or whether it must be on \textit{C} or on \textit{I} is an independent
question: presumably locality considerations will rule out this featural combination on E from occurring anywhere where it cannot be checked by a +wh, +Q C⁰. This restricts E to C or I. If E is on C, the effects follow directly. If it is on I, then we must posit head-to-head or feature-to-head movement to check E. For present purposes, the decision between these two options is immaterial.)

The phonology of E, in broad terms, should be such that E instructs PF not to parse its complement. Assuming E to be located on C at the relevant point in the derivation (when the syntactic structure is parsed by the mechanisms operative at PF), we can view E to be something like a suprasegmental, but with the unusual effect of parsing its complement IP into a prosodically unrealized category. In SPE terms, E has the effect of triggering a context sensitive rule of the following sort, however we may wish to conceptualize this in more recent phonological theories:

(3) $[\Phi/IP] \rightarrow \alpha/E\_\_

PF-"deletion," in this view, is the result of a feature in the syntax, not of a freely operating "deletion transformation."

Finally, the semantics of E must capture the traditional identification of the ellipsis site, ideally encoding all and only those requirements that regulate under what conditions an XP can be deleted. Two general approaches to this question have been pursued: one, that the elided XP must be syntactically (LF-structurally, in current formulations) isomorphic to an antecedent, and two, that the elided XP must be semantically equivalent to an antecedent. Both views have weaknesses: generally, the syntactic isomorphism approach undergenerates, while the semantic identity approach overgenerates. I will follow here the approach advocated in Merchant (2001), where the semantic identity condition is defined as e-givenness: an XP α can be deleted only if α is e-given. Once we have defined such a predicate that applies to (the meanings of) XPs, we can use it to define a semantics for the E feature. (By a similar token, we could easily define a predicate over LF structures and impose a syntactic isomorphism condition in the same way, assuming presumably a structured meaning approach.) The semantics of E will therefore be a partial identity function on propositions, a semantic filter which allows the composition to proceed only if its argument is e-given. In the notation of Heim and Kratzer (1998), the semantics of E is the following:

(4) $[E] = \lambda p: p \text{ is e-given. } p$

A sluice like (5), then, will have the structure in (6).

(5) Abby was reading, but I don't know what.

(6) Abby was reading $t_2$

At the relevant point in the computation, E will take IP as its argument, as in (7). This expression will return a value (namely Abby was reading $x_2$) only if Abby was reading $x_2$ is e-given, otherwise it returns no value. Computation up the tree therefore proceeds only if E’s IP complement is e-given, as desired.

(7) $[E][[[IP]]] = \lambda p: p \text{ is e-given. } p(\text{Abby was reading } x_2)$

The greatest advantage of using E, encoded as a partial identity function, to impose the identity requirement is that it localizes ellipsis identification, and allows us to dispense with the more usual formulations of the requirement on ellipsis which essentially postulate a separate "ellipsis module" in the grammar (i.e. a global, late, well-formedness condition imposed just on the structures containing ellipsis) parallel to the Binding Theory module (cf. Giannakidou’s 1998; 2001 elimination of a “polarity” module by encoding polarity requirements as local, lexical semantic well-formedness conditions, using type-combinatorics).

A second important advantage is that the licensing (the local featural requirements of E) and identification (the semantic condition E imposes on its complement) requirements on ellipsis can be directly linked. Most theories posit no direct link between these requirements at all.

In sum, I will assume that a unified theory of PF-deletion based on semantic identity is possible (indeed, desirable), and that Sluicing instantiates PF-deletion of an IP out of which wh-movement has occurred.

6.2 Sluicing and wh-extraction out of islands

Assuming that wh-movement of the usual kind occurs in Sluicing brings us directly to the puzzle that has inspired much of the work on the topic since Ross (1969) first discovered it: the wh-movement found in Sluicing (at least with certain kinds of correlates) is insensitive to syntactic islands (see Baker and Braine 1972; Chomsky 1972a; Lakoff 1972; Chung et al. 1995; Lasnik 2001; Merchant 2001). Examples for the major kinds of
syntactic islands are given (occasionally with nonelliptical controls) in (8)–(16).

(8) Relative Clause island:
   a. They want to hire someone who speaks a Balkan language, but I don’t remember which.
   b. *I don’t remember which (Balkan language) they want to hire someone [who speaks ... ].

(9) Left-branch (attributive adjective case):
   a. She bought a big car, but I don’t know how big.
   b. *I don’t know how big she bought [a — car].

(10) Derived position islands (subjects, topicalizations):
   a. A biography of one of the Marx brothers is going to be published this year — guess which!
   b. *Guess which (Marx brother) [a biography of —] is going to be published this year.

(11) COMP-trace effects: (Chung et al. 1995 (90), (91a); Perlmutter 1971: 112)
   a. It appears that someone will resign, but it’s not yet clear who.
   b. Sally asked if somebody was going to fail Syntax One, but I can’t remember who.

(12) Coordinate Structure Constraint:
   a. They persuaded Kennedy and some other Senator to jointly sponsor the legislation, but I can’t remember which one. (Chung et al. 1995: (88b))
   b. Bob ate dinner and saw a movie that night, but he didn’t say which.

(13) Adjuncts:
   a. Ben will be mad if Abby talks to one of the teachers, but she couldn’t remember which.
   b. *Ben will be mad if Abby talks to one of the teachers, but she couldn’t remember which (of the teachers) Ben will be mad [if she talks to —].
   c. Ben left the party because one of the guests insulted him, but he wouldn’t tell me which.

(14) Complement to nouns: (Chung et al. 1995: (84c))
The administration has issued a statement that it is willing to meet with one of the student groups, but I’m not sure which one.

(15) Sentential subject: (Chung et al. 1995: (84b))
That certain countries would vote against the resolution has been widely reported, but I’m not sure which ones.

(16) Embedded question: (Chung et al. 1995: (84a))
Sandy was trying to work out which students would be able to solve a certain problem, but she wouldn’t tell us which one.

In (8a), for example, the wh-phrase which has moved out of the relative clause, interpretationally parallel to its unelided but ungrammatical counterpart in (8b). Similar remarks apply to the remaining islands.

One possibility for accounting for the fact that deletion of the island rescues the sluice from ungrammaticality is to posit that the PF interface cannot parse crossed island nodes. One way of formalizing this, following in essence Chomsky (1972b), is to assume that crossed island nodes are marked with some PF-uninterpretable feature, call it *. This general approach, "* as a feature of island nodes," has been pursued in one form or another by Ross (1969), Lasnik (2001), and Kennedy and Merchant (2000) (the latter for the Left Branch Condition only).

For the example in (8a), repeated in (17a) on the facing page with the structure in (17b), the account works as follows. Wh-movement (cyclic or otherwise – I suppress here for simplicity possible intermediate traces) extracts the DP which from its base position (marked by t_i), moving it to the highest specCP. In doing so, the relative clause island is crossed. As such, it is marked with *. In (17b), the island node is assumed to be the CP adjoined to the NP, and this CP is marked with *. In nonelliptical cases, when this *CP reaches PF, it will cause a PF-crash, since * is by hypothesis PF uninterpretable.

Under this formulation, ellipsis will have the desired effect: deletion of the boxed IP in (17b) eliminates the *CP as well, preventing the * from triggering a PF-crash. The structure is therefore saved, and surfaces as the grammatical sluice in (17a). This general solution applies mutatis mutandis to the other islands in (9)–(16) as well.

6.3 VP Ellipsis and wh-extraction out of islands

As appealing as the solution sketched above is, it faces a serious problem when we turn our attention to parallel extractions out of elided VPs. In the same environments, with the same correlates and the same islands, extraction out of a VP Ellipsis site is no more grammatical than in nonelliptical cases. The relevant data from VP Ellipsis, contrasting with the sluices in (8)–(16) above, are given in (18)–(25).
(17) a. They want to hire someone who speaks a Balkan language, but I don't remember which.
b. IP-Deletion eliminates *CP

(18) Relative Clause island:
*They want to hire someone who speaks a Balkan language, but I don't remember which they do.

(19) Left-branch (attributive adjective case):
*She bought a big car, but I don't know how big she did.

(20) Derived position islands (subjects, topicalizations):
*He said that a biography of one of the Marx brothers is going to be published this year – guess which he did!

(21) COMP-trace effects:
a. *It appears that someone will resign; it's just not clear how it does.
b. *Sally asked if somebody was going to fail Syntax One, but I can't remember who she did.

(22) Coordinate Structure Constraint:
a. *They persuaded Kennedy and some other Senator to jointly sponsor the legislation, but I can't remember which one they did.
b. *Bob ate dinner and saw a movie that night, but he didn't say which he did.

(23) Adjuncts:
a. *Ben will be mad if Abby talks to one of the teachers, but she couldn't remember which he will.
b. *Ben left the party because one of the guests insulted him, but he wouldn't tell me which he did.

(24) Complement to nouns:
*The administration has issued a statement that it is willing to meet with one of the student groups, but I'm not sure which one it has.

(25) Embedded question:
*Sandy was trying to work out which students would be able to solve a certain problem, but she wouldn't tell us which one she was.

In fact, the problem of *wh-extracting out of VP Ellipsis sites is even more severe, as highlighted in Lasnik (2001), where the data in (26)-(28) are given. Movement of a *wh-phrase whose correlate is an indefinite out of an elided VP is degraded even when no island is involved.

(26) They said they heard about a Balkan language, but I don't know
a. which they said they heard about. No Ellipsis
b. which. Sluicing
c. *which they did. VP Ellipsis

(27) They attended a lecture on a Balkan language, but I don't know
a. which they attended a lecture about. No Ellipsis
b. which. Sluicing
c. *which they did. VP Ellipsis

(28) They studied a Balkan language, but I don't know
a. which they studied. No Ellipsis
b. which. Sluicing
c. ??which they did. VP Ellipsis
One possibility for accounting for this range of data, including the data in (18)–(25), would be to claim simply that there is a general ban on wh-extraction out of VP Ellipsis sites, à la Sag (1976) and Williams (1977). Unfortunately, such a claim is too strong, as the following examples demonstrate.

(29) a. I know what I LIKE and what I DON'T.
    b. I know which books she READ, and which she DIDN'T.
    c. What VP Ellipsis CAN do, and what it CAN'T.

    (Johnson 2001)

(30) a. GREEK, you should take; DUTCH, you shouldn't.
    b. I know which books ABBY read, and which ones BEN did.

(31) a. I think YOU should ride the TALLEST camel, but I don't know which one PHIL should.
    b. I think you SHOULD adopt one of these puppies, but I can't predict which one you actually WILL.

    (Schuyler 2001: (48))

    (Schuyler 2001: (49))

    c. ABBY took GREEK, but I don't know what language BEN did.
    d. We know that Abby DOES speak [Greek, Albanian, and Serbian] – we need to find out which languages she DOESN'T <speak t>!

    (Merchant 2001: 115 fn. 5 (ii))

    e. (I know) ABBY wants to take GREEK, but I don't know what language BEN does <want to take>.
    f. ABBY, she took GREEK, but I don't remember what language ZETH did <say she took>.
    g. ABBY attended a lecture on KEATS, but I don't know what poet BEN did.

What distinguishes the examples in (29)–(31) from (18)–(25), (26c), (27c), and (28c) is the presence in the former of an element in the elliptical clause which contrasts with some element in the antecedent clause. The observation is simple: it appears that some kind of contrast is required in the cases where VP Ellipsis is licit (see Schuyler 2001 for one formulation). When such contrast is absent, as in (18)–(28), VP Ellipsis is disallowed.

Perhaps, as Merchant (2001) and Lasnik (2001) suggest, there is a ban on eliding less than possible under wh-extraction (whose ultimate source remains obscure). For the present, let us capture this ban in the form of an inviolable constraint, MaxElide (this constraint may be in part derivable from economy, since putting the E feature higher in a given structure allows for less pronunciation). Roughly put, it states that if ellipsis applies in a structure with a wh-trace, ellipsis should target the largest constituent possible. More accurately, it requires that if ellipsis targets an XP containing an A′-trace, XP must not be properly contained in any YP that is a possible target for deletion.

(32) MaxElide [Definition]

    Let XP be an elided constituent containing an A′-trace. Let YP be a possible target for deletion. XP must not properly contain YP (XP ⊈ YP).

The VP Ellipsis in the (c) examples of (26)–(28) violates MaxElide; in (26c), for example, the VP <say they heard about t0> contains a wh-trace and the VP is properly contained in the IP <they did say they heard about t0>, which, as (26b) shows, is itself a possible target for deletion. The possibility for deleting the containing IP, then, blocks deletion of any contained VP (likewise for the more deeply embedded VP: *They said they heard about a Balkan language, but I don’t know which they said they did.).

Parallel reasoning applies to the examples in (18)–(25), in which the deleted VP happens to contain an island: in each case no contrasting material is present, and the IP containing the deleted VP is a possible target for deletion (as witnessed by the Sluicing counterparts in (8)–(16)).

The examples in (29)–(31) differ in precisely this regard. Consider (29a): the deleted VP is <like t>, which is properly contained in the IP <I don’t like t>. But this containing IP is not a possible target for deletion: there is no antecedent which would license deletion of the sentential negation (technically, in the theory assumed here, the IP is not e-given). Hence the containing IP is irrelevant to the deletion of the VP, MaxElide doesn’t apply, and VP Ellipsis is not blocked in this case. In other words, if the material outside the VP Ellipsis site contrasts in some way with the antecedent clause, the contrasting material cannot be deleted (since it is not e-given), and hence no larger constituent will be a possible target for deletion. This contrasting material can be in the auxiliary domain (negation as in (29), (30a), or modal as in (31b)), or the subject (as in (30b), (31a, c–f)), or elsewhere external to the VP but internal to the IP (see Schuyler 2001 for further examples and discussion).

MaxElide seems also to be responsible for the unexpected oddity of examples like (33b, d) (from Merchant 2001: 58). While VP Ellipsis targeting the highest VP is fine, as in (33a), and while ellipsis is not required, as in
(33c), ellipsis targeting either the most deeply embedded IP (33b) or VP (33d) is distinctly degraded. Both these latter examples are in violation of MaxElide: taking XP in (33b) to be the matrix VP whose subject is Charlie, MaxElide is violated by the ellipsis of the embedded IP, properly contained in the matrix VP. Likewise in (33d): MaxElide prohibits deletion of the embedded VP <invited t> since the matrix VP <know who she invited t> is a possible target of deletion itself.

(33) a. Ben knows who she invited, but Charlie doesn't.
   b. ??Ben knows who she invited, but Charlie doesn't know who.
   c. Ben knows who she invited, but Charlie doesn't know who she invited.
   d. ??Ben knows who she invited, but Charlie doesn't know who she did.

Similar reasoning applies to the example in (34), from Williams (1986), where the possible deletion of the higher IP blocks deletion of the lower one:

(34) John knows how to do something, but I don't know what (*he knows how).

It is crucial to note that MaxElide applies only to XPs that contain a wh-trace, since no similar blocking effect is found in the absence of wh-movement:

(35) a. Ben knows that she invited Klaus, but her father doesn't.
   b. Ben knows that she invited Klaus, but her father doesn't know that she did.

The contrast between XPs containing wh-movement and those that do not appears in the data in (36) as well, from Merchant (2001: 89 note 9):

(36) a. Abby knew that he had quit, but Beth didn't know that he had.
   b. Abby asked if he had quit, but Beth didn't ask if he had.
   c. ?? Abby knew when he had quit, but Beth didn't know when he had.
   d. ?? Abby asked when he had quit, but Beth didn't ask when he had.

The fact that MaxElide applies only to XPs containing A'-traces also allows us to set aside the possible objection raised in Lusnik (2001) on the basis of an example with subject extraction: as he points out, the possibility for Sluicing in (37a) does not preclude the VP Ellipsis variant in (37b).

(37) a. Someone solved the problem.
   b. i. Who?
      ii. Who did?

The reason that Sluicing doesn't block VP Ellipsis here follows from the definition of MaxElide. In particular, in (37b) MaxElide has no provenance, since the elided VP does not contain a wh-trace, assuming the structure in (38) with the type of the two traces of the subject notated as superscripts.

(38) [cp who [p r' [vp r' solved the problem]]]

Constraints similar to MaxElide are discussed in Hirschbühler (1978) and Tancredi (1992: 123) for a related set of data; nevertheless, at this point, a satisfying theoretical reduction of MaxElide remains elusive. For present purposes, I will be concerned only with the obvious effects it has.

At this point, we have constructed a coherent account of some surprising differences between Sluicing and VP Ellipsis, and it might appear that the ungrammaticality of the examples in (18–25) with which this section began do not, after all, pose a difficulty for the *-as-a-feature-of-island-nodes view that accounted so elegantly for the lack of island effects in Sluicing in section 6.2. Unfortunately, this view makes the following prediction: if the island node is internal to a deleted VP and if MaxElide is satisfied (by the presence of contrasting material external to the VP and internal to the IP), then we should find that ellipsis has the same island-ameliorating effect that we find in Sluicing. This prediction is incorrect, as the following examples show (see also the examples and discussion in Merchant (2001: 114–115).

(39) Relative Clause island:
   *Abby DOES want to hire someone who speaks GREEK/a certain Balkan language, but I don't remember what kind of language she DOESN'T.

(40) Left-branch (attributive adjective case):
   *ABBY' bought a big car, but I don't know how big BEN did.

(41) Derived Position islands (subjects, topicalizations)
   *Abby DID say that a biography of HARPO is going to be published this year – guess which Marx brother she DIDN'T.

(42) COMP-trace effects:
   a. *It appears to ME that SEN. HATCH will resign, but I don't know which senator it does to YOU.
   b. *ABBY asked if I was going to fail Syntax One, but I can't remember who BEN did.
(43) Coordinate Structure Constraint:
   a. *They got the president and thirty-seven Democratic Senators
to agree to revise the budget, but I can’t remember how many
Republican ones they DIDN’T.
   b. BOB ate dinner and saw five movies that night, but he didn’t
say how many ABBY did.

(44) Adjuncts:
   a. *BEN will be mad if Abby talks to Mr. Ryberg, and guess who
CHUCK will.
      (Merchant 2001: 115 (15))
   b. *BEN left the party because Charlene/some guest insulted
him, but God only knows which guest ABBY did.

(45) Complement to nouns:
   *The dean’s office has issued a statement that it is willing to meet
with Students for a Democratic Society/a certain student group,
but I’m not sure which student group the provost’s office has.

(46) Embedded question:
   *Sandy was trying to work out how many students would be able
to solve problem #4/a certain problem, but she wouldn’t tell us
which problem she wasn’t.

To see how the account fails on these examples, consider (39). The presence
of negation in the IP beneath the wh-phrase means that MaxElide is
satisfied (vacuously). The island node, as in (17b) above, is the *CP of
the relative clause internal to the deleted VP <want to hire someone who
speaks t>. Hence the deletion of the VP prevents the *CP from reaching
the PF-interface, and this example should have the same status as its
Sluicing counterpart in (17a), contrary to fact.

Instead of thinking of * as a feature of island-nodes, I suggest we think of*
as a feature of traces. In particular, let intermediate traces of island-
escaping XPs be marked with the * feature. One way of implementing this
is to say that each link in a chain of wh-movement must be licensed either
by locality or by being in a spec-head relation with a C (or perhaps simply
by being pronounced). As each new copy of an XP is generated (via
“re-merge,” i.e. Move), the relevant locality restrictions are checked (sub-
jectivity, etc.) and if locality is not respected (e.g. if an island node is crossed),
the new copy is given the feature * (compare the γ-marking of Lasnik and
Saito 1984; 1992 and Chomsky and Lasnik’s 1993 *-marking; see Kitahara
1999 for an alternative viewpoint). All later copies of this *XP will them-
selves also be *-marked. Finally, I assume that the *-feature can be erased
(checked) in the final spec-head relation that a +wh XP comes to be in with
a +wh C (perhaps * should in fact be thought of as feature of [wh], in some
conceptions); in other words, the [+wh, +Q] C that licenses the movement
of a wh-XP to its specifier also checks the *-feature on that copy, eliminating
the * from the PF-representation of the highest copy. (Perhaps even the E
feature itself checks the *-feature.)

Under this conception, a standard island effect will come about whenever
intermediate *-traces survive until PF. I assume, following Chomsky (1986a), Fox (2000), and López and Winkler (2003) (see also
Koster 1978 for a precursor), that wh-movement proceeds by adjunc-
tion to intervening maximal projections (VPs and IPs at the least; other
intermediate landing sites will not be crucial here). I illustrate the work-
ings of this system with the schematic derivation of (47) in (48), giving
only the steps of interest here:

(47) *What language do they want to hire someone who speaks?

(48) 1. Merge wh-DP [what language]:
   speaks [what language]

   2. Move DP out of relative clause, marking it with *:
      [*what language] [CP who speaks [what language]]

   3. Move DP to intermediate landing site, adjoined to matrix VP:
      [*what language] [VP want to hire someone [*what language]
      [CP who speaks [what language]]]

   4. Move DP to intermediate landing site, adjoined to matrix IP:
      [*what language] [IP they [*what language] [VP want to hire
      someone [*what language] [CP who speaks [what language]]]]

   5. Move DP to specCP, erasing * on final (namely, the highest)
copy of DP:
      [CP [what language] [IP [*what language] [IP they [*what
      language] [VP want to hire someone [*what language] who
      speaks [what language]]]]

In the final representation, all the intermediate traces between the CP
island node and the final landing site are *-marked. Since the *-feature
is by hypothesis PF-uninterpretable, the final structure will cause a
PF-crash, yielding the ungrammaticality of (47) as desired.

Conceiving of the *-feature as a feature of traces, however, allows us to
make the necessary distinction between IP and VP Ellipsis. In a Sluicing
example like (49a), deletion of the highest IP eliminates all *-traces from
the PF-representation, yielding the attested amelioration of island effects
under Sluicing. In VP Ellipsis, on the other hand, the elided VP does not
Table 6.1. Overview of the data with respect to MaxElide and *t

<table>
<thead>
<tr>
<th>Examples</th>
<th>MaxElide</th>
<th>*t</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Slicing</td>
<td></td>
<td></td>
<td>Island-containing</td>
</tr>
<tr>
<td>(8)-(16)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(26b), (27b), (28b)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>VP Ellipsis</td>
<td></td>
<td></td>
<td>Non-island-containing</td>
</tr>
<tr>
<td>*t(18)-(25)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>*(26c), (27c), (28c)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(29)-(31)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>*(39)-(46)</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

contain all the offending traces: in particular, the trace of a conjunction to the highest IP (*t''_2 in (49b)) remains in the PF-representation, causing a crash just as in (47).

(49)  a. They want to hire someone who speaks a Balkan language, but I don't remember which.

b. 

```
  ...CP
 [DP which]_2  
  C'       
    C       
    IP      
     IP     
     t_2    
     IP     
     t_2    
  they    
  VP      
    do     
    VP     
    want to hire
 [NP someone[who speaks t_2]]
```

The proposed system, consisting of MaxElide and of * as a feature of traces, captures a wide range of data, both in Slicing and VP Ellipsis, and both in structures containing islands and those without. All the relevant data presented here is tabulated by example number in Table 6.1. The top two rows include all the Slicing data examined. These examples satisfy MaxElide and no *-traces survive to PF (indicated by check marks in the columns labeled MaxElide and *t respectively). The VP Ellipsis examples fall into four groups, given in the final four rows of Table 6.1. The only grammatical examples are those in (29)-(31), which satisfy both MaxElide and do not present any *-traces. All other kinds of examples fail on one or both counts, as indicated by the * in the respective column. Interestingly, the kinds of examples which have been taken as most indicative of the inability of VP Ellipsis to repair islands, namely (18)-(25), are doubly bad: not only do their island-violations survive the deletion, but they also violate MaxElide.

Notice that this account locates the deviance of the island-containing VP Ellipsis examples in the clause that hosts the *wh*-movement. An intermediate *-trace of the successive cyclic *wh*-movement survives VP Ellipsis, but not IP Ellipsis. *wh*-movement in the elliptical clause in both cases is successive cyclic. In this regard, the account thus differs from Fox and Lasnik (2003), who account for the difference between Slicing and VP Ellipsis with respect to island repair by assuming that successive cyclic *wh*-movement occurs only in the VP Ellipsis case, but not in the Slicing case; they trace the different status of the examples to this posited difference, and its effects on the resulting structures for satisfying an Lf-isomorphism constraint (in the cases they examine, the correlates are all indefinites, which Fox and Lasnik assume are interpreted as choice functions and do not move).

Up to this point, the correlate to the moved *wh*-phrase has played no role. In Slicing, we have seen mostly examples where the correlate was an indefinite, though it is known that other kinds of correlates are possible. Both names and quantifiers can be correlates in Slicing (see Chung et al. 1995 and Romero 1998 for relevant discussion).

(50)  a. Abby speaks GREEK, but I don’t remember what OTHER languages.

b. She met RINGO, but I don’t know who else.

c. He said he talked to ABBY, but I don’t know who else he talked to.

d. John met most applicants, but I can’t remember exactly which ones.

The same holds for VP Ellipsis, as the examples in (31) above demonstrated, three of which are repeated here in (51a-c):

(51)  a. ABBY took GREEK, but I don’t know what language BEN did.

b. We know that Abby DOES speak [Greek, Albanian, and Serbian], we need to find out which languages she DOESN’T speak! (Merchant 2001:115 fn. 5 (ii))
c. (I know) ABBY wants to take GREEK, but I don’t know what language BEN does want to take.

d. ABBY interviewed two-thirds of the applicants, but I don’t remember exactly how many of them BEN did interview.

This state of affairs is expected under the present account, since the wh-movement in the elliptical clause violates no islands, and the semantic identity condition based on e-givenness is satisfied (since e-givenness is defined using F-closure, the focused correlates will be replaced by variables; see Merchant 2001: 35–37). The data can also be accommodated on Fox and Lasnik’s account, assuming that the focused correlates scope at LF. For them, the antecedent clause in e.g. (50a) and (51a) must have at least the following structure (assuming scoping to IP, and an intermediate landing site at VP): GREEK λx[IP Abby [vp x λx’ [speaks x’ ]]]. This structure will license, via their posited LF-identity requirement, deletion of the IP or VP as needed, assuming that the wh-movement in the elliptical clauses is likewise successive cyclic.

Now we are in a position to appreciate the puzzle that arises with the examples in (52). In these cases, the focus correlate occurs inside an island, and the resulting slice is ungrammatical. We have already seen that this holds for VP Ellipsis in some of the sentences in (39)–(46) above, two of which are repeated here in (53).

(52) a. *Abby wants to hire someone who speaks GREEK, but I don’t remember what OTHER languages she wants to hire someone who speaks.

b. *The radio played a song that RINGO wrote, but I don’t know who else.

(53) a. *Abby DOES want to hire someone who speaks GREEK, but I don’t remember what kind of language she DOESN’T.

b. *BEN will be mad if Abby talks to Mr. RYBERG, and guess who CHUCK will.

(Merchant 2001:115 (15))

In short, Sluicing with indefinite correlates repairs islands, but Sluicing with focused correlates does not. Given the account presented so far, this contrast is puzzling. We cannot assimilate the deviance of the slices in (52) to that of the VP Ellipsis cases in (53), since only in the latter does the *-trace remain after ellipsis has applied. Nor does the Fox and Lasnik (2003) account fare any better: their account, like the present one, is based on the idea that *XPs (intermediate VPs and IPs, for example) are PF-uninterpretable (which is why PF-deletion repairs the islands). In their presentation, “avoiding an intermediate landing site . . . yield[s] an island violation . . . [unless] the island is deleted” (p. 12); I take this to be a member of the *as-a-feature-of-island-nodes family of analyses. Consider now (52a): the correlate must scope out of the island in one fell swoop: GREEK λx[Abby wants to hire someone who speaks x]. This structure licenses the deletion of the IP beneath what OTHER languages, which has similarly undergone non-successive-cyclic wh-movement, by hypothesis. Under their account, therefore, we again expect the slices in (52) to be grammatical. These slices’ ungrammaticality cannot be assimilated to that of the VP Ellipsis examples in (53) on Fox and Lasnik’s account either (long-distance wh-movement in the latter is ruled out by virtue of the surviving *IP in the elliptical clause). Alternatively, Fox and Lasnik could assume that the focus movement must be successive-cyclic, while the movement in the Slicing cases is not; in this case, the Slicing examples would be ruled out (as violations of LF-parallelism), but the VP Ellipsis cases should then be grammatical.

One possibility that would rescue both the present account and Fox and Lasnik’s is that focus movement is island-sensitive for other reasons. If so, then the movement needed to generate the LF structures (on Fox and Lasnik’s account) or to provide the semantic antecedent (on the e-givenness account) would be illicit in any case. If this were so, the examples in (52) and (53) would be ruled out because ellipsis parallelism could not be satisfied (and the examples in (53) redundantly also because of the * in the elliptical clause).

The problem with this idea is the usual assumption that focus is not, in fact, island respecting (Chomsky 1972b; Rooth 1985; Kratzer 1991), whether this is accomplished by island-insensitive scopal movement or in situ. Kratzer (1991a) in particular provides examples involving VP Ellipsis which seem to indicate that a focused item can be ‘scoped out’ of an antecedent VP for purposes of ellipsis as well.

(54) I only talked to the woman who chaired the ZONING BOARD because you did.

As Kratzer points out, (54) has a reading as follows: the only x such that I talked to the woman that chaired x because you talked to the woman who chaired x is the zoning board. This reading necessitates island-violating scoping of the focused element zoning board in order to allow the bound reading in the elided VP. The example in (54) suffers from a slight defect, however: since the pitch accent falls on the final DP in the relative clause, it is difficult for some speakers to distinguish this from focus on the entire DP the woman who chaired the zoning board (cf. Drubig 1994 and Winkler 1996). Kratzer’s claim can be seen perhaps more clearly in the following
example, where no focus percolation from within the relative clause to the containing DP is possible:

(55) I only played a song that RINGO wrote because you did.

This example, parallel to (54), has a reading that can be paraphrased as follows: the only x such that I played a song that x wrote because you played a song that x wrote is RINGO.

These latter facts are unsurprising under the view of islands that emerges from the study of Sluicing especially: islands are essentially PF phenomena, so the movement necessary for the focus in (54) and (55), since it has no PF consequences, is not expected to result in island violations (see Rooth, 1996). Likewise for wh-in-situ inside islands, if these require long-distance movement (see Simpson 2000 for a recent overview).

We seem to have reached an impasse: the evidence from 'contrast' sluices in (52) indicates that focus movement is island-sensitive, while the data in (54) and (55) seems to indicate that it is not. But this is, luckily, not the only difference between (52) and (54)–(55): in the former, but not the latter, we also have an instance of wh-movement out of the ellipsis site. This wh-movement has scopaal properties of its own, and I would suggest we can capitalize on these properties to rule out (52) while allowing (54)–(55). Wh-movement out of an ellipsis site forces its correlate to take scope over the entire antecedent clause, just as the wh-phrase itself does in its own clause (see Chung et al. 1995 and Romero 1998 for extensive analysis of this fact; see Fox 2000 for general discussion of scopal parallelism in ellipsis). All the data examined so far indicate that when focus movement extracts a focused XP out of an island, the focus movement can no longer target the highest clause node, but is limited to the VP. Metaphorically speaking, it is as though escaping from an island cripples or hobbles further focus movement; it can only limp along up to VP, not to IP. I thus suggest that the differences seen above are trace to these differing possibilities for satisfying the identity condition on ellipsis.

In the non-island cases, focus movement of the correlate in the antecedent can scope the focused XP to IP, as we saw above (yielding an LF-parallel structure or a structure which satisfies e-givenness, assuming that the existential binder is inserted at the locus of the moved focused XP). For these cases, the antecedent after focus movement will be the following (for (50a); mutatis mutandis for (51a))

(56) GREEK_P x [IP Abby [x λx’ [speaks x’]]]

This structure has the correlate in a parallel position (clause-external) to the moved wh-phrase in the elliptical clause, satisfying parallelism:

(57) what OTHER languages λx [IP Abby [VP x λx’ [speaks x’]]]

Consider now the example in (55). Here, focus movement must scope RINGO to the highest VP (above the VP-adjoined adjunct because clause), satisfying the parallelism requirement for the deletion of the VP in the because clause:

(58) I only *RINGO_P x [VP [VP played a song that x wrote] because you did] play a song that x wrote]

The movement in (58), though island-violating, targets the matrix VP, not IP. Finally, in the puzzling cases of (52) and (53), the stipulation that island-escaping focus movement cannot target the highest IP will prevent the correlate from attaining the necessary scopal parallelism with the wh-phrase (clause-external), and hence these clauses can never satisfy the identity requirement needed to license deletion. The highest the focus movement can go is VP, yielding the following as the antecedent clause, by hypothesis (intermediate traces suppressed):

(59) [IP Abby [VP GREEK_P x [VP wants to hire someone who speaks x]]]

A structure like (59), unlike (56), does not have a scopolal element in a position parallel to the moved wh-phrase, and hence will license neither deletion of IP nor of VP.

These considerations will not apply in cases where the correlate is an indefinite, since indefinites are known to be able to freely take wide scope, even out of islands (see Farkas 1981). Hence such indefinites will always be able to provide scopally parallel antecedents for wh-phrases moved out of ellipsis sites, as Chung et al. (1995) discuss (whether or not these indefinites are interpreted as choice functions or not). So scopal parallelism is satisfied in the Sluicing cases we began with in (8)–(16) as well as in the VP Ellipsis cases in (8)–(25) and those with indefinites in (39)–(46). The latter, therefore, still necessitate the theory of islands developed above, since the focus-based restriction that rules out (53) will not apply to those cases with indefinite correlates.

A related point comes from correlatives that are interpreted as generalized quantifiers, forcing them to scope via QR, known to be even more local than merely island-respecting (though not quite, as often assumed, clause-bound: see Farkas and Giannakidou 1995). These provide possible antecedents for Sluicing and VP Ellipsis, as in (50d) and (51d) above (repeated here as (60a,b) respectively).

(60) a. Abby met most applicants, but I can’t remember exactly which ones.
b. ABBY interviewed exactly two thirds of the applicants, but I
don’t remember how many of them BEN did.

The grammaticality of these examples is expected, since movement of
the correlates via QR will provide appropriate antecedents for the elliptical
clauses. This holds both for an e-givenness-based account and an LF-
isomorphism account of Fox and Lasnik (2003). The fact that such corre-
lates are island-bound is also expected, given the local nature of QR:

(61)  a. *If most senators resign, Abby will stop her hunger strike, but I
can’t remember exactly which ones.

b. *If exactly two thirds of the senators resign, ABBY will stop
protesting, but I can’t remember how many of them BEN will.

But notice again that something more must be said to rule out examples
like (62), which differs from (60b) only in not having a contrasting element
in the clause hosting the VP Ellipsis.

(62) ?? Abby met most applicants, but I can’t remember exactly which
ones she did.

Recall that Fox and Lasnik rule out examples similar to this one in which
however the correlate is an indefinite – such as (28c) above – by positing
that the indefinite, unlike the wh-phrase, does not move, and hence the
elliptical clause does not satisfy LF-parallelism (since it, unlike the ante-
cedent clause, contains traces of successive cyclic movement, by hypo-
thesis). But as we have just seen, when the correlate is a quantified or focused
DP, not an indefinite, successive cyclic movement must be posited, in order
to rule in (60b) and (51a–c). The current account, employing MaxElide,
rules out (62) and (28c) on a par, while one based on the postulated
presence vs. absence of successive cyclic movement apparently must be
supplemented by MaxElide in any case.

6.4 Conclusions

This chapter has investigated a number of surprising asymmetries in island
repair between Sluicing (IP Ellipsis) and VP Ellipsis, and has argued that these
fall out from taking certain island effects to be due to ill-formed intermediate
traces at the PF interface. The conclusion, then, is that a number of deviances
that have been ascribed to other parts of the grammar (derivational con-
straints, LF output constraints) may best be located at the PF interface (at
least part of the effects of island-violating extraction). Prima facie paradoxically,
it seems that the nature of PF and the constraints that operate there can
be illuminated by investigating structures that have no PF exponence.

Certain elements of the analysis presented here remain at present theo-
retically unsatisfying, in that the effects encoded in the various constraints
(especially MaxElide and the restriction on island-escaping focus move-
ment) have yet to be reduced in an insightful way to the theoretical
primitives they presumably derive from. This project is ongoing, and I
mention here a number of other areas in which repair effects of ellipsis
seem to be indicated:

1. lack of complementizer agreement in Bavarian Sluicing

2. lack of Wackernagel clitics in S. Slavic Sluicing

3. multiple Sluicing in Germanic, Greek, and Turkish (and perhaps in
   Bulgarian, Japanese, Russian, and Serbo-Croatian as well)

4. remnant movements in Gapping (Johnson 2003, Richards 1998)

5. remnant movements in Pseudogapping (Johnson 2001)

6. lack of verb movement in Pseudogapping (Lasnik 1995, 2001)

7. swiping in English, Norwegian, Danish (Merchant 2002)

8. “vehicle change” effects in anti-pronominal contexts (Potts 1999)

9. long-distance reflexives in English (Kennedy and Lidz 2001)

10. wh-movement in wh-in-situ languages

11. lack of I-to-C movement in matrix sluices in Germanic (Lasnik 1999b
    and Merchant 2001)

12. lack of the otherwise obligatory complementizer in Irish sluices
    (Merchant 2001).

This potpourri of effects has emerged mostly recently from investigating
elliptical structures from the perspective of repair, and it is not surprising
that they have yet to be made to follow from primitives of any theory. It
appears that we have just begun to uncover a new domain that may help us
shed light on phenomena that have been traditionally investigated only
with respect to their pronounced manifestations.
15. Alternatively, we could define two types of antecedents: one with respect to linking (coreference) configurations; one with respect to binding configurations, and retain Higginbotham’s definition of dependence. The end result would be the same, but would not reflect as clearly as refining the notion of dependence the fact that we are really making a distinction between binding and coreference.

16. By allowing outer indices to appear on every maximal projection in the extended projection of N, i.e., on both DP and NP (see Grimshaw 1991), we can assume that both DP and NP are potential binders, modulo the e-command constraint on binding.

17. Note that a structure in which the higher NP is the binder of an element inside VP is ruled out for the same reason that DP cannot be a binder – the bound element in VP would be e-dependent on itself.

18. The basic idea behind Higginbotham’s analysis is that in the translation of a structure like (62) into a logical representation, the relative clause is mapped into the restriction of the determiner every, while the VP is mapped into its nuclear scope (see Heim 1982; Diesing 1992b). Empty categories bound by the DP in the matrix sentence and those bound by the relative operator in the relative clause are interpreted as co-bound, i.e. bound by the determiner every. A more intuitive, semantic representation of (62), where the variable x has been substituted for the index 2, is given in (vi).

(vi) \forall y [\text{man}(x) \land \text{clean}(x, y, \text{Sari})] \land [\text{happy}(y)]

19. For clarity, I will only show indexing on the relevant DPs (the ones involved in the argument containment configuration). Although the subjects of the matrix VPs will necessarily bear the same outer indices (because the matrix VP is what is copied), this index does not play a role in the e-dependence relation involved in the ungrammatical sentences.

20. The analysis proposed in this chapter has no account of the ungrammaticality of (76) in terms of ill-formed indexing configurations (and would, in fact, predict it to be grammatical). Given the grammaticality of examples like the ones in (77), I will assume that an explanation of (76) will be found outside of conditions on indexing configurations, and will leave a resolution of this issue for future work.

CHAPTER 6: VARIABLE ISLAND REPAIR UNDER ELLIPSIS

1. The Sluicing correlates that will interest us for the moment are all indefinites; see Chung et al. (1995) for discussion. Other correlates are possible (the “contrast” sluices of Merchant 2001, for example), but these show strong locality effects (stronger, in fact, than more sensitivity to islands: closer to the kinds of locality found in Gapping and multiple Sluicing). The picture is further complicated by the fact that certain Sluicing-like structures in some languages appear to retain island-sensitivity, as reported in Hoji and Fukaya (1999).


3. This example shows that such contrast-licensed wh-extraction from VP Ellipsis can cross tensed clause boundaries; nevertheless, such examples are highly sensitive to the nature of the intervening material: if the embedded subject is not bound by the matrix subject, the matrix reading for the elided VP becomes extremely difficult to get (cf. identical bound subject restrictions in cross-clausal Gapping and multiple Sluicing as discussed in Nishiguchi 1998; Johnson 2001; and Merchant 2001).

CHAPTER 7: ON BINDING SCOPE AND ELLIPSIS SCOPE


2. The two dichotomies, wide vs. narrow scope, and de dicto vs. de re, are well known to match only partially (wide scope implies de re, but narrow scope is compatible with both de dicto and de re). This inaccuracy does not affect the exposition in the text, though, as the diagnostics to be used do not test for referential transparency.

3. See, e.g., Hankamer (1973b) and Pinkham (1982). In what follows, than will – for expository convenience – be bracketed as a coordinator in a ternary structure. For discussion of coordinate properties of comparatives and the structure of comparative coordination see Lechner (2004) and references therein.

4. Wide scope in (8) cannot be attributed to a non-scoping mechanism (choice functions or non-local binding of world variables for de re readings; see section 7.2.2), as these devices do not evacuate the relative clause from the e-command domain of the antecedent. Moreover, on the standard assumption that VPE may only elide a single node, and not e.g. multiple terminals, reading (8b) leads to regress, as the antecedent still contains the ellipsis site.

5. Similar ambiguities have been discussed in McCawley (1998: 688) and Pinkham (1982: 130). McCawley assumes that the two readings in examples similar to (11) are not derivationally related, but arise from different deep structures. Pinkham considers (i), and concludes that reading (ii) is the product of a construction specific deletion process, while (ii) represents a base-generated PC which does not include any elliptical structure. As will be seen shortly, both views differ from the one advanced here, according to which all PCs derive from a clausal source (potentially a small clause, as in (ib)).

(i) John seems taller than Bill.  
   a. d = seems d-tall  
   b. d = is d-tall

6. In the narrow reading (= (11a)), the verb is also underspecified for tense. Note incidentally that atemporal readings are subject to the additional, curious restriction that the correlate has to be structurally higher than the comparative NPs (see Lechner 2004 for a partial analysis):

(ii) John will subject more students to this year’s exam than d to last year’s exam.  
   d = John will subject/*subjected d-many students