Type-Logical Syntax

The title-level DOI for this work is:

doi:10.7551/mitpress/11866.001.0001

10 On Functional Constraints on Extraction: The Status of Island Constraints

In examining the range of arguments and counterarguments for various approaches to the analysis of the grammatical patterns surveyed in previous chapters, readers will likely have been struck by the frequency with which appeal is made in the literature to supposed facts about islandhood. This is unsurprising: most analyses of coordination and ellipsis in mainstream generative grammar posit covert structure (though the details often differ considerably). Islandhood has long been regarded as the gold standard for arguments on behalf of specific configurations, based as it is on specific locality conditions that filler-gap linkages must satisfy—conditions which can be invoked as evidence for specific configurations. The critical assumption here, of course, is that islandhood is actually a reflection of purely *syntactic* conditions on locality—an assumption which has been predominant from an early stage of generative grammar research, and one shared very widely not only across phrase structural frameworks until relatively recently but also in parts of the CG research community. Thus, we find in both Combinatory Categorial Grammar (cf., e.g., Steedman 2000, 2012) and the Displacement Calculus, a version of TLCG (cf., e.g., Morrill 2010, 2017), the premise that the CG combinatory formalism must be expanded to capture island effects.

Nonetheless, the premise that island effects are entirely (or even primarily) structural in their origin has, during the past twenty-five years, been increasingly abandoned by many theorists. The significance of this general change in research practice for the approach advocated in this volume can hardly be overstated: it can now no longer be taken for granted that, for example, the null hypothesis for the appearance of island-like effects in patterns of ellipsis is the presence of covert syntactic configuration in the clause displaying such ellipsis. In the first part of this chapter, we provide an overview of the history of syntactic accounts of island effects and summarize the current controversy regarding the status of these effects in the grammar. We conclude that the alternative accounts that take island effects to be non-syntactic are more promising than the traditional syntactic view of island effects. Based on this discussion, in the second part of this chapter, we assess the validity of specific arguments made in the

literature for analyses of (subtypes of) coordination and ellipsis that crucially invoke islandhood data. This critical review reveals that even if one accepts the premise that islandhood is indicative of specific phrase structural configurations, in almost all cases, such arguments are empirically defective: there is extensive counterevidence for the data that are supposed to support a specific configurational analysis. This leads us to the conclusion that the true explanations for the empirical patterns (or tendencies) first observed by the proponents of structural analyses are more likely to lie elsewhere.

10.1 The History of Syntactic Island Effects

10.1.1 Ross's Constraints: CNPC, CSC, Sentential Subject, and Left Branch Constraints

Early on in the modern history of grammatical theory, Chomsky (1964) noted data suggesting that *wh* and related extraction phenomena cannot occur in certain contexts, of the kind illustrated in (486), and proposed to account for it by what he called the A-over-A principle, given in (487).

(486) a. John discussed [your concerns relating to Mary's rash statements].
b. *Which statements did John discuss [your concerns relating to __]?
c. [Whose concerns relating to Mary's rash statements] did John discuss __?

(487) [I]f the phrase X of category A is embedded within a larger phrase ZXW which is also of category A, then no rule applying to the category A applies to X (but only to ZXW).

This formulation is, of course, far too general; data such as (488) show that (487) is empirically untenable:

(488) a. John came across [a remarkable new book about a nineteenth-century author].
b. That's the author that John came across [a remarkable new book about __].

Examples such as (488b) are hardly obscure or dodgy. Nonetheless, another three years elapsed before the inadequacy of Chomsky's account was clearly demonstrated in the literature and an alternative set of proposals offered to account for a much larger range of data than had been previously considered. Ross (1967) provided the first detailed consideration of supposedly syntactic conditions on extraction phenomena, listing a number of specific phrase structure configurations prohibiting any linkage between displaced material and the gap site originally occupied by that material. These configurations became widely known, following Ross's phrasing, as *syntactic islands*.

Probably the most influential of Ross's proposed constraints was the *Complex NP Constraint* (CNPC), intended to explain the distribution of data displayed in (489)–(492).

(489) a. Mary bought a book which John wrote for MIT Press.
b. *Which publisher did Mary buy a book which John wrote for __?

(490) a. I believe Mary's promise to debug the program.
b. *Which program did you believe Mary's promise to debug __?
c. *I'll show you the theorem prover that I believed Mary's promise to debug __.

(491) a. I became aware of the rumor that John had worked for Mary.
b. *Who did you become aware of the rumor that John had worked for __?
c. *The critic who I became aware of the rumor that Mary had challenged __ to a duel entered the room.

(492) a. John is too unreliable to keep his promise to debug that theorem prover.
b. *Which theorem prover is John too unreliable to keep his promise to debug __?

On the basis of facts such as (489)–(492), Ross formulated the restriction summarized in (493), which he called the Complex NP Constraint.

(493) **Complex NP Constraint:** Nothing can be extracted out of an NP where the gap site is under S in the configuration $[_{NP} \ldots S \ldots]$.

The CNPC was frequently and enthusiastically cited in the syntactic literature; for several decades after Ross's thesis appeared, the only remaining issue seemed to be whether or not his formulation was sufficiently general. For much of the history of modern syntax, the "default" position has in fact been that it is not and that the facts motivating the CNPC can be shown to fall out from a more general condition called Subjacency, which encompasses a number of phenomena that Ross treated separately from the CNPC or did not address at all.

In contrast, the condition that Ross labeled the *Coordinate Structure Constraint* (CSC) has for the most part been regarded as a stand-alone phenomenon from the outset. It comprises two quite separate conditions:

(494) **Coordinate Structure Constraint**
a. *Conjunct Constraint:* Conjuncts may not be extracted.
b. *Element Constraint:* No element may be extracted from a proper subset of the conjuncts in a coordination.

The Conjunct Constraint was intended to account for facts like those in (495).

(495) a. I like raspberries and papayas in fruit salad.
b. *What do you like raspberries and __ in fruit salad?
c. *What do you like __ and papayas in fruit salad?
d. *What do you like __ and __ in fruit salad?

In these cases, whole constituents showing up in the pre- and/or post-conjunct particle position have been extracted, leading to markedly bad results.

What distinguishes the Conjunct Constraint from the Element Constraint is that in (495), extraction of all conjuncts does not help in the least, so far as the status of the result is concerned. Compare this pattern with the pattern illustrating the Element Constraint in (496):

(496) a. I play cards with the gang and go to night classes on alternate Thursdays.
b. ??/*Who do you play cards with __ and go to night classes on alternate Thursdays?
c. Who do you play cards with __ and attend lectures by __ on alternate Thursdays.
d. Florence is the city that I visited __ last year and would like to return to __ in the near future.
e. To whom have you presented your plans __ or shown your prospectus __?

Just as with the CNPC, we find the CSC in all extraction constructions:

(497) Visit Paris, John definitely will __ this summer and I really should __ sometime next year.

(498) a. It is easy (for us) to please John and offend Mary.
b. John and Mary are easy (for us) to please __ and to offend __, respectively.
c. *John is easy (for us) to please __ and offend Mary.
d. *John is easy (for us) to please Mary and offend __.

The restriction imposed in (494b) is often referred to as the *Across The Board* (ATB) condition: a gap in a coordinate structure anywhere entails that an extraction dependency holds in *every* conjunct in the coordination, regardless of the category of the mother. In the examples of the CSC given above, the conjuncts were all VPs. But note cases such as the following:

(499) a. Who does John like (*Ann) but Mary always criticize __? (S *and* S)
b. What is John worried about __ and troubled by (*the tax audit)? (AP *and* AP)
c. Which painter did John buy a book about (*Victor Hugo) and a biography of __? (NP *and* NP)
d. Who did you carry messages to (*John) and from __? (PP *and* PP)

We see in all these examples the effect that Ross aimed to capture via the ATB requirement: extraction is permitted in all of these cases as long as each conjunct hosts a gap linked to the same filler.

The third constraint Ross lists in the core chapter 4 of his dissertation is the *Sentential Subject Constraint* (SSC), which can be simply stated as follows:

(500) **Sentential Subject Constraint:** No constituent may be moved out of a clausal subject.

This condition can be illustrated via the examples in (501):

(501) a. *I know which people [$_S$ that John invited __] bother you.
b. *Who do you suspect that [$_S$ inviting __] was a bad idea?
c. *Who do you think that [$_S$ (for us) to invite __] would be a mistake?

Ross did not in general regard extraction from a subject as ill-formed in itself but rated all examples in which a gap appears in a subject with clausal structure as uniformly bad.

Finally, Ross posited the *Left Branch Constraint* (LBC), which had the specific intent of blocking the extraction of an NP appearing on the left branch of its mother NP—essentially, a prohibition on extracting possessor NPs as in (502).

(502) a. *Whose did you buy [$_{NP}$ __ book]?
b. *John's I bought [$_{NP}$ __ book].
c. *John's was too expensive to buy [$_{NP}$ __ book].

10.1.2 Island Constraints Following Ross (1967)

One striking aspect of how syntactic theory developed over the decades following Ross's watershed dissertation is the way in which the generalization of one of the four constraints he proposed, namely the CNPC, came to dominate the search for increasingly abstract and supposedly general principles of syntactic well-formedness. The main goal of this line of research was to link the conditions on extraction to supposedly separate grammatical domains, such as anaphora and coreference, with the SSC eventually assimilated to a subcase of the Subjacency Condition that Chomsky (1973) proposed to account for the breakdown of filler-gap connectivity in the configurations specified in the CNPC. There have also been efforts to derive the LBC as a consequence of various principles posited in later developments of transformational grammar. The CSC, in contrast, has for the most part been treated as a stand-alone restriction unconnected to more fundamental principles.

10.1.2.1 Principles and Parameters: Subjacency and weak vs. strong islands

Chomsky (1973), the next major development in the theory of constraints opened up as a research area by Ross (1967), recast the CNPC in more general terms as a locality condition on movement defined with respect to a certain class of category types. In earlier work, transformations had been taken to apply in cyclic fashion in both clausal and nominal domains, whose respective maximal projections, $\overline{S}$ and NP, were therefore identified as

"cyclic nodes," and it was just these nodes—an apparent natural class—which defined the locality domains for movement. Chomsky's proposed reanalysis of the CNPC is the condition in (503), which he called *Subjacency*.

(503) **Subjacency:** In the structure $[_{\alpha} \ldots [_{X} \ldots C. \ldots] \ldots]$, α cyclic, no extraction rule moving a category C out of X can apply unless C is subjacent to α, that is, embedded no more than a single cyclic node below α.

(503) has the status of an analytic primitive in connection with a revision of the phrase structure of clauses in which the Comp node (which hosts words such as *that*, *for*, and *if*) is a sister to the uncomplementized clause S. Certain other conditions that Chomsky proposed in this paper imposed conditions which suggested that this Comp node could act as a gateway through the barriers to movement imposed by those conditions but was only available to *wh* phrases; these, in turn, once in Comp, could *only* move to a higher Comp node. This restriction, according to Chomsky, accounted for the fact that (504) is ill-formed, since the lack of a Comp node associated with the NP means that *who* must move from its position in the lowest Comp through the immediately dominating $\overline{\text{S}}$ and then directly across NP (both of which are taken to be cyclic nodes under the theory of the transformational cycle generally assumed at that time), violating Subjacency:

(504) *Who did he believe the claim (that) John saw __?

As Chomsky noted with specific reference to (503), "in this way we can explain many of the examples that fall under the complex noun phrase constraint" (Chomsky 1973, 104).

But in fact, on the assumptions in Chomsky (1973), the Subjacency condition appears to mispredict dramatically, since Ross's key example (505), which he had invoked in his argument against Chomsky's earlier A-over-A constraint, is likewise incorrectly ruled out under (503):

(505) What books does the government prescribe $[_{\text{NP}}$ the height of $[_{\text{NP}}$ the lettering on $[_{\text{NP}}$ the covers of __]]]?

The reasoning here is the same: *what books* must pass through three NP nodes in order to reach Comp, in clear violation of (503). While Chomsky argues that (506) supports the Subjacency condition, he acknowledges that there appears to be a conflict with cases such as (505) but says nothing about the empirical challenge to Subjacency posed by (505) beyond the observation that "I see no obvious explanation for an apparent difference in degree of acceptability."[1]

1. As made clear in later work by Robert Kluender (1998), discussed below, there is indeed an obvious difference between the two examples, viz., the level of referential specificity carried by the extracted *wh* phrase. Suppose we replace (506) with (i):

(506) *??What do you receive [$_{NP}$ requests for [$_{NP}$ articles about __]]?

Yet despite the fact that the empirical basis for the Subjacency condition provided in Chomsky (1973) is not just sparse but empirically dubious, it was widely accepted among theoreticians as preferable to Ross's CNPC, largely, it seems, on the basis of its greater generality and abstractness. Further wrinkles involved changes in the assumptions about which English categories embodied that same property. In particular, it was recognized that if one took S, rather than $\overline{S}$, to be the clausal bounding category, then the following configuration fell negatively under the scope of Subjacency,

(507) [$_{\overline{S}}$ *wh* [$_S$ . . . [$_{\overline{S}}$ X [$_S$ . . . __]]]]

where X's occupancy of the position under Comp has the effect of blocking the *wh* phrase's movement to that position en route to the matrix Comp. This is in fact the description of a class of examples exemplified in (508):

(508) a. *What did you wonder who John gave __ __?
b. *Where did you figure out when John put the receipts __ __?

In these and similar cases, a *wh* phrase occupies a Comp node through which a subsequently moved *wh* phrase would need to pass en route to its final position in the structure.[2] The data pattern in (508) was taken to be an instance of something called a "*wh* island," which, by replacing $\overline{S}$ with S among the set of bounding nodes in English, was automatically reduced to just another consequence of Subjacency.

Further research on what Chomsky began to call "bounding theory" came from the work of Rizzi (1980) and others, who presented evidence that could be understood to support the variability of the nodes which count in any given language as having the bounding property with respect to Subjacency. The limited freedom of languages to determine which nodes counted as bounding nodes was historically perhaps the major influence leading to the concept of parameterization: universal grammar as a set of

---

(i) Which natural disasters do you receive requests for articles about __?

Structurally, there is no difference between (i) and (506) in relevant respects, and the lexica in the extraction domain are identical in the two cases, but (i), in marked contrast to (506), is close to impeccable. The properties of the respective extractees themselves constitute the sole difference. This shows that whatever the source of the ill-formedness in (505) and (506) might be, it cannot be an attribution of "bounding" status to NP nodes. Also, while Chomsky (1971, 1980, 1981, 1982, 1986b) repeatedly refers to the A-over-A constraint as though it were not just a viable but generally accepted hypothesis, it is plainly falsified by the data set in Ross's thesis, which includes (505). See Levine and Postal (2004) for discussion of some of the implications involved.

2. There is an alternative derivation in which the higher *wh* phrase moves first, passing through the lower Comp before moving to the higher one, but this would entail the *wh* in the lower position to replace the trace left by the first *wh* movement, which was generally assumed to be ruled out as a possibility in transformational grammar (though it is difficult to find a principled argument in the literature undergirding this assumption).

underspecified principles containing parameters with certain possible settings. The child then sets the value of each parameter on the basis of the positive data s/he is exposed to.

The picture of "bounding theory" that had emerged by the end of the 1970s proved relatively stable until the appearance of *Barriers* (Chomsky 1986a), in which it became considerably more complicated. In particular, the notion "barrier," which corresponded to the bounding nodes of earlier work, was no longer completely fixed in advance for the movement of any constituent, but rather depended for its status on the extraction path followed by that constituent prior to encountering the potential barrier. *Barriers* provides an intricate and in places somewhat opaque set of definitions, whose crucial content is given in the following passage (Chomsky 1986a, 14–15):

> We first define *blocking category* (BC) as in (25) and then define *barrier* in terms of BC as in (26):
>
> (25) γ is a BC for β iff γ is not L-marked and γ dominates β.
> (26) γ is a barrier for β iff (a) or (b):
> a. γ immediately dominates δ, δ a BC for β;
> b. γ is a BC for β, $\gamma \neq$ IP
>
> We understand γ in (25) and (26) to be a maximal projection, and we understand *immediately dominate* in (26a) to be a relation between maximal projections (so that γ immediately dominates δ in this sense even if a nonmaximal projection intervenes). In case (26a) the category γ inherits barrierhood from a BC that it dominates; in case (26b) γ is a barrier intrinsically, by virtue of its own status as a BC.

L-marking refers to the relationship of lexical selection: X L-marks Y if X is a lexical head and Y is a constituent which saturates one of X's valence requirements; IP corresponds to the early GB assumption that an Infl node was the head of the clause, so that what was written "S" in previous work was actually to be analyzed as IP, the maximal projection of Infl. The point of (26b) is that, while other blocking categories automatically count as barriers for movement, the IP node only becomes a barrier by "inheriting" barrierhood from a daughter (blocking) category.

The best way to see how Chomsky applies (25) and (26) to the characterization of islandhood is to take two contrasting examples: extraction from a subject NP and extraction from an object NP. So far as the first is concerned, since NP is not L-marked, it is a blocking category, and since all blocking categories except IP are automatically barriers, NP is also a barrier. Thus, by Chomsky's (26a), IP is also a barrier for anything originating within the subject NP. Hence any constituent originating within the subject which moves to the nearest privileged landing site for *wh* movement crosses two barriers and is therefore, under Chomsky's finer-grained characterization of Subjacency, in "2-subjacent" violation of bounding theory. An object NP, on the other hand, is L-marked by the head of its VP, so that IP does not inherit barrierhood from anything,

and thus no barriers are crossed by the extraction—at least at first glance. Similarly, in the case of extraction from a relative clause attached to a nominal head, a *wh* phrase internal to that clause can move unproblematically to its Comp node; but the CP above it is both a blocking category (since it is not lexically selected) and an inherent barrier, and the NP above the CP inherits barrierhood from the latter. Hence, movement out of the relative clause crosses two barriers and therefore incurs a major bounding theory violation.

But the above scenarios omit a critical detail: the same line of reasoning appears to make *any* extraction from IP illicit for the same reason as extraction from a subject: VP is, after all, no more L-marked than its NP subject is, and hence by (25) is a blocking category for anything it dominates; hence by (26b) is a barrier for anything it dominates. But there is a special escape hatch that allows such extraction after all: the movement

(509) $[_{\mathrm{VP}} \ldots \alpha \ldots] \rightarrow [_{\mathrm{VP}}\ \alpha\ [_{\mathrm{VP}} \ldots t_{\alpha} \ldots]]$

is permitted via adjunction, and since, according to a novel interpretation of domination introduced in *Barriers*, not all the "segments" of the VP containing α actually dominate it, the VP as a whole does not dominate it. Thus the only maximal projection dominating α at this point is the IP itself, which is a blocking category, but now not a barrier, and so movement of α out of IP is now permitted. Why then cannot a constituent originating under the subject NP node, or within a relative clause attached to a nominal head, adjoin to NP as its first move and so escape IP and NP in the same way? In effect, the answer is that it can't because it can't. Adjunction to VP is possible; adjunction to NP is not.

It is evident in retrospect that the whole development of the theory of islandhood in the two decades between Ross (1967) and *Barriers*, originally predicated on the need to derive Ross's descriptive characterization of island environments from more general and abstract fundamental principles, had instead culminated in an extensive set of stipulative conditions and *ad hoc* exceptions. The use of adjunction to eliminate dominance relations where such relations would otherwise block legal movements and similar add-on mechanisms are certainly far more abstract than Ross's conditions. But they do not emerge from a simple set of fundamental principles any more than his do. The theory of islandhood in *Barriers* actually becomes considerably more ramified and intricate than we can detail in the space available, but the upshot can be summarized as follows: the "murky" questions that Chomsky wished to address in *Barriers* did not receive answers in which foundational simplicity and naturalness were much in evidence. One can in fact plausibly read the ready acceptance of the Minimalist framework within a decade after the appearance of *Barriers* as a sign that few researchers within the theoretical syntax community felt much would be lost by, in effect, scrapping the elaborate superstructure of the *Barriers* framework and starting again more or less from scratch.

10.1.2.2 The Empty Category Principle (ECP) Immediately after Ross's dissertation appeared, work by David Perlmutter (1968, 214–215) pointed out yet another perplexing restriction on English extraction. Perlmutter's examples included the data in (510):

(510) a. What did he say that Laura hid __?
b. *Who did he say that __ hid the rutabaga?
c. Who did he say __ hid the rutabaga?

As (510a) shows, there is no general problem with extraction from a clause with a *that* complementizer. And there is no problem with extraction of subjects, as (510c) itself makes quite clear. The problem arises just when the filler is linked to a subject gap position following a complementizer. This phenomenon became widely known (and is still typically referred to) as the *that-trace effect* (or *that*-t effect), although it also arises when subjects following *if* and *whether* are displaced (**Who did you wonder if/whether __ would get the job?*). For a decade or more after it was discovered, the *that*-t effect was taken in much influential work to be yet another primitive condition, a kind of "surface filter" on extraction which just had to be assumed as part of the grammar of English, with no obvious deeper source.

A line of thinking which began in the early 1980s seemed to suggest, however, that such a source might well exist. The details are complex, but the general idea was that a trace cannot just appear anywhere. It has to occur in a context in which it has a particular relationship to a selecting head or, as a secondary possibility, in a configuration which *resembles* head-selection contexts in the right way. For example, in a structure such as (511), which will be part of the representation of the sentence *I wonder which book you reviewed*, the head V is in the right position, as the left sister, to be a lexical selector of an NP complement.

(511) $[_S\ [_{NP}$ which book$]\ [_S$ you $[_V$ reviewed$]\ [_{NP}\ t]]]$

Whether or not this selection actually takes place, the "left sister" configuration was taken to correspond to special licensing properties that made the appearance of a following trace legal. But literal sisterhood wasn't necessary. For various reasons, it was assumed at this time that the following configurations defining structural relations between a lexical head X and a selected complement YP as in (512a) were at some abstract level equivalent to the more distant relationship in (512b), characterizable as instances of a single notion of *government* under the right circumstances (with bolded text indicating the constituents in the government relation), depending on what the intervening node Z is.

(512) a. $[_{XP} \ldots \mathbf{X\ YP}]$
b. $[_{XP} \ldots \mathbf{X}\ [_Z\ \mathbf{YP} \ldots]]$

For example, in the case of (513a), the complement of *expect* is assumed to be an infinitive clause with an accusative subject, as in (513b), under standard transformational assumptions.

(513) a. I expected her to get the job
b. expected [$_S$ *her* [$_{VP}$ *to get the job*]]

The subject of such clauses is in a parallel relationship to YP in (512a), since in, for example, *I expected her at 9:00 a.m.*, accusative case also appears. It was understood that accusative case assignment was determined by a lexical head in a government configuration with an NP, with both cases in (512) as subspecies of that configuration.[3] The configuration in (512b) was further assumed to subsume the relationship between a moved constituent and lower traces in Spec, although in this latter case X does not have the status of a selecting head.

This last point bears in a crucial way on analyses seeking accounts of the *that*-t effect at a deeper level than the simple surface filter account assumed through the late 1970s. Such analyses started from the premise that one or the other of the configurations in (512) had to hold if YP were a trace. For example, in (514), we have the trace in a properly governed position:

(514) Who$_i$ did you expect t_i to get the job?

There is an evident problem with this approach, however: it is clear that the subject position of *finite* clauses does not satisfy the restrictions on Z in (512b). We do not, for example, get *I expected him would get the job*. Yet extraction from finite subject position, as we have seen in (510), is unproblematic as long as the complementizer *that* is not present. Hence, examples such as (514) are actually misleading; if being properly governed is necessary for traces to appear, simply being in the position of YP in (512b) must not in itself guarantee proper government.

Transformational theorists working in this early-to-mid-1980s framework, therefore, made a further assumption. In the kind of analysis of filler-gap linkage assumed, as we have seen, from Chomsky (1973) onwards a constituent in a certain position moves in a series of steps, always upward and to the left, appearing in some "protected" position on the periphery of the clause and then moving up/leftward again, leaving a trace behind. The result is a chain of movements from one protected position to a higher protected position, and the typical structure of filler-trace or trace-trace linkages along this chain is displayed in (515):

3. A critical aspect of this analysis was the assumption that while subjects of embedded finite clauses are "shielded" from government by an external head via an intervening clausal boundary node ($\overline{C}$), infinitival complement clauses either lose this node by deletion or never possess it in the first place. The subject of infinitival clausal complements to heads such as *expect* are therefore subject to government by those heads.

(515) $[_S\ XP_{i_3}\ [_S \ldots [_S\ \mathbf{t}_{i_2} \ldots [_S\ \mathbf{t}_{i_1}\ [_S \ldots \mathbf{t}_{i_0} \ldots]]]]]$

The idea is that the relation between hierarchically adjacent traces in this configuration, such as the bolded pairs $\langle \mathbf{t}_{i_{n+1}}, \mathbf{t}_{i_n} \rangle$ in (515), or between the filler and its highest trace, bears *some* relation to the configuration in (512b), enough that it can be seen, at a quite abstract level, as an instance of the same tree-geometric relation depicted in (512b). And, just as in (512b), much depends on just what intervenes between the lower trace and the structure higher up. In particular, if the "landing site" for the movement leaves the higher trace in a position where a complementizer is present, it is stipulated that the lower trace is separated from the upper trace by a barrier which nullifies the connection between the two traces, and the movement fails. If this rather elaborate story were correct, then we would have an account of why the structure in (510b) is ill-formed when *that* is present: the presence of *that* was assumed to interrupt the linkage between the filler and the trace, with the resulting configuration failing the requirement that the "proper government" relationship must hold between traces on the one hand and either the filler or a higher trace left by the filler on its upward movement path. Note that the presence of the complementizer does not block extraction from the object position, since object NPs satisfy "proper government" by being in the privileged configuration with respect to its lexical selector via (512a); thus, examples such as *Who does John think that Terrence should hire t?* are correctly licensed.[4]

But now a new problem arises: adjuncts can freely move through higher clauses to filler positions, even when a clause marked with *that* intervenes:

(516) a. When [S do you suppose [S that **t** John will leave **t** on his next trip]]?
b. How fast [S would you say [S that **t** Mary can expect to run one hundred meters **t**]]?
c. Tuesday, [S I don't think [S that **t** we're doing very much of anything **t**]].

In all of these cases, the movement chain of the fronted filler passes safely through a complementizer-marked clause, yet the examples are good. The reason cannot involve licensing in the structure (512), since adjuncts do not appear as selected elements. These data appear to be clear counterexamples to the "privileged configuration" explanation.

To circumvent this difficulty, transformationalists added further wrinkles to the system. One quite-often-cited proposal, suggested in Lasnik and Saito (1984, 1992), consisted of two parts: (i) the stage at which adjuncts are licensed by filler-gap chains

4. "Proper government" was disjunctively defined as either local selection by a lexical head in order to satisfy a valence requirement or a specific local structural configuration (so-called antecedent government). No convincing arguments were ever provided to suggest that these two distinct conditions represented any kind of natural class.

occurs later than the point at which arguments of the verb are licensed, and (ii) *that* is "edited out" for purposes of chain licensing before this later stage (but not until argument licensing has been determined). It follows that at the stage where the status of the adjunct chain is determined, the structure of (516a) will look like (517):

(517) When [$_S$ do you suppose [$_S$ ~~that~~ **t** John will [$_{VP}$ [$_{VP}$ [$_{VP}$ leave] **t**] on his next trip]]]?

At this point, although *that* is present in the phonological representation, it is no longer visible at the particular syntactic location where adjunct traces are checked in terms of whether they satisfy the privileged-configuration criterion. Various versions of this by now extraordinarily complex and only rather vaguely spelled out scenario appeared during the later 1980s and early 1990s, but the essential features of the approach sketched in this section are preserved in later variants. By this point, the excessive complexity and epicyclic quality of the principal syntactic accounts on offer to explain islandhood had become evident even to many of those approaching the problem from a transformationalist perspective.

10.1.2.3 Ross's constraints under Minimalism The current Chomskian framework, the Minimalist Program (MP), officially inaugurated in Chomsky (1995), represents a dramatic break with the previous version of the Principles and Parameters architecture, with most components of the Government and Binding version that preceded it summarily removed from the analytic toolkit—a point that has been recognized both within and outside the Principles and Parameters research community. For example, Culicover and Jackendoff (2005, 93) note that "the MP lacks an account of most of the phenomena handled by GB/PPT [Principles and Parameters Theory] and other syntactic theories," citing on this point the following remark by Hilda Koopman (2000, 2), a prominent contributor to research in the Government and Binding framework:

> [C]ompared to the GB framework, the Minimalist Program led to relatively few new insights in our understanding of phenomena in the first half of the nineties. This is probably because it did not generate new analytical tools, and thus failed to generate novel ways of looking at well-known paradigms or expand and solve old problems, an essential ingredient for progress to be made at this point.

A showpiece example of Koopman's and Culicover and Jackendoff's point is the notion of government, which provided the crucial conceptual platform on which the *Barriers* framework rested and which was jettisoned completely in the MP. Elimination of government as an analytic concept rendered the notion of barriers and other components of that analysis unavailable. But the obvious need for a new source from which the descriptive generalizations originally proposed in Ross's thesis could be derived was largely unmet in the earlier phases of the Minimalist project; Chomsky (1995) mentions island phenomena only a handful of times, with no actual explicit

specification of island environments in terms of the new framework given anywhere. A decade later, Radford's (2004) textbook, *Minimalist Syntax*, offers nothing that sheds any further light on the origins of islandhood. Much of the syntactic literature for a considerable period after the appearance of Chomsky (1995) gives the strong impression that syntacticians were simply taking for granted the islandhood of the standard stock of configurations from previous incarnations of transformational grammar, with little concern about just how these were to be characterized under Minimalist assumptions.

Probably the most serious effort to derive island constraints from Minimalist assumptions is the analysis offered in Müller (2010), which makes critical use of the notion of "phase," referring, somewhat ambiguously, to both a set of category types and a subsequence of operations in a derivation upon whose completion the resulting object is "transferred" to LF for semantic interpretation and PF for prosodic realization and becomes opaque to further syntactic operations. In particular, in categories which have phasal status, the complement of the head of such categories undergoes transfer to SpellOut, leaving only the head itself and its specifier(s) syntactically active. For example, simplifying considerably, a transitive verb V may undergo the Merge operation with a DP to yield a VP, which in turn serves as the argument to an abstract "light causative verb" *v*. Successive applications of Merge give rise to the structure in (518).

(518) $[_{vP}\ DP_2\ [_{v'}\ v\ [_{VP}\ V\ DP\ ]]]$

In order to project this structure, Müller makes the innovative proposal that Minimalism incorporate the argument-saturation specification typical in categories in HPSG and CG. On this view, (518) is built up by saturating a DP argument requirement in V (notated in Müller (2010) as $V\langle D\rangle$); the resulting Merge is then $[_{VP}\ V\ DP\ ]$. The VP constituent so derived then saturates the VP specification of the light verb *v*, and finally *v*'s last-in DP requirement is satisfied by a DP which is Merged into *v*'s Spec position, giving rise to the structure in (518). A further derivation step adds $T\langle vP\rangle$, with an unsaturated *v*P argument requirement, and Merge then gives rise to the structure in (519).

(519) $[_{TP}\ T\ [_{vP}\ DP_2\ [_{v'}\ v\ \boxed{[_{VP}\ V\ DP\ ]}\ ]]]$

The boxed material in the last stage of the derivation represents the portion of the structure which has been transferred to the PF interface of the grammar and is no longer available for syntactic operations. Note that the saturated phasal head *v* and its specifier remain outside the box, notating their continued syntactic accessibility to processes dictated at higher levels of structure as the derivation proceeds. Under standard MP assumptions, this accessibility persists until the point at which the *v*P in (518) is itself part of a structure which is transferred to the syntactically opaque interface domains. The derivation outlined for (519) will be part of the syntactic history of an ordinary English sentence such as *John saw Mary*.

Given the fact that argument lists have been part of nontransformational frameworks from the beginning, it might seem that they add little in the way of explanatory resources that could be brought to bear on the origin of islandhood. In Minimalist analysis, however, certain additional mechanisms are available which Müller exploits for this purpose—in particular, the treatment of extraction as an instance of valence satisfaction. In many MP analyses, *wh* movement is a possibility dependent on the presence of what is commonly referred to as an EDGE feature in the specifications of the phasal head *v*. This feature, sometimes written [*wh*-P(eriphery)], will on some versions of Minimalist syntax be a freely added specification on *v*, in effect licensing an "extra" specifier position arising from the movement of the transitive object *who*. Müller adapts these features to his account of islands, treating them literally as valence specifications, displayed as an HPSG-style list, licensing "internal Merge" operations in the same way that *v*'s VP and DP argument features license ordinary syntactic structure building via "external Merge," based on material in the numeration.

Permitted cases of extraction, such as *Who did John think that Mary criticized?*, point to the *wh* word's escape from the lower VP complement to *v* at some point prior to the transference stage exhibited in (519). An edge feature must therefore be present in *v*'s valence list at that point, allowing a *wh* phrase in the object position to reMerge as the specifier to *v*—a position from which it will be syntactically accessible to subsequent reMerges, that is, movement to a still higher Spec, even after the subsequent transfer of *v*'s VP complement to the SpellOut interface.

This general approach to extraction appears in principle somewhat perverse, given the island facts (as assumed in mainstream syntactic theorizing over the past half century). The material which becomes syntactically opaque at the end of each phase is the complement to the phasal head; the specifier of that head is still syntactically active. A *wh* phrase contained in the former can in principle become trapped there via LF/PF transfer, while the same phrase contained in the specifier is still accessible to the movement-triggering features added during successive applications of Merge. In other words, the default situation is that the complements of *v* are predicted to be more likely to have island status than the specifier of *v* which becomes the subject of the sentence after raising to [Spec, T]. Thus, the phase-based approach to syntactic islands appears to fly in the face of the universally observed pattern that subjects are more difficult to extract from than complements, all other things being equal.

Müller's proposal offers, however, a detailed set of premises which are claimed to correct this anomaly and yield what has been called the Condition on Extraction Domains (CED). The CED, introduced in Huang (1982), essentially posits a unitary explanation for the putative prohibition on extraction from subjects and adjuncts based on the distinction between complements and non-complements of a selecting head. Objects and other valents of a verb are thus extractable, but dependents which do not qualify

as directly selected arguments are not. The key assumptions that Müller makes, which supposedly achieve the needed result, are the following:

1. The order of valence satisfaction is strictly determined by stacks, or lists, of features borne by all heads (with later-satisfied features written to the left of any that must be saturated earlier);
2. "edge" features—*wh* phrase attractors—can be added freely to phasal heads (which Müller takes to include all category types), as supposedly required by languages with multiple *wh* extraction to the same CP;
3. but only while these heads are still syntactically active, that is, have not had all their arguments saturated; and
4. when an "edge" feature is added to a head's list of valence features, it is added just to the right of the top of the list—that is, as *v*'s penultimate argument specification—and hence must be saturated before the head's specifier requirement.

Specifiers are then the arguments that saturate the "top" specification in this list and hence are Merged with their selecting heads as the very last step in the saturation of the latter. The result of these assumptions is that as long as the specifier argument of *v* is not saturated, a *wh*-attracting edge feature can still be legally added. The steps involved that underwrite such a derivation will then be something along the lines of (520):

(520) a. $[_{vP}\ v[\langle \text{DP} \rangle]\ [_{VP}\ \text{V DP}\ ]] \xrightarrow{by\ 2\text{–}4}$
b. $[_{vP}\ v[\langle\ \text{DP}, \textit{wh}\text{-P} \rangle]\ [_{VP}\ \text{V DP}\ ]]$

This addition to the *v* head's valence list has the consequence (by 1) that before *v* can saturate its Spec position argument, some *wh* phrase α must be attracted to the left edge of *v*′ to "clear" the *wh*-P feature; otherwise this uninterpretable feature will crash the derivation at LF. But where can α originate? Clearly, it cannot be internal to *v*′'s specifier, destined to be the subject of the sentence, since at the stage of the derivation depicted in (520b), that DP specifier does not yet exist and can only be Merged into the derivation *after* the edge feature added to *v* in (520b) has been saturated and eliminated from *v*'s argument list. Hence, the only way to complete the saturation process yielding *v*P is to move α from inside VP to eliminate the *wh*-P specification. But of course once this occurs, and the specifier DP is then added to form *v*P, no further edge features can be added, since now *v*'s valence list is completely saturated and *v* is inert, meaning no further edge features can be added, as per 1. Hence, nothing internal to *v*'s specifier—which as noted will surface as the subject of TP—can be moved.

In effect, under Müller's analysis, extraction from subject position faces a kind of catch-22: by 1 no content can be instantiated in subject position till the penultimate edge feature specification has been saturated, but after that, instantiating such content in *v*'s Spec position deprives any *wh* phrase within that content of an escape hatch to

the edge, since, by 3 no new edge features can be added at that point. Müller's conclusion appears to be that the only alternatives are either to add an edge feature and move material out of the VP (exclusively) or to saturate all of *v*'s arguments and not add an edge feature, with nothing moving out. No matter which choice one makes at this point, all subjects thus wind up displaying island properties—as has been standard doctrine at every phase of transformational grammar following the appearance of Chomsky (1973).[5]

But the second option—saturate *v* without adding any edge features to its argument list—has, as Boeckx (2012) points out, a rather different possible outcome from the one Müller envisages. Suppose that, instead of inserting a [*wh*-P] feature into *v*, we simply saturate the DP argument of *v* with a constituent which happens to contain a *wh* phrase, call it *wh*β, notating this DP specifier as $\text{DP}_{\ldots\beta\ldots}$. Nothing in Müller's proposal blocks this move, which simply represents satisfaction of the *v* head's specifier valence requirement by a phrase such as *a painting of which mountain* and which by 4 will yield the structure in (521):

(521) $[_{v\text{P}}\ \text{DP}_{\ldots\beta\ldots}\ [_{v'}\ \boxed{v\ [_{\text{VP}}\ \text{V DP}\ ]}\ ]]$

The derivation can freely proceed as sketched earlier. It is true that movement to the edge of *v*P is prohibited at this point. Suppose, however, that we introduce a functional head T from the numeration, but, prior to Merging it with the *v*P argument on its valence list, we add an edge feature, as allowed by 2, so that the head has the form T[⟨DP,*wh*-P,*v*P⟩], allowing it to combine first with *v*P, then with a filler, and finally with its own specifier DP as the last-in argument, per Müller's proposal. The result is the structure in (522):

(522) $[_{\text{TP}}\ \text{T}\langle\text{DP},wh\text{-P}\rangle\ [_{v\text{P}}\ \text{DP}_{\ldots\beta\ldots}\ \boxed{v\ [_{\text{VP}}\ \text{V DP}\ ]}\ ]]$

In order for the derivation to yield an acceptable input to LF, the edge feature in T must be canceled—but this can now be achieved simply by moving wh_{β} out of $\text{DP}_{\ldots\beta\ldots}$ to the left edge of the TP, yielding (523):

(523) $wh_{\beta}\ [_{\text{TP}}\ [_{\text{T}'}\ \text{T}\ [_{v\text{P}}\ \text{DP}_{\ldots t_{\beta}\ldots}\ [_{v'}\ \boxed{v\ [_{\text{VP}}\ \text{V DP}\ ]}\ ]]]]$

And now $\text{DP}_{\ldots t_{\beta}\ldots}$ can itself undergo A-movement into the DP specifier position of TP:

(524) $wh_{\beta}\ [_{\text{TP}}\ \text{DP}_{\ldots t_{\beta}\ldots}[_{\text{T}'}\ \text{T}\ [_{v\text{P}}\ \text{t}\ [_{v'}\ \boxed{v\ [_{\text{VP}}\ \text{V DP}\ ]}\ ]]]]$

5. Very few analyses in the P&P tradition acknowledge the markedly gradient nature of extraction from subjects, which Ross was well aware of, and which led him to restricted subject islandhood to clausal constituents, where the restriction on movement appeared to be more sharply defined. We return to this issue in section 10.2.

As Boeckx notes, the import of this possible line of derivation is that even assuming, as per Müller's scenario, that you can't move wh_{β} to [Spec,*v*P] (since, ex hypothesi, you haven't added an edge feature to *v*'s argument list), you can *still* carry out unbounded extraction from the subject position by simply waiting till the next operation of Merge with an edge-argument-seeking head. In fact, exactly such a derivation must be involved in the extraction *of* subjects, which in the absence of a preceding complementizer is typically unexceptionable. It seems, then, that free instantiation of edge features on functional heads, cited by Müller as necessary to allow multiple *wh* fronting as reported in much of the literature over the past two decades, winds up making it very difficult to enforce subject islandhood.

The consequences for Müller's proposal are thus quite drastic. As Boeckx (2012, 66) points out,

> invoking the PIC [Phase Impenetrability Condition] *alone* cannot account for the Subject Condition. For that to be feasible in a PIC-based story, the external argument should be included in the transfer domain, but that would only happen if the *v*P *in full* were transferred upon merger of T. (emphases added)

In short, once a phase head is completely inert—that is, there are no valence or other features triggering syntactic operations left to discharge—the *whole* phase would be transferred to the interface, not just the head's complement. This at least would put complements and subjects on an equal footing, so that the final step in the derivation would then be not (522) but (525):

(525) $[_{TP}\ T\langle DP,wh\text{-}P\rangle\ \boxed{[_{vP}\ DP_{\ldots\beta\ldots}[_{v'}\ v\ [_{VP}\ V\ DP]]]}\ ]$

Thus, even if T contains a feature attracting a lower *wh* phrase to [T,Spec], it would not be able to reach inside the now opaque *v*P to a target constituent within the specifier. But this of course means that nothing else could move either. On Boeckx's scenario—in effect, a reductio of Müller's proposal—a *wh* phrase within the VP complement will move to the [Spec,*v*P] position, awaiting its chance to move to the Spec of TP and finally to [Spec,CP], but that chance will never come: as soon as *v*P is Merged with T, it becomes syntactically opaque, having undergone transfer to the prosodic interface, and anything that was within it, even in its Spec position, is now inaccessible. The rather undesirable result is the prediction that there are no extraction phenomena.

Given that in this case the cure seems to be worse than the disease, we have to conclude that this latest attempt to derive subject islandhood from general principles of the current version of the P&P framework is no more successful than previous efforts.[6] In any case, the CED only accounts for a portion of Ross's original constraints. Presumably, the supposed prohibition on extraction from relative clauses would be an instance

6. And it is obvious that the same is true of adjunct-internal *wh* phrases, notwithstanding Müller's (2010, 46) observation that

of Müller's adjuncts-as-specifiers analysis (see footnote 6), but the ill-formedness of extractions from complement clauses within complex NPs, for example, does not seem to receive a natural account from either the standard CED or Müller's phase-based derivation of it based on the interplay of edge features with other aspects of the already highly elaborate syntactic machinery provided by the Minimalist Program.

On the whole, then, it is difficult to see the Minimalist effort to derive island effects as natural consequences emerging from the foundations of syntactic theory as any more successful than earlier attempts to achieve that objective. Far from providing conceptual and analytic resources which make sense of the diversity (and, as has become increasingly obvious, the variability) of island effects, the current mainstream grammatical architecture appears to have given up the chief tools employed in its previous avatars to capture these effects, without providing effective replacements: specified syntactic configurations are excluded in the definition of island boundaries, even in the relativized form of the *Barriers* framework, but no plausible alternatives have been proposed and generally accepted within the research community built on Chomsky (1995) and its elaborations. Rather than subsuming Ross's descriptive characterizations of island environments under deep, natural principles, restrictions on extraction have taken on the appearance of a series of new epicycles attached to previous, already overly complex characterizations of island configurations. In the sciences generally, this sort of situation is typically taken as a warning that the research strategy leading up to it has been on the wrong path and needs urgent—and drastic—rethinking.

10.1.3 Challenges to Syntactic Islandhood

It is therefore not particularly surprising that, within the past decade and a half, a fundamental shift in perspective has become increasingly evident in syntacticians' attitudes toward constraints of the kind discussed above. This new line of thinking breaks decisively with the assumption that the prohibitions on filler-gap connectivity sketched above are syntactic in nature. It now seems clear that the three decades following the appearance of Ross's thesis have not provided anything like confirmation—or even

---

> [t]he barrier status of adjuncts follows immediately if we assume that adjuncts are to be reanalyzed as last-merged specifiers of special functional projections (see Alexiadou 1997; Cinque 1999). The reasoning is then identical to that given before for subjects.

No actual argumentation is provided to motivate the assumption that adjuncts are components of "special functional projections," or to document their syntactic parallelism with other syntactic objects that are generally assumed to be specifiers; rather, the primary motivation for this treatment of adjuncts seems to be a need to provide some generality for the elaborate and otherwise special-purpose machinery proposed earlier in the paper to account for a single island environment. For a more recent Minimalist analysis of adjunction that urges an analysis in which complements and specifiers, which are introduced into a derivation by Merge, are radically distinguished from adjuncts—which are *not*—see Hunter (2015). There appears to be little or no current acceptance among Minimalists for the identification between adjuncts and specifiers that Müller's argument here depends on.

an increase in plausibility—for the position that such prohibitions, to the extent that they apply at all, arise from restrictions within the combinatoric system itself. On the more recently developed approach to the data, the unacceptable (and sometimes uninterpretable) sentences in question are syntactically well-formed. That is, nothing is *structurally* deficient about them; rather, the problems that arise are due to independent (but occasionally interacting) sources. The ill-formed examples can be shown either to violate restrictions imposed by some non-syntactic part of the grammar (e.g., prosody, semantics, and pragmatics) or to incur costs that have little to do with the grammar directly, but rather with the nature of psychological mechanisms (depending, for example, on certain kinds of memory) required to process linguistic information in real time. Such processing events are easily derailed—possibly to the point of failure—by certain kinds of interference and complexity. On this view, it is the latter set of interacting factors, rather than any kind of prohibitions on structures, which give rise to the ill-formedness previously attributed to configuration-based constraints.

10.1.3.1 Island effects via processing: Kluender (1998, 2004); Chaves (2013); Chaves and Dery (2019) The far-reaching reassessment of islandhood just alluded to was in no small measure due to the work of Robert Kluender, whose innovation takes as its point of departure a well-known effect called *center embedding*, arising from the repetition of specific structures underlying the form of a certain kind of relative clause illustrated in (526).

(526) The man [$_S$ the host knew __] left the room.

Clearly, there is nothing wrong with this structure (the name *center embedding* comes from the fact that the relative clause is embedded after the nominal subject head and before the VP). But when this structure is iterated—that is, when the relative clause subject itself contains a center-embedded structure—things get bad very quickly. Thus, a typical paradigm illustrating the gradual decline in acceptability is displayed in (527):

(527) a. The man left the room.
b. The man the host knew left the room.
c. The man the host the intelligence people investigated knew left the room.
d. The man the host the intelligence people the reporter interviewed investigated knew left the room.

(527b) is still fine, but (527c) is considerably more difficult to process without a very deliberate use of intonation to make clear the intended structure; and (527d) is still worse. Such examples were known at quite an early stage of syntactic research (see, e.g., Yngve [1960, 460], where center-embedding data are referred to as examples of "regressive structures") and became a somewhat clichéd piece of evidence that whether or not some string of words was acceptable only indirectly implicated its status as a

possible output of the rules of the grammar. The assumption for many years has been that there is something about the nature of center-embedding structures which taxes the short-term memory resources available to keep track of linguistic structures. In the case of (527), these effects make it difficult for the hearer to link particular nominal structures in the string preceding the verbs to the associated gap site. The particular configuration in (527d) requires a certain correspondence pattern—crossing dependency—to hold between the nominal heads and their respective correlated verbs and gaps. Crossing dependencies are independently known to invoke processing difficulties in other domains—an often-cited example is the difficulty of interpreting patterns of crossing filler-gap linkages addressed by Fodor's (1978) Nested Dependency Condition.

All this was old news even in the 1970s. But Kluender's remarkable results during the 1990s made it clear that much of the thinking about center-embedding constructions had missed possibly the most important point: these structures could be improved significantly by certain purely lexical adjustments. A hierarchy of intelligibility is shown in (528), where we follow Kluender's (1998) usage in notating relative acceptability with inequality markers, so that X < Y indicates greater acceptability for Y as compared with X.[7]

(528) a. The woman [the man [the host knew __] brought __] left.
b. < The woman [that man [the host knew __] brought __] left.
c. < The woman [a man [the host knew __] brought __] left.
d. < The woman [someone [the host knew __] brought __] left.
e. < The woman [someone [he knew __] brought __] left.
(Kluender 1998, 254)

The improvement between the first and the last of these examples is quite striking. Yet nothing about the structure has changed. What has happened to yield this unexpected improvement?

In passing from (528a) to (528d), note the progressive reduction in the definiteness of the NP subject in the highest relative clause. *The man* conveys uniqueness more strongly than *that man*, while *a man* indicates no uniqueness at all. In (528d) we find

7. Kluender attributes the discovery of such examples to Thomas Bever (1970). But careful inspection of the latter source fails to reveal any data pattern along the lines of (528). Bever did note that the acceptability of center embedding is variable to some extent and reflects lexical choice, but the examples he provides have no bearing on the kind of effect reflected in the acceptability gradient in Kluender's paradigm. What is critical about that set of judgments is not only the fine-grained gradient effect Kluender notes but, most significantly, the relationship of this gradient effect to referential specificity—the same variable that, as we discuss directly, is responsible for the dramatic amelioration of island effects. It is this connection—that amelioration of a clearly processing-based effect is faithfully mirrored in the judgments of putatively syntactic island conditions—that constitutes the initial empirical basis for Kluender's breakthrough, and it appears that the crucial paradigm in (528) which establishes the linkage between the center-embedding facts and the gradience of island effects represents an original observation by Kluender himself.

a completely indefinite NP, *someone*, from which all we can infer is that an unspecified human being is being referred to. Finally, in the final example, the subject of the lower relative clause is replaced by a pronoun *he*, carrying far less information than the definite NP *the host*. Cumulatively, what has happened is that the intermediate NPs between the highest NP and the lowest gap site, into which this NP must be interpreted, have been in some sense diminished in terms of their information content, in particular, their referential specificity. This reduction has the concomitant effect, of course, of increasing the referential specificity of the filler which has to go the furthest distance to find its gap site, relative to any intermediate filler(s) occupying positions at clause boundaries along the way.

Apparently, then, the problem with center-embedded relative clauses is not structural in essence; while there are structural aspects to it, such as the location of clause boundaries in relation to where the various NPs occur, these structural factors alone do not determine the difficulty of psychologically processing such relative clauses successfully. Rather, they only create the possibility for a high degree of difficulty in that task, depending on what else is going on, and it is that "what else" that turns out to play the critical role. The determining factor seems to be the degree of difficulty that the hearer encounters in trying to link a filler to an increasingly deeply embedded gap site, when such linkage encounters processing tasks along the way that distract the parser. If these tasks cumulatively place too much demand on limited processing resources, the speaker in effect loses track of the filler in topmost position, making completion of the linkage impossible. Linkage of fillers to gap sites, in the case of center-embedded relative clauses, is thus more successful to the degree that the filler can be made more informative and its intervening NP competitors made less so.

One must bear in mind that from the outset, as already noted, the diminished acceptability/intelligibility effects in center-embedding constructions were taken to reflect facts about the mechanism by which speakers establish the relationships among parts of the sentence required for interpretation, rather than whether or not such relative clauses were sanctioned as legal by the grammar. Kluender's discoveries about the improvements in comprehensibility in these constructions were therefore altogether plausible, even expected, *if* we make certain specific assumptions. We can say in advance that the "repair" strategy exhibited in (528) makes sense on the assumption that judgments of a legal structure which is psychologically difficult to process in real time can be dramatically improved through means which have nothing to do with structural factors. Certainly the center-embedding gradient discovered by Kluender constitutes evidence for such improvement. If the same kind of gradient profile were reflected in the case of classic islandhood phenomena, then general considerations of parsimony appear to put a substantial burden of proof on advocates of a syntax-based account in which ungrammatical utterances, corresponding to structures never generated by the grammar,

somehow acquire syntactically acceptable and semantically interpretable status by as-yet-unknown mechanisms. We return to this point below.

These methodological considerations have immediate relevance in the face of the data such as the following, exemplifying the amelioration effects with CNPC.

(529) a. What do you need to find the professor [who can translate __]?
b. < What do you need to find a professor [who can translate __]?
c. < What do you need to find someone [who can translate __]?
d. < Which article do you need to find someone [who can translate __]?
e. < Which article do you need to find someone [to translate __]?
(based on Kluender [1998, ex. (12)])

Here, (529a) is a typical example of CNPC violation and, unsurprisingly, strikes the ear as extremely awkward at best. But acceptability increases gradually—and markedly—in (529b–e). Note moreover that the factors which lead to this improvement are the very same ones that yield the successful examples of center embedding in cases such as (528): acceptability gradually increases by making the intervening NPs referentially more specific ((529a–d)), and we obtain a nearly impeccable sentence (529e) by changing the finite tense to an infinitive (where the finite/nonfinite contrast is arguably a counterpart of definiteness in the verbal domain; cf. Partee 1984).

Similar patterns of improvement can be observed with many other examples as well. For example, in (530), decreasing the complexity of intervening NPs on the boundary of the lower clause (*a random assortment of tradespeople* → *anyone*, *friends of whom* → *whom*) leads to a distinct improvement in acceptability.

(530) a. John, I can't think of a random assortment of tradespeople [friends of whom like __]. <<
b. John, I can't think of anyone [who likes __]?

Another example strongly suggesting that what's involved here is a processing-based effect is the contrast in (531).

(531) a. Euthanasia is a topic which$_2$ I can never find anyone who$_1$ I can argue with t_1 about t_2 . <
b. Euthanasia is a topic$_2$ I can never find anyone$_1$ to argue with t_1 about t_2 .

Note that here, (531a) is already fairly acceptable, despite the fact that it involves a CNPC configuration. This already casts considerable doubt on the syntactic status of the CNPC. But the example can be improved still more by replacing the finite modal to an infinitive and removing the two explicit relative pronouns (*which* and *who*), as in (531b). Since (531a) is already a good (enough) example, the increased acceptability in (531b) can be nothing other than a processing effect.

In short, CNPC fits the profile not of a structural condition (blocking the licensing of word strings which violate that condition) but of a performance effect which inhibits the processing of a legal sentence (reducing its comprehensibility by interfering with the identification and retention of possible reference targets and preventing the latter from surviving long enough in the processing task to be linked to the filler).

But the explanatory reach of these discoveries extends well beyond the case of CNPC. Let's reconsider the *wh* island cases we considered earlier, where few would argue that (532) is an acceptable example.

(532) Who did you wonder what Mary said __ to __?

(533) a. Who did you wonder what you should say __ to __?
b. Who did you wonder what to say __ to __?

(534) a. Which of the people at the party did you wonder what you should say __ to __?
b. Which of the people at the party did you wonder what to say __ to __?

(535) a. John is someone who I never know what I should say __ to __.
b. John is someone who I never know what to say __ to __.

Acceptability increases gradually in (533)–(535), and in each pair, the infinitive version is somewhat better than the finite clause version involving a modal. This gradual increase of acceptability is completely expected on an extragrammatical account of island effects. There are two key factors involved: pragmatic felicity (of the sort we discuss in more detail in the next subsection) and processing factors of the sort discovered by Kluender (such as referential specificity).

Note first that the content of the question is much more natural in the examples in (533) than (532) (where one can imagine contexts in which the sentence might be uttered much more easily for (533) than for (532)). The slight improvement in the infinitive version (533b) over the finite modal version (533a) can be attributed to something analogous to referential weight in the verbal domain: while the modal form is already less "definite" than simple finite past tense (where the latter is about some specific past event, whereas modal statements are about possible situations), infinitives (with no overt subjects) arguably carry still less specific content. (534) improves on (533) by making the extracted material referentially heavier, and (535), involving relativization, improves still more by eliminating any presuppositional content associated with an interrogative, which increases processing burdens (invoking extra accommodation process unless appropriate contextual background is already provided).

The acceptability of these island violation examples and the gradual effect of amelioration observed in (533)–(535) correlating with factors independent of syntactic structure seriously call into question the elaborate machinery that has standardly been posited in the syntactic literature to account for the alleged ungrammaticality of such

structures and point instead to a non-syntactic treatment in terms of performance and pragmatic factors.

It is of course true that such cases were not entirely unmentioned in the mainstream generative literature, and it is instructive to consider how they were approached. Examples such as the contrast between (533b) and (534b) were recognized in, for example, Pesetsky (1987), who argued that they reflect a difference in *D-linking* ("discourse-linking") properties, with D-linking defined as a restriction of the possible answers to the question to members of some contextually salient set. D-linking was subsequently adopted in Cinque (1989) and Rizzi (1990), but as Kroch (1989) pointed out, the critical distinction underlying the notion was semantically not particularly well-founded: a supposedly non-D-linked "bare" *wh* word such as *who* in (533b) "*does* constrain answers to membership in fixed sets. The sets are only rigidly (i.e., semantically), and very broadly, rather than contextually, and more narrowly, defined" (emphasis added). More fundamentally, it is not clear why an essentially semantic/pragmatic distinction should allow any kind of exception to a purely syntactic restriction.[8]

In subsequent work, Kluender extended his processing-based account of CNPC and *wh* islands to subject islands, arriving at the conclusion that "the same general processing factors will apply in a slightly different way to subject islands to render them difficult-to-impossible to interpret" (Kluender 2004, 102). The essence of his proposal is that a subject island—necessarily already a "heavy" structure with significant referential content (since it must contain enough material to constitute a context for the extracted filler)—imposes the kind of processing costs that, without amelioration, create islandhood effects similar to those observed in relative clauses, *wh* islands, and other cases covered in Ross's (1967) thesis. Kluender presents evidence from a wide range of data sources—adult and child language production research, as well as work with elderly speakers on processing difficulties—that establish that among the various possible sites for phrasal constituents to occupy, subject position presents special parsing difficulties quite independently of whether or not its occupant contains a gap. The result is, predictably, a strongly negative assessment of subject-internal gaps. However, as Kluender (2004, 491) demonstrates, "Subject islands can in fact be made

8. A model that incorporates gradience in acceptability in the architecture of the grammar itself may provide a suitable framework for explicitly modeling such interactions between semantic and pragmatic factors and the combinatoric component of syntax (see, for example, Keller [1998] and the papers collected in Fanselow et al. [2006]). While such approaches seem to generally share the same motivations and goals with processing-based accounts of island constraints advocated by Kluender (1998), Michel (2014), Chaves (2013), Hofmeister and Sag (2010), and others, it is currently an open question whether/to what extent this type of gradient architecture of grammar is compatible with the processing-based accounts discussed in the present chapter (the latter of which seem to generally assume a more traditional distinction between narrow grammar [violations of whose constraints are strictly categorical] and the processing component), either at the specific level or at the more general, foundational level.

more transparent to extraction" through the same kind of techniques shown in Kluender (1998) to materially improve CNPC violations (eliminating referentially heavy subjects, or indeed any subjects at all, in the island context; changing finite [i.e., semantically definite] tensed clauses to infinitive clauses; increasing the linear distance between the gap and the filler, and so on). One observation of particular interest in Kluender's findings is that "[semantic] association of the filler with . . . the main assertion of the current sentence, i.e., the main clause predicate," as well as the gap site, typically results in significant improvement in the acceptability of extractions from subjects—a point that is taken up and developed in Chaves (2013) and Chaves and Dery (2019).

In his 2013 paper, Chaves begins by reprising the parsing model in Kluender's 2004 paper, observing that, in the face of complex subjects, "the speaker's memory resources are strained sooner in the sentence, and longer, since those resources are not available for processing the remainder of the sentence," so that "filler-gap dependencies in them are harder to maintain in working memory without additional support" (Chaves 2013, 14). But these considerations in themselves do not explain why processing gaps within subjects should become so much easier in parasitic gap constructions or why certain rather short examples (such as **What did the owner of __ sneeze*), involving minimal processing burdens, are nonetheless judged severely ill-formed. Chaves's (2013, 28–29) solution hinges on the following premise:

> Given that processing complex subjects is cognitively more strenuous than processing complex objects, and that certain pragmatic conditions restrict the use of filler-gap dependencies in general, . . . speakers avoid the use of sentences with subject-internal gaps. This leads to extremely low (near zero) frequency. In turn, extremely low frequency may cause the language processor to develop a conventionalized processing heuristic: expect gaps to be in the verbal structure, not in the subject phrase. . . . This gapless subject expectation cannot be seen as a grammatical condition. Precisely because it is a parsing expectation rather than a grammatical rule, it can be dampened by the presence of prosodic, pragmatic and contextual cues that signal the correct parse.

Based on this background, Chaves documents a wide range of acceptable subject gaps, noting the effect of three factors mentioned by Kluender in distinguishing such cases from ill-formed examples: (i) prosodic cues, (ii) referential specificity, and (iii) a relevance relationship between the filler and the gap.

On this type of account, parasitic gaps enable the establishment of connectivity between the filler and the subject gap in two ways. On the one hand, the parser's expectation-based search for a nonsubject gap site—that is, one internal to the VP—is rewarded. On the other hand, as Chaves (2013, 37) puts it, "multiple gaps in close proximity . . . reactivate the same filler," strongly facilitating the backtracking necessary to link the filler to the subject gap venue.

The line of reasoning pursued by Kluender, Chaves, and a number of other theorists based on a range of experimental results thus provides a major step toward reducing is-

land effects to the interaction of largely extragrammatical factors. Nonetheless, a number of other lines of inquiry have been pursued during the same recent phase of investigation. For example, in the processing-based accounts we have summarized above, the role of pragmatic and semantic factors in the processing-based work is largely confined to the influence of such factors in enabling or inhibiting the course of parsing in an island environment. But in other work, these factors are argued to independently contribute effects which account for the apparent islandhood of various classes of extraction prohibitions. This line of research, to which we now turn, was largely pioneered in research by Kroch (1989), highlighting the role of speech act presuppositions, pragmatically determined implicatures, and other kinds of felicity conditions on discourse that are altogether independent of syntactic configuration.

10.1.3.2 Island effects via discourse pragmatics: Kroch (1989); Kehler (2002); Kubota and Lee (2015); Culicover and Jackendoff (2005); Oshima (2007) Kroch (2007) argued that in the case of ill-formed data such as (536), a failure of presupposition rather than any kind of syntactic condition was responsible for blocking the "long movement" of an amount-quantified NP:

(536) *How much did Bill wonder whether to pay __ for the book?

The essence of Kroch's proposal is that the question in (536) presupposes the existence of a specific dollar figure that Bill had in mind as a possible purchase price for some book and was wondering whether to pay exactly that amount for it. Kroch suggests that such a state of affairs is pragmatically very implausible, hence unlikely to occur to anyone hearing (536) as part of whatever discourse background was involved. He shows that in contexts where such presuppositions are actually part of the pragmatic context, interrogatives such as (536) become markedly better, even unexceptionable. This analysis should probably be credited for opening what has become a very active and widely accepted line of explanation for many island effects—one which takes them to be emergent phenomena, reflecting a basic conflict between the interpretation associated with extraction on the one hand and the broader semantics of the constructions from which the extraction has occurred on the other.

A particularly persuasive illustration of this line of analysis is the variety of evidence pointing to the origins of the CSC as a fatal deviation from contextual expectations not unlike the presupposition failures associated with interrogatives which Kroch posited for *wh* extraction of amount or quantity *wh* phrases. Certain coordinate structures support extractions which fail to apply to all the conjuncts in a coordinate structure. Consider, for example, the following cases:

(537) a. This is the house that I plan to dress up as a water use inspector and break into __.

b. How many political debates can you listen to __ and not become completely cynical?
c. This is the cereal that I plan to eat __ for breakfast every morning and live to be 100.

These data were discovered in the 1970s and '80s. See, for example, Schmerling (1972), Goldsmith (1985), and Lakoff (1986), which offer quite different accounts of the possibilities reflected in (537). On the one hand, Goldsmith assumed the correctness of Ross's syntactic formulation of the CSC but proposed that examples such as those in (537) reflect not coordination but subordination; on the other hand, Lakoff argued that the data in (537) are acceptable because they reflect a single complex predication on the extracted element. In Lakoff's view, the CSC could not be a purely syntactic constraint, since it evidently can be overridden by conditions which he characterized as semantic and hence cannot be structure-dependent; rather, it must be reformulated as a semantic requirement that, in effect, only a unitary predicate can apply to an extracted referent, where a condition on a propositional function being unitary is that all of its parts are in some respect relevant. *To listen to X and not become completely cynical* denotes a unitary predicate, in Lakoff's view, because what is being predicated crucially depends on the satisfaction of the truth conditions on both conjuncts. In contrast, (496b) displays a propositional function comprising two conjuncts, the second of which—expressing the fact that the hearer goes to night classes on the days when he doesn't play poker—is irrelevant to the definition of the set of individuals with whom the hearer plays poker on the other days. Thus, the acceptability of (537) and the unacceptability of (496b) are said to reflect the difference between the two clauses with respect to the relevance of the denotation expressed by the gapless conjunct vis-à-vis the propositional function carried in the clause containing the extraction site.

Despite the persuasiveness of Lakoff's examples—which do indeed make a syntactic source for the CSC appear very unlikely—the concept of relevance that he calls on and its relation to the technical notion of the predication relation are somewhat elusive, as noted by Kehler (2002, 111). Without making this point more precise, Lakoff's proposal as it stands faces the charge that it is devoid of predictive power. Indeed, Kehler's (2002) own account, which links exceptions to the ATB pattern to a wide range of other phenomena in the syntax-semantics interfaces of Gapping, ellipsis, anaphora, and the semantics of tense, could be viewed as offering a particular line of refinement of Lakoff's by making more explicit the role of *coherence* in determining the well-formedness of discourse.

The essential point of Kehler's account is that extraction out of coordinated VPs have to obey constraints imposed by the coherence relations between the two conjoined constituents. Building on David Hume's philosophical work, Kehler identifies three basic coherence relations: Resemblance, Contiguity, and Cause-Effect. In the case of Re-

semblance, one can identify sets of participants in two situations as bearing the same semantic roles with respect to either the same predicate (*resemblance*) or to predicates which are in some clear sense opposites. The Cause-Effect coherence relation holds between two situations where one is, roughly speaking, a necessary precursor to the other. Contiguity between situations is a relationship in which one is the occasion, pretext, or noncausal precurser to the other.

The badness of (538) (= (499a)) is then straightforward.

(538) Who does John like (*Ann) but Mary always criticize __.

Here, there is no inherent semantic relation between the two conjuncts, either in the Cause-Effect type or the Contiguity type. The use of *but* as the conjunction marker suggests a contrastive relation, which is a type of Resemblance relation. In this case, extraction out of a single conjunct is prohibited since the presence of a filler-gap linkage in one conjunct versus an absence thereof in the other conjunct breaks the "equal" status imposed on the two conjuncts by the Resemblance discourse relation. Thus, when the discourse relation is Resemblance, filler-gap relations into a coordinated structure must apply across the board.

In contrast, consider (539):

(539) This is the house that I plan to dress up as a water use inspector and break into __.

The specific relationship between the first and second clauses is something we might call "precursor/outcome": the implication is that a particular disguise (water use inspector) will expedite an illicit activity (a house break-in), which makes the sentence coherent as an instance of Contiguity. Unlike the Resemblance relation, Contiguity does not require the two conjuncts to have "equal" status to each other. This means that asymmetrical extraction does not violate any constraint and hence is predicted to be perfectly acceptable, which indeed is the correct prediction.

A similar account goes for cases involving the Cause-Effect relation, such as the following:

(540) This is the cereal that I plan to eat __ for breakfast every morning and live to be 100.

Here, the underlying relation that makes the whole sequence coherent is that the habit described in the first clause is the cause which (supposedly) brings about the consequence described in the second clause. Just like the Contiguity relation, the Cause-Effect relation does not require the cause and the result to have an equal status (if anything, they are in an asymmetrical relation in that the result would never obtain without the cause). Thus, asymmetrical extraction is fine with the Cause-Effect relation too.

A useful follow-up to this last point is that any nonstructural change which alters the interpretation possibilities of a clausal discourse environment from an unequivocally parallel interpretation to a means-end, cause-effect, or some similar nonsymmetrical interpretation will allow the same structure to support non-ATB extraction. Compare, for example, the following:

(541) a. I caught a bus last week and attended a lecture this morning.
b. ??*Which lecture did you catch a bus last week and attend __ this morning?

(542) a. I caught the bus and attended a lecture this morning.
b. Which lecture did you catch the bus and attend __ this morning?

Suppression of the first conjunct's temporal adjunct *last week* makes it possible to plausibly construe the VP *catch the bus* as identifying an event which is something like the means to the end of attending the lecture, thus ensuring the possibility of non-ATB extraction. For a more detailed discussion of this kind of contrast in possibilities, see Kehler (2002, 129–132).

A further prediction follows. If what non-ATB extraction hinges on is (non-)parallelism in the semantic relations between the coordinated constituents, rather than the particular meaning of the coordinating particle, we would expect to see the same correlation between parallel readings and ATB extraction on the one hand and between nonparallel readings/non-ATB extractions on the other in the case of conjunct particles other than *and*. Certainly, if this pattern were restricted to *and*, one might well suppose (as does, e.g., Steedman [2012, 95], among others) that the syntactic behavior noted by Kehler is a mere idiosyncrasy of a single lexical item. However, as shown below, the patterns predicted by Kehler's account hold across all cases of syntactic coordination, regardless of which conjunct particle is chosen. We list some examples with the disjunction particle *or* and the "contrastive" conjunction particles *but*/*yet* in (543)–(546) (the slightly reduced acceptability for *yet* is possibly due to its somewhat stylized "learned" register requirements).

(543) a. John will go to Seattle this week or visit Chicago next week. (parallel)
b. ??*Which city will John go to Seattle this week or visit __ next week?

(544) a. Chris will have to be polite to the guests or face a stern talking-to. (nonparallel, cause-effect)
b. John and Mary are the people that Chris will have to be polite to __ or face a stern talking-to.

(545) a. ??*What did John have __ for lunch but ate soup for dinner? (parallel)
b. ??*Who did John vote for __ but Mary voted for Obama? (parallel)

(546) a. Aspirin . . . THAT's what I went to the store but forgot to buy __! (nonparallel)

b. What are we to make of the fact that it was this prescription medicine that John took __ but/(?)yet got sick anyway? (nonparallel)

c. The situation arose because there were a number of problems that John vaguely knew about __ but at that point was just wandering around in his usual foggy state. (nonparallel)

The contrast here between the parallel non-ATB extractions and the nonparallel ones is striking. Again, then, we find that non-ATB extraction becomes completely acceptable when the intended reading is based on a nonparallel semantic relation between the conjuncts and that this pattern is far from restricted to coordinations with *and*. Two nice examples from actual text are the following:

(547) [He] regards the limitless abundance of language as its most important property, one that any theory of language must account for __ or be discarded. (Campbell 1982, 183)

(548) Penitence abroad is little worth. There where we live lie the temptations we must defeat __, or perish. (Reade 1869)

We therefore have a solid basis for taking the CSC to reflect the interaction of the semantic effect of extraction as the establishment of a predication relation, with the pragmatic understanding of the basis for the relationship between the conjoined clauses. This is, again, not a configurational account but a *functional* account, based on how speakers make sense of sets of propositions rather than on phrase structure hierarchical relations. The content differs from the functional account Kluender gives for various island constraints, but what the two analyses have in common is that syntactic factors play a distinctly secondary role in these restrictions on the speaker's judgments.

Significant support for Kehler's reinterpretation of the CSC as a semantic/pragmatic effect comes from certain analogous patterns in Japanese and Korean relative clause and *wh* question constructions presented in Kubota and Lee (2015). What distinguishes the Japanese-Korean paradigm from the English case is that, on the one hand, the syntactic relationship between the two clauses is clearly some kind of subordination, and on the other, the gaps involved not only fail to correspond to *wh* extraction but are themselves optional. That is, there is a version of the *wh* question construction in which no gap at all appears; rather, the relationship is one of construal. Yet, just as in English, the ATB requirement correlates with the parallel coherence relation, and the suspension of that requirement depends on the involvement of other coherence relations instead. See Kubota and Lee (2015) for details (showing, inter alia, that any attempt to treat the Japanese-Korean facts via a covert movement analysis encounters insuperable empirical problems and will in addition render the CSC empirically vacuous).

Finally, we note that from time to time one encounters assertions that the counterexamples which are taken to show the nonexistence of the CSC as a syntactic effect are

suspect, because they are all restricted to cases of conjoined VPs rather than full-fledged clauses. The examples in (549), displaying acceptable non-ATB extraction from conjuncts, show that this claim is empirically unfounded.[9]

(549) a. [Which guy]$_i$ did they say [$_S$ [$_S$ they would hire __$_i$] and [$_S$ that would make everything all right]]?
b. The guy who they said they would hire and that would make everything all right was Harry.
c. How many barrels of toxic waste did he say/do you remember he said [they had dumped __ into the inlet and the Coast Guard had done nothing about it]?
d. Which people did Ann say that Nigel had fired __ and as a result everyone came to their defense?

Here, in each case, both the gapless clause and the clause containing the gap are under the scope of *say/said*.

The foregoing remarks all pertain to the Element Constraint, where the evidence seems extremely strong that non-syntactic factors interact to yield the ATB effects noted in the literature. But what of the Conjunct Constraint, which is far stricter, blocking *any* extraction of a complete conjunct at all? This phenomenon has occasionally been appealed to as an example of the inadequacy of functional approaches (sometimes labeled, somewhat misleadingly, as "reductionist treatments," as in, e.g., Boeckx [2012] and Ott [2014]) to islandhood; thus Boeckx (2012) cites the example in (550) as "evidence" that functional accounts cannot be the complete story of the CSC.

(550) *What does John like __ and oranges?

This assessment seems to be based on the unfounded premise that the Conjunct Constraint must have the same basis as the Element Constraint—a position for which there is neither a conceptual nor an empirical basis. It could just as well be that the Conjunct Constraint is a phonotactic phenomenon owing nothing to coherence conditions on discourse. The former possibility is in fact a plausible candidate for an explanation of the restriction illustrated in (550): Suppose, for example, that English requires all conjuncts to bear at least one nuclear stress. Such a requirement will immediately account not only for the Conjunct Constraint but the fact that conjuncts cannot consist of unstressed contracted forms:[10]

9. Such cases also rule out treatments of ATB counterexamples as conjunctions of a clause containing a filler-gap dependency with some other constituent type lacking such a gap.

10. This kind of prohibition is strongly reminiscent of Zwicky's (1986) discovery that the syntactic antipronominality of dative second objects in English, frequently invoked in the syntactic literature (see, e.g.,

(551) a. I talked to əm about safety equipment.

b. I want us and { them / *əm } to get along, understand?

We therefore have at least one promising candidate for a purely prosodic explanation of the Conjunct Constraint.

Still another interesting application of the line of analysis introduced by Kroch is the explanation provided in Culicover and Jackendoff (2005, 335–336) for the so-called *factive islands*, an effect alluded to (though not illustrated for extraction constructions) in Kiparsky and Kiparsky (1971), illustrated in (552).[11]

(552) a. What does Ginny think that Don bought __? >>>

b. What does Ginny regret that Don bought __?

Culicover and Jackendoff note Chomsky's "purely stipulative" syntactic treatment of this kind of difference as a possible manifestation of a syntactic difference between verbs that allow movement from S to S′. They further observe that there is an independent, purely semantic basis for the difference in well-formedness manifested in (552). Factive verbs such as *regret* presuppose the truth of the proposition expressed by their complements and therefore the existence of the event which is the clausal argument of *regret*. In the case of (552b), the use of this factive therefore creates the implicature that the event is in the common ground, so that what it was that Don bought is presumably known to both speaker and hearer. Hence, even if the hearer shares the speaker's knowledge that Ginny regrets Don's purchase of the item in question—which itself might not be the case—the fact that the speaker is querying something which by virtue of the use of the factive should be something s/he already knows the answer to (i.e., the object(s) that Don bought) makes the question distinctly anomalous.

The example in (552a) is quite different. Here, it isn't necessarily the case that Don actually bought something, only that Ginny believes that there was something Don bought, and it is perfectly possible for Ginny to hold that belief without the speaker knowing what she imagined that object to be. Hence the speaker's request for the

---

Cinque 1990; Postal 1998), is in fact explicable on the basis of what are familiar, independently needed *phonotactic* grounds; see the discussion below.

11. Szabolcsi and Zwarts (1993) offer a semantic account of factive islands as a subcase of their general theory of weak islands, where *wh* elements can only be extracted (and therefore scope wide) from contexts in which the denotation domain of the *wh* material is defined. As noted in Oshima (2007), their account has a number of analytic problems: it hinges critically, for example, on the claim that syntactic coordination of manner adverbials forms not intersections but "sums," i.e., unions, despite the fact that all semantic indications (such as entailment relations) point to the former, not the latter, analysis. They also introduce a type distinction between singleton sets and sets with larger memberships for no apparent reason other than to exclude *wh* extraction of manner predicates from contexts corresponding to non-iterable events, a point identified as a serious weakness by Schwarz and Simonenko (2018).

hearer to identify that object is pragmatically quite reasonable if there is reason for the speaker to believe the hearer to possess that knowledge.

This pragmatic account predicts that, as long as the right context is provided, acceptability of extraction out of factive islands will improve. Culicover and Jackendoff offer exactly this sort of evidence by embedding the question in a discourse context where a multiplicity of possible purchase objects is made explicit and where there is a strong implicature that the hearer has knowledge of Ginny's attitude toward these purchases that the speaker partly lacks:

(553) Yes, I KNOW that Ginny was happy about Don buying the laptop and external memory, but what I want to know is, what did she REGRET that Don bought?

Examples like (553) are, as Culicover and Jackendoff note, indeed much better than analogues along the lines of (552b).

On this account of "factive islands," the problem therefore once again arises from a question which is extremely odd given the network of background assumptions shared by the speaker and hearer. Culicover and Jackendoff show that not only island effects but a range of data exemplifying the definiteness effect, as well as certain cases of the CNPC, find natural accounts on the basis of this kind of pragmatic infelicity. A related proposal is given in Oshima (2007), which is further elaborated in Schwarz and Simonenko (2018), where the data set is somewhat expanded and refined from the Culicover and Jackendoff account; Oshima derives the decreased acceptability of manner adverbial *how* extraction from factive contexts, the absolute prohibition on *why* extraction, and the relative acceptability of spatial/temporal *wh* extraction from the basic scenario he sketches for the important minimal pair first cited in Szabolcsi and Zwarts (1993):

(554) a. (?)To whom do you know that John showed a letter?
b. *??From whom do you know that John got a letter?

On Oshima's (2007) and Schwarz and Simonenko's (2018) approach, the ill-formedness of (554b) is essentially due to inevitable pragmatic infelicity. In a nutshell, the felicity condition on *wh* questions requires that the question has at least one true answer. At the same time, the pragmatics of question-answer exchange requires that the answer to the question be informative. And in (554b), there is no way in which these two conditions can be satisfied at the same time.

The account goes as follows. (554b) involves a non-iterable event (for any given letter, there can be only one receiving event; by contrast, with (554a), the event is iterable since one can show the same letter to different people at different occasions) under the complement of a factive predicate. This means that, due to the factive presupposition, the speaker and the hearer both know about the relevant letter-receiving event. And whoever the sender of the letter, the hearer either knows or does not know the iden-

tity of the sender. The felicity condition that there be at least one true answer to the question, jointly with the factivity the question attributes to the hearer's knowledge, requires that the hearer does actually know the identity of the sender. Assuming that the speaker is posing the question (554b) on the (usual) assumption that its felicity condition is satisfied, the speaker should also already know that the hearer knows the identity of the sender. But then, the only one answer that the hearer can give for (554b) is not informative at all with respect to what is being asked, namely, the hearer's epistemic state (which is what the *wh* question of the form "who *do you know* . . .?" is all about). Thus, whatever the discourse context, (554b) inevitably leads to infelicity, and this is why this sentence is unacceptable.

There thus appears to be a range of plausible pragmatically based research on factive islands which derives such island effects from the interaction of felicity conditions on discourse and the lexical semantics of a certain class of verbs (which properly includes those imposing a factive presupposition). This result is of course important in itself. But its broader lesson is that, as in other island effects, functional factors can induce negative judgments of acceptability that—as in the case of (554b)—are as strong as what we would expect from examples that were syntactically ungrammatical. The factive island pattern thus illustrates the explanatory potential of the discourse-based line of investigation into islandhood phenomena inaugurated in Kroch's (1989) watershed paper.

10.1.3.3 Island effects via prosody: Kandybowicz (2006, 2009) The fate of the ECP in the twilight of the GB era represents a kind of case study of how long-held assumptions about the syntactic origins of some grammatical patterns can be drastically undermined by a few uncontroversial counterexamples and consequently abandoned in short order across much of the field. A 1993 *Linguistic Inquiry* squib by Peter Culicover (1993) sharply challenged more than a decade's reliance on the ECP to account for a wide variety of phenomena, based on its seeming indispensability in accounting for the *that*-t effect. The crucial counterexamples Culicover adduced are given in (555):

(555) a. Robin met the man {Op$_i$ that/who$_i$} Leslie said that *(for all intents and purposes) t_i was the mayor of the city.
b. This is the tree Op$_i$ that I said that *(just yesterday) t_i had resisted my shovel.
c. I asked what$_i$ Leslie said that *(in her opinion) t_i had made Robin give a book to Lee.
d. Leslie is the person who$_i$ I said that under no(rmal) circumstances t_i would run for president.

Such data severely jeopardize the claim that the *that*-t effect represents a failure of proper government with respect to the gap in subject position.[12] In order to preserve the ECP account of *that*-t ungrammaticality, one must assume that adverbials such as *for all intents and purposes* and *in her opinion* govern the gap in subject position. But the network of assumptions about phrase structure in the GB framework did not readily provide for a way to govern into a "canonical" position from an adjoined site. Culicover himself discusses, and rejects, various possibilities along these lines, concluding with the bleak assessment that "in view of the data discussed here it appears that the original Chomsky and Lasnik (1977) proposal for a *that*-t filter is empirically more adequate than a standard ECP-type account."

In the years following the publication of Culicover's squib, no convincing way around the difficulty based on plausible, independently motivated syntactic representations ever emerged. Inevitably, syntacticians began to suspect that the *that*-t effect is not the consequence of a dedicated syntactic principle or set of principles, but is rather a by-product of the interaction of a variety of factors, with syntax only one of the elements in the mix, and not necessarily the primary factor.

The work by Jason Kandybowicz (2006; 2009) is a good example. Kandybowicz points out a number of constructions, in addition to those noted by Culicover (1993) (and still earlier by Bresnan [1977]), which seem to at least ameliorate the severity of the effect, in certain cases quite markedly, and several other subtle aspects of the effect which point to a solution in terms of intonational properties in the relevant data. The gist of his analysis is that an empty syntactic category may not appear on the left edge of a syntactic constituent which corresponds to a phonological phrase and that, in the unacceptable cases such as (510b) (with the complementizer *that*), a complementizer falls on the wrong side of an intonational phrase boundary from a trace that immediately follows it. The effect of this prohibition is that a certain separation between *that* and the trace is necessary. This separation can be achieved in various ways, each of which contributes to the reduction of the *that*-t effect's severity. The intervening material in the kinds of data presented in Culicover's paper has the effect of moving the subject trace away from the boundary of the intonational phrase following *that* and

12. What is especially interesting is that data of just this sort were actually discovered a decade and a half earlier by Joan Bresnan, who mentioned them in a brief footnote in Bresnan (1977). At the time, the *that*-t effect had the status of little more than an isolated, idiosyncratic restriction, a "surface filter" as in Chomsky and Lasnik (1977). Since the effect neither followed from, nor had any deep significance for, any principles or premises taken to be fundamental at that time, Bresnan's exceptions had little theoretical resonance. By the time of Culicover's rediscovery of such examples, the status of the *that*-t phenomenon was of course quite different.

other complementizers, with the result that the syntactic and prosodic phrases are no longer in conflict.[13]

Kandybowicz backs up his argument with crosslinguistic support based on the Benue-Congo language Nupe. In Nupe, whose relative clauses behave in ways quite parallel to those in English, clausal complementizers also block extractions as in the ECP data displayed above, but there is a dedicated relative clause complementizer which can appear directly before a VP just as English *that* can in (556):

(556) John is the kind of person that looks for trouble and then seems surprised when he finds it.

Kandybowicz demonstrates that this apparent exemption from ECP effects is predicted on his account, because in Nupe, the relative clause complementizer is *part of* the relevant phonological phrase that corresponds to the syntactic clause from which the subject has been extracted. The subject gap in the relative clause is not on the left edge of that phrase; instead, the phonology of the relative complementizer begins that phrase, and the gap follows it. Hence there is no "missing" phonological material on the left edge of the crucial phonological phrase, and the latter is well-formed. If something similar were true in the case of English relative clauses, but not *that* complements, we would get the discrepancy between the two that is at issue.

One highly suggestive fact which supports such a parallel analysis for English is that both the relative pronoun and the finite declarative complementizer with an in situ subject are pronounced in normal-paced speech with a markedly reduced vowel—[ðə t] rather than [ðæt]—suggesting that these markers have cliticized to the following word, that is, have been integrated into the phonological phrase corresponding to the clauses they mark.

(557) a. Rocco's the guy [ðət] was here yesterday evening.
b. I think [ðət] Rocco needs to get his story straight.

But when speakers attempt to pronounce sentences in which a subject gap in a clausal complement follows *that*, the prosody of the complementizer is typically the full [ðæt] form:

13. It is very tempting to speculate that there is a linkage between this proposal on the one hand and, on the other, the general approach taken by Kluender and those pursuing his research paradigm in terms of processing bottlenecks at the edges of clauses which can be ameliorated in various ways, and to see the source of the *that*-t problem in the difficulty posed by the close linkage of the complementizer and the trace for recognition of the trace *as* a gap site. It seems likely that not only the kinds of factors noted by Kluender but various others as well, including matters of prosody, play a role in expediting or inhibiting the processing mechanism's efforts to connect the filler to a site within the structure where it can receive a coherent interpretation; work along the lines sketched here ultimately is likely to lead to a general theory of filler-gap processing in which morphological, semantic, pragmatic, and prosodic factors play significant roles.

(558) *Who do you think [ðæt] likes pizza?

The contrast indicates that the marker has failed to cliticize—that is, to attach itself, as a kind of prosodic dependent—to *likes*, and so remains outside the prosodic phrase—whose leftmost element will therefore be the phonologically null gap position, and hence ruled out.

A natural first reaction to the prosodic treatment of the *that*-t effect along the lines sketched above is that something more must be involved than just "sounding wrong" in the strong negative judgments speakers form in response to *that*-t examples. But there are precedents for just such a prosodic explanation masquerading as a syntactic effect. The following kinds of data had long been taken to reflect a syntactic fact about the "second object" position associated with verbs such as *tell*, *give*, and *show*, which take two NP complements. The second of these NPs cannot be a weak (i.e., necessarily unstressed) pronoun:

(559) a. I told John the facts.

b. *I told $\left\{ \begin{array}{c} \text{John} \\ \text{her} \end{array} \right\}$ them.

It has been assumed, for most of the history of modern syntax, that this pattern reflects a lexical property of such verbs ruling out pronouns from the second NP position (for example, Wasow [1975] cites the prohibition on weak definite pronouns in the dative object position as "well known"). A number of fairly intricate arguments have attempted to use this supposed fact about such verbs to construct claims about the nature of filler-gap dependencies. But Zwicky (1986) demonstrated that the effect in (559) had a far simpler basis. As it happens, a number of other phenomena display patterns similar to the "double object" construction exhibited in (559). What they all have in common are morpho-syntactic properties that prevent the unstressed pronoun from attaching itself phonologically to a preceding or following lexical item. They must, therefore, stand on their own as independent intonational units. Such units necessarily contain a major stress. Since the second NP in double object constructions has the status of an intonational phrase, it follows that in the normal course of things, weak definite pronouns—so called because they cannot receive normal intonational stress—will not appear in this second object position.

Zwicky's solution, which is completely consonant with standard views of intonational phrasing, eliminates what otherwise would be a very eccentric stipulative lexical restriction on a particular class of verbs. But the kind of negative judgment this prosodic violation induces is much more like what one expects from syntactic ungrammaticality. The so-called antipronominality of the dative second object position is a textbook illustration of the fact that a phonological source can have an effect indistinguishable from a syntactic failure.

10.1.3.4 Defenses of syntactic islandhood and their shortcomings The increasingly large body of evidence in favor of islandhood as an epiphenomenon of interactions among a wide range of independently attested functional factors has not gone uncontested. A range of results in experimental syntax has been offered which purportedly undermines the conclusions drawn by Kluender and other proponents of islandhood as functional in origin, as outlined in the preceding section. The major empirically driven objections to processing-based accounts of islands are given in two widely cited sources, Phillips (2006) and Sprouse et al. (2012). The thrusts of these two sources are in a sense complementary: whereas the conclusion that Phillips defends is that cognition itself, as embodied in the parser, consults the constraint pool defined by the grammar in the course of its operations, the point of Sprouse et al. (2012) is that perceptions of islandhood do not vary among subjects along the lines that processing-based accounts would lead us to expect. The findings reported in these studies seem to have been taken in some quarters as a definitive basis for maintaining the syntactic origin of islands (see, e.g., Boeckx 2012), but in fact, as more recent research has shown, there are crucial shortcomings in both of these works which make them ill-suited to serve as defenses of a syntactic view of islands. In this section we focus on what we take to be the major content of the first of these and provide our own assessment of how successful it is in rebutting the arguments for processing accounts of island effects.

Phillips (2006) reports the results of an experimental suite comprising two studies. The first of these is a kind of screening test, establishing the well-formedness of parasitic gaps and the ill-formedness of extractions from relative clauses within subject NPs, regardless of the tense of the clause. The second experiment, constituting the major contribution of the suite, uses a self-paced reading task to determine whether or not gaps are posited in the incremental parsing of extraction dependencies involving island domains, and whose methodological pivot is best explained by example. Consider the variants in (560):

(560) The school superintendent learned which high school students the proposal $\left\{\begin{array}{c}\text{a. to expand}\\ \text{b. that expanded}\end{array}\right\}$ drastically and innovatively upon the current curriculum would motivate during the following semester.

Expand is optionally transitive and therefore a target for filler-gap linkage in *wh* extraction, but the semantics of the verb preclude *which high school students* as the direct object. Under an incremental processing scenario, if the parser projects a gap following *expand(ed)* which it attempts to fill with the only available *wh* NP, the incoherence of the result will force a backtracking which in effect returns the latter to storage until a semantically compatible linkage is possible. This backtracking will be reflected in a slowdown in reading times. The absence of such a slowdown might therefore be taken as an indication that the parser does not posit a gap within the subject. Phillips

argues that on a strictly processing account of islandhood, the badness of unsupported filler-gap linkages into subjects entails that the parser should not be able to posit a gap within subjects under any circumstances, which would lead to the prediction that no matter what form a subject-internal "garden path" took, there should be no slowdown observed in the reading-time experiment. What actually happens is, however, quite different: while the finite version of the "spurious gap" context in (560b) indeed showed no slowdown in reading time following *expanded*, "[i]n the infinitival conditions, . . . average reading times . . . were 27 ms slower in the implausible conditions" (Phillips 2006, 812).

On the basis of these results, Phillips concluded that the parser projects gaps within subject-internal infinitival relative clauses but not their finite counterparts. This is, in Phillips's view, indirect but strong confirmation of the syntactic nature of the subject island constraint, since he saw the simplest explanation of the discrepancy as an architecture in which the parser is guided by the different statuses assigned by grammatical principles to extraction from finite and infinitive relative clauses. Thus, islandhood originates in the grammar and *determines* the behavior of the parser.

But a careful reading of Phillips's argument provides ample grounds for skepticism about his conclusion. The putative difference in acceptability between finite- and infinitive-clause-internal parasitic gaps is the indispensable empirical linchpin of his argument; if it turns out to be spurious, then obviously there would be nothing for there to be a "strictly syntactic" basis *for*. Yet at the same time that he states this difference in acceptability as an unequivocal fact, Phillips (2006, 803) acknowledges that there are clear counterexamples—the "classical examples of acceptable parasitic gaps discussed by Kayne (1983)," offering, however, no explanation for how these examples evade suppression by the grammatical filter guiding the parser.[14]

Such examples cut the ground out from under Phillips's line of reasoning; the fact that there is a subset of such examples which are uncontroversially well-formed makes it particularly difficult to maintain a syntactic constraint which suppresses their possibility. But there might still be a systematic *preference* for parasitic gaps out of infinitive contexts arising from semantic/pragmatic factors. A pointer to such factors is in fact given in Kluender (1998), as we have already noted.

Phillips's claims are, however, fundamentally undercut by the experimental results reported in Michel (2014) for *whether* islands, and most strongly and convincingly rebutted by the results reported in Chaves and Dery (2019), as discussed below. Michel (2014, 334) notes that his high-span subjects adjusted the probability of a gap down-

14. Phillips claims that "all of Kayne's examples . . . used relative clauses with a quantificational head NP," with further discussion of the relevance of this fact to his claims about finite clause subject extraction. In fact, it is not difficult to find examples, parallel to Kayne's in all critical respects, in which the head of the NP displays no quantification at all; see Chaves and Dery (2019) for examples.

ward when it occurred in a *whether* clause, but the experimental results still showed that an association between the filler and the gap site was posited:

> [T]he data reported in the ERP experiment indicate that both [low- and high-span] groups successfully identify the gap (P600 . . .) and associate the filler and gap (LAN . . .). . . . We simply see a difference in predictability: gaps are less predictable in *whether*-islands. Even so, when that less predictable gap is encountered, the parser still associates the filler with it, as evidenced by the post-gap LAN response in all four conditions. *Whether*-islands are unexpected, but not unprocessable.

Phillips is of course writing about subject-internal gaps and Michel about *whether* clause gaps. But on Phillips's account, a grammar-guided parser should be expected to fail to posit a gap in a *wh* island context such as a *whether* clause just as such a parse is expected to in the case of a subject-internal gap. And this prediction is strongly contraindicated by Michel's experiments: a gap site supposedly rendered unavailable to the parser because of the latter's guidance by the syntax has been shown, through extensive and careful experimental protocols, to be in fact available.[15]

The discrepancy between Phillips's earlier results and Michel's findings in his follow-up experiment naturally raises important questions about the source of this fairly dramatic divergence. Much of the answer to these questions can be inferred from the recent work reported in Chaves and Dery (2019), based on a suite of experiments probing Phillips's claims. Chaves and Dery found Phillips's conclusions wide of the mark on essentially every point, demonstrating that speakers do indeed posit and fill gaps within subject islands and that, when certain critical factors overlooked in Phillips's original experiments are taken into account in designing the stimulus examples, speakers find such data as acceptable as the uncontroversially grammatical control materials.

The difference between Phillips's finding and Chaves and Dery's basically comes down to the following factors:

- Phillips's test materials do not take into account the constructional conflict inherent in extractions from subjects. As noted repeatedly in the literature, subjects typically have the topic role of introducing old information, while extracted elements correspond to "information focus," the new or added information that is part of what is predicated *of* the topic.
- More than half of the tensed-clause embedded gaps in subject islands in Phillips's test materials were semantically anomalous in their *un*extracted variants.
- The constructional conflict is resolvable when the extracted material bears a plausible semantic/pragmatic relationship to the main predication.

15. This same point is reinforced by the findings reported in Chaves and Dery (2019), discussed below.

- The acceptability of subject island extractions further increases markedly with repeated exposure to such examples.

When the constructional clash noted is reduced to a minimum and the experimental subjects are exposed to several such examples, the discrepancy between normal, fully acceptable sentences and the subject island violations largely disappear. This result is documented not only for the simple extractions but for the supposedly proscribed subject parasitic gaps within tensed clauses, once the anomalous instances have been replaced by examples based on well-formed non-extraction sources.

Chaves and Dery's assessment of Phillips's study, and their own experimental results, are mirrored in the critique of the conclusions offered in Sprouse et al. about the putative support their investigation lends to the position that island effects in general reflect violations of structural constraints rather than functional factors.

The gist of Sprouse et al.'s (2012) claim is that the performance of speakers displaying a sizable range of working memory capacities show, according to certain metrics for working memory under laboratory conditions, a uniformity in processing activity with respect to island contexts which does not reflect the variation that should be expected on the basis of that range. This conclusion, they claim, runs counter to the reduction of island effects to restricted memory resources.

As noted in Michel (2014, 114), however, there are a number of problems with Sprouse et al.'s reasoning here. The most fundamental problem is that the key assumption underlying Sprouse et al.'s argument—which Michel calls the "cognitive covariation intuition" (CCVI)—is based on premises which Michel's experiments show to be empirically deficient. The CCVI is broadly stated as follows:

> [I]f island phenomena are due to working memory related processing costs . . . then individuals who have greater working memory capacities should be able to process the island violation sentences better and thus rate them as more acceptable. (Michel 2014, 105)

But as Hofmeister et al. (2013) and Michel (2014) argue, the choice of cognitive metrics that Sprouse et al. employ to test the working memory capacities of their subjects are not the right ones, testing only memory storage size and not (as in the Just and Carpenter model of working memory) active computation capacity.

Michel additionally raises a number of concerns about the statistical methods Sprouse et al. use to extract conclusions about the processing account of islands from their experimental results. Moreover, one foundational assumption they rely on—a so-called "push the limits" view that increasing processing difficulty will translate in a simple linear way into decreasing acceptability scores—is not supported by prior research on this specific correlation. In these and several other respects, Michel adduces strong evidence that the reasoning behind the Sprouse et al. (2012) conclusions was based on a number of dubious, overly simplistic premises.

In his own replication of Sprouse et al.'s experiments, Michel systematically corrects for these flaws in the experimental design: Michel's test suite uses the reading span test that Just and Carpenter themselves take to be the optimal probe for working memory, along with three others, including, crucially, the so-called memory lure diagnostic, which "provides a measure of how well participants can suppress distractors that compete with items in recent memory" (Michel 2014, 113). What emerges from Michel's much-improved test suite is the inadequacy of the Just and Carpenter model of working memory and its limitations due to storage cost. As Michel notes, his own results in themselves correspond to a trade-off between the constrained-capacity interpretations of working memory limitation on the one hand and the content-accessibility interpretation on the other, and is—along with the less-refined version of his Experiment 1 in Sprouse et al. (2012)—essentially neutral on the relationship between processing and combinatorics so far as island effects are concerned. He further notes that his own experiments do point to a relevant cognitive factor that may well be significantly implicated in island effects, having to do not with storage, as in the Just and Carpenter model, but with *processing* capacity—in particular, the ability of speakers with a higher ability to filter out distractors in real-time syntactic processing to parse long-distance dependencies. One of Michel's key points is that the relationship between processing ease and positive acceptability judgments is still too complex and poorly understood to conclude a simple linear correlation between the two.

The state of the art in the experimental syntax of islands thus, in our view, supports two conclusions which are entirely congenial to the functional view of islandhood and its origins.

- There is no evidence supporting the claim that speaker variation data in responding to island violations favor a syntactic origin for the latter. The reported facts, rather, suggest that a more fine-grained and probably more complex model of cognitive limits on processing capacity is called for.
- There is positive evidence that speakers both hypothesize and fill gaps in syntactic contexts in accordance with their probabilistically based expectations, that islandhood strongly corresponds with those expectations, and that revising them radically ameliorates islandhood effects.[16]

16. There is, conversely, also some reason to believe that subjects' expectations can be revised in the opposite direction, leading to perceptions of ill-formedness in constructions which other evidence suggests should be regarded as grammatical. Phillips's (2006) use of defective sources for testing subject island extractions from finite contexts is a good example; Chaves and Dery (2019, 30) observe that "because more than half of the items in the tensed condition were odd . . . it is possible that comprehenders adapted to the near-systematic oddness of the sentences in the tensed condition and simply refrained from attempting to fill gaps in that condition." Yet their results show that when the anomalous tokens are eliminated from the experimental suite, the status of the same class of examples approaches the same level of acceptability as the nonfinite relatives.

10.1.4 The Status of Syntactic Islandhood: Summing Up

The foregoing discussion has touched on only a small portion of the research on the non-syntactic sources of island effects. Nonetheless, we believe that the evidence presented to this point puts a strong burden of proof on advocates of an approach to island effects which seeks a syntactic explanation for these effects, along the lines of research in the grammatical tradition following Ross (1967) or various adaptations thereof to non-movement frameworks (e.g., Gazdar 1981; Gazdar et al. 1985; Steedman 2012; Morrill 2017). As time goes on, work along the lines described in the sources cited above, including Deane (1991), Hofmeister and Sag (2010), Gibson (2018), Culicover and Winkler (2019), and many others, has revealed a steadily widening network of independently attested extragrammatical effects sufficient to account for the classical island phenomena and other constraints that have been proposed over the past half century since Ross's thesis appeared. The comprehensive critical overview of research on islandhood during this period, Chaves and Putnam (2020), argues that upon close examination, the evidence that has accumulated against a unitary syntactically determined source of islandhood is now overwhelming.

This state of affairs has come to be recognized even by supporters of mainstream generative grammar, as attested in the following remarks in Newmeyer (2016, 207):

> The explanation of island phenomena has been a central feature of formal grammatical theory practically since its inception. However, a growing number of linguists have provided explanations for these phenomena that are not based on purely syntactic constraints. Some linguists have proposed alternative explanations that appeal to information structure or to semantic information. Others find the basis for island constraints to lie in processing: In brief, sentences that violate island constraints are difficult to parse. In the course of the last few decades, more and more formal syntacticians have concluded that an exclusively syntactic approach to islands is overly ambitious, but there is broad uncertainty about how to construct a general theory of island phenomena, or even whether a general theory is a possibility.

Against this increasingly skeptical background, we examine in the following section specific islandhood-based arguments putatively supporting the transformational analyses of coordination and ellipsis phenomena we have considered in previous chapters. We demonstrate that, consonant with the empirically problematic status of syntactic islandhood, these argument simply fail to hold up.

10.2 Islands, Coordination, and Ellipsis

The history of coordination and ellipsis analyses in mainstream generative grammar has seen frequent appeals to island effects, a fact which is hardly surprising in view of the nature of the constructions involved. In ellipsis, for example, the principal issues center on the existence of structure which has no phonological expression and is therefore

only indirectly detectable. On the assumption that islands are essentially syntactic in nature, it is certainly logically possible that the effects of the deletion operation would leave any ill-formedness due to island violations unaffected, but matters are a bit more complex if, as the preceding discussion suggests, syntactic form plays only a partial or secondary role in restricting extraction possibilities. In particular, it becomes more likely that island effects in coordination and ellipsis in the literature either are spurious or can be shown to arise on the basis of independent factors not taken into account in the original sources.

In the remainder of this chapter, we revisit the movement-based analyses of coordination and ellipsis considered in previous chapters and demonstrate that the evidence based on island effects given in support of these proposals does not, in fact, stand up to closer examination. The importance of this discussion in the present context is that once the islandhood-based arguments are removed, the movement-based analyses lose perhaps the strongest basis for their validity, leaving them distinctly unfavorable in relation to the much simpler analyses that eschew covert structural representations we have already given (specific problems of these analyses that don't pertain to the issue of islandhood have already been discussed in detail in previous chapters).

10.2.1 Gapping

We begin with the island-based argument in the analysis of Gapping supposedly motivating the low VP coordination analysis. Advocates of the low VP coordination approach in one or another version have attempted to construct such an argument by invoking alleged parallelisms between syntactic island effects in extraction and in Gapping.

The unacceptability of the following examples from Johnson (2004), for example, is offered in an effort to establish the adherence of Gapping to island constraints, and hence the plausibility of the movement-based analysis (see also Toosarvandani [2013, 18] for a similar argument).

(561) a. *John wondered what to cook today and Peter ~~wondered what to cook~~ tomorrow. (*Wh* Island)
b. *John must be a fool to have married Jane, and Bill ~~must be a fool to have married~~ Martha. (Adjunct Island)
c. *I read out the order to fix the tortillas, and Mary ~~read out the order to fix~~ beans. (Complex NP Constraint)
d. *Stories about Frankenstein terrified John, and ~~stories about~~ Dracula ~~terrified~~ Peter. (Subject Island)

However, counterexamples are easy to find. Examples such as the following, some of which are from Culicover and Jackendoff (2005), are structurally parallel to those in (561), but they seem well within the bounds of acceptability.

(562) a. [Wife of a couple discussing who decides what to cook for which meal:] Ok, how about this: I get to decide what to cook for LUNCH, and you, for DINner. (*Wh* Island)

b. ROBIN believes that everyone pays attention to you when you speak FRENCH, and LESLIE, GERMAN. (Adjunct Island [Culicover and Jackendoff 2005, 273])

c. ROBIN knows a lot of good reasons why DOGS are good pets, and LESLIE, CATS. (Complex NP Constraint [Culicover and Jackendoff 2005, 273])

d. I don't think we need worry about John harassing us. Threats directed at ME would offend his *wife*, and at YOU, everyone else! (Subject Island)

Examples like these contraindicate the supposed parallel between Gapping possibilities and syntactic extractability (see also Repp [2009, 13] for a similar conclusion).

A note is in order here regarding speaker variability. Some readers may find the examples in (562) not fully acceptable. While we have not done any systematic survey, we have consulted seven speakers on the status of (562a) and (562d), and the latter was unanimously found impeccable, while the former showed some variability in acceptability. In relation to Gapping in particular, one possible reason for speaker variability is prosody. In all of the examples in (562), strong contrastive/parallel stress is required on the corresponding elements in the antecedent and gapped clauses, along with distinct but quite short pauses before the second of these elements. For example, in (562a), *lunch* and *dinner* should be contrastively stressed, with the former receiving high/falling pitch, and the latter, following a clipped pause between *you* and *for dinner*, pronounced with steady mid-level pitch.

Having addressed the speaker variability issue (which may already provide part of the answer to the following question), the next question is where the difference lies between the (allegedly) unacceptable examples in (561) and the structurally parallel good examples in (562). We have already provided extensive evidence to support the position that processing, coherence, and other functionally based principles are the source of island effects in the typical extraction contexts. It seems highly likely that similar factors are at play in the variability exhibited by specific examples with respect to acceptability, though we have to leave a detailed examination of this effect in the case of Gapping for future study.

Note that there are also cases of acceptable extraction with the corresponding Gapping examples ill-formed:

(563) a. There were certain cars of which only [the windows __] were damaged in the explosion. (slightly modified from Ross [1967, 242])

b. ??The windows of the van were cracked, and of the cars, shattered.

(563a), discussed in Ross (1967) as evidence *against* a general restriction on extraction from subjects, is perfectly good. But the structurally parallel Gapping example in (563b) sounds severely degraded. This kind of example strongly suggests that the real-time interpretation of Gapping involves mechanisms that are somewhat different from those involved in filler-gap dependencies.

10.2.2 "Respective" Predication

In chapter 5, we provided a detailed analysis linking *respectively* interpretations, symmetrical predicates, and summative predicates to the effect of a single **resp** operator which, when applied to two multisets of cardinality *n* and a "continuation," relates the members of the two multisets in a pairwise manner with respect to the continuation, preserving the original multiset structure (so as to enable recursive applications of **resp**). This analysis involves a covert-movement-like operation, formalized via the vertical slash in Hybrid TLCG. In the literature, the interdependency between non-adjacent expressions in terms of semantic interpretation found in "respective" predication and related phenomena has sometimes been analyzed in terms of a similar nonlocal movement (or movement-like) operation (as in Barker's [2007] analysis of *same* via "parasitic scope"; cf. section 5.4.1.2) and sometimes via a chain of local operations (as in Gawron and Kehler's [2004] analysis of "respective" readings; cf. section 5.4.1.1). One might then wonder whether any empirical argument can be made for one or the other type of analysis.

There are some discussions in the literature relevant to this question, though the investigations in this domain are considerably less thorough than in other, well-studied domains such as extraction. Thus, Carlson (1987) claimed that the internal readings of *same/different* obey the same syntactic island constraints as filler-gap dependencies, but this observation was disputed by Dowty (1985) and Moltmann (1992). We concur with these latter authors. Note in particular the following cases, all of which seem to induce the internal reading for the symmetrical predicates relatively easily:

(564) a. Robin and Leslie believe that the acquisition of $\left\{ \begin{array}{l} \text{the same} \\ \text{different} \end{array} \right\}$ skill sets is crucial to success in the business world.
(Subject Island violation on the *de re* reading)

b. [Robin and Leslie usually agree about who's a bad singer at a karaoke party, and they both get immediately mad when such a person starts singing.]
But today, something weird happened: Robin and Leslie got mad when different people started singing. (Adjunct Island violation)

c. The Smiths$_1$ and the Joneses$_2$ go to the same$_1$ psychiatrist and to different$_2$ psychiatrists respectively.
(Coordinate Structure Constraint violation; Dowty 1985, 6)

Similar facts seem to hold for the "respective" reading (though so far as we can tell, examples such as (565) have not been discussed in connection with island effects in the previous literature):

(565) a. Robin and Leslie thought that studying category theory and intuitionistic logic respectively would be all that was needed for success.
(Subject Island violation)

b. Robin and Leslie got home before the train and the bus stopped running respectively. (Adjunct Island violation)

c. Robin and Leslie named someone who was innocent and guilty respectively.
(Complex NP Constraint violation)

It is somewhat difficult to tell what conclusion to draw from data such as those in (564)–(565). One thing we can say is that if one takes island effects to be by-products of functional factors such as constraints on real-time processing and felicity conditions on discourse, what the above data suggest is that the processing and discourse factors which come into play in filler-gap dependencies are not the same ones which govern the interpretation of "respective" and symmetrical predicates. In the former case, island violation *does* result in reduced acceptability (but can be ameliorated via various strategies); in the latter case, island violation simply does not seem to arise to begin with, or, at least, the effects seem considerably weaker in general.[17] In any event, like the case of Gapping discussed above, a more careful investigation is needed to determine what is really going on in these data.

10.2.3 Pseudogapping

Miller (2014, 82–83) notes a variety of attested examples in which pseudogapping displays insensitivity to island restrictions (note that (566b) is a case of antecedent-contained deletion (ACD); as discussed in chapter 6, we take pseudogapping and ACD to be licensed by the same mechanism).[18]

(566) a. According to current ideas, the frothiness of space **retards the arrival of** a burst's highest-energy photons more than it does ~~retard the arrival of~~ *the lowest-energy photons*. [Subjacency]

b. **Bring the same kind of carry-ons** when traveling by train as you would ~~bring __ when traveling~~ *by air*; you're allowed two per person.
[Adjunct Island]

17. Note in this connection the robustness of semantic islands, which presents yet another pattern, discussed in section 10.3.

18. Miller (2014) labels (566a) as Complex NP, but Subjacency seems more appropriate.

In order to derive these examples via movement + ellipsis, the movement operation prior to ellipsis would have to evacuate the remnant by moving it across an island.

It should be noted in this connection that, although islandhood is often invoked to argue for a movement-based analysis of ellipsis (a representative and very clear instance of this can be found in Kennedy [2003]; see chapter 8 and the discussion in the next subsection), the relationship between ellipsis and islandhood is actually much more intricate. In fact, since Ross's (1969) celebrated observation that sluicing is completely insensitive to islandhood, the literature has struggled to account for the apparent dilemma that some cases of ellipsis are (apparently) island-sensitive while others are not. One typical (and influential) explanation offered in the mainstream literature for this dilemma, a version of which can be found in Merchant (2008a), involves the "size" of the structure that undergoes ellipsis (the key idea is that as long as what's elided is "large enough," the offending island-inducing structure gets elided via the ellipsis process and thus the resultant string is well-formed).

Thus, we might conclude that, after all, pseudogapping's insensitivity to islandhood does not have any direct implication for the viability of a movement-based analysis. This is true to some extent at a general level. But there are two important points which are worth keeping in mind. First, even though island insensitivity is not a dead end for a movement-based approach, it significantly weakens the motivation for a movement-based analysis by removing a key piece of evidence for assuming covert syntactic structure. Second, the classification of island-sensitive and island-insensitive types of ellipsis should be done consistently inside movement-based approaches, and in this respect, Miller's (2014) data seem highly problematic. This is because Merchant's (2008a) classification identifies VP ellipsis as a type of ellipsis that does *not* admit island repair. Since pseudogapping is usually taken to be an instance of VP ellipsis in most mainstream approaches, this apparently poses a serious tension in the kind of structure-based account of island repair of the sort proposed by Merchant.[19]

10.2.4 Extraction Out of Ellipsis and Islandhood

The issue of islandhood is relevant in two distinct cases in Kennedy's (2003) argument for covert structural representations underlying ellipsis. We review these two cases and conclude that in neither case does the argument go through.

19. One might think that if the contrast between VP ellipsis and pseudogapping with respect to island repair is real, the empirical problem itself remains unaccounted for in non-movement approaches as well. This is true, but note that non-movement approaches (including ours) start from the premise that non-syntactic factors play major roles. In the case of VP ellipsis and pseudogapping, seeking explanation in independently observed differences in the discourse structural properties of the two constructions (cf., e.g., Hoeksema 2006) seems to be a promising direction.

10.2.4.1 *Wh* extraction As noted above, Kennedy's argument for covert structural representations in VP ellipsis is a version of the familiar type of logic in the transformational literature invoking the island sensitivity of structure that is invisible on the surface string. The argument is crucially based on the observation that examples of the following sort are unacceptable:

(567) a. *Sterling criticized every decision that Doug was upset because Lou did.
b. *Dogs, I understand, but cats, I don't know a single person who does.

According to Kennedy, this fact follows straightforwardly from the assumption that these examples are derived in the following manner, where the fronted *wh* phrase (or the relativization operator) is extracted from an island environment which at a later stage of derivation gets deleted via VP ellipsis. On this view, the ungrammaticality follows from the violation of island constraints in the *wh* movement involved:

(568) a. *Sterling criticized every decision that Doug was upset [$_S$ because Lou did ~~criticize *t*~~]. (Adjunct Island)
b. *Dogs, I understand, but cats, I don't know [$_{NP}$ a single person [$_S$ who does ~~understand *t*~~]]. (Complex NP)

However, this argument is problematic since there is an alternative explanation of the unacceptability of these examples that does not make any reference to the notion of islandhood. To see this point, note that examples such as the following, in which VP ellipsis takes place in an environment in which the target clause is embedded, are at least as unacceptable as (567a):

(569)*STERLING criticized every decision (that) Doug was unhappy (that) [$_S$ LOU did ~~criticize *t*~~].

Note that in (569), unlike the examples in (567), the extraction site is not located within an island. Thus, the relevant factor here is not islandhood, but simply the level of embedding of the elided material. In fact, theorists from widely divergent approaches have noted that such embedding corresponds to semantic/pragmatic difficulties which are in themselves sufficient to guarantee that such examples are perceived as anomalous (see, e.g., Jacobson [2016] on short-answer ellipsis and Toosarvandani [2016] on Gapping). In fact, Kennedy himself had, in prior work on the identity conditions holding between antecedent and ellipsed VP (Kennedy 1994/2008), argued that examples such as (569) are ill-formed because they violate a complex set of essentially semantic requirements in which islandhood considerations have no role at all. Thus, Kennedy's own account of ACD in Kennedy (1994/2008), by unifying the patterns displayed in (567) and (569), decisively undercuts his own invocation of islandhood—and the concealed structure which it would entail—in the analysis of VP ellipsis in Kennedy (2003).

Moreover, there seems to be considerable speaker variation in the judgments on at least some of the relevant examples. For example, the judgment given to (567b) by Kennedy (2003) was disputed by our informants, according to whom the example is only somewhat awkward at worst. In fact, an even more natural example that is structurally parallel to (567b) can be constructed without too much difficulty:

(570) Life, I like to think about, but death, I don't and I don't know $\left\{ \begin{array}{c} \text{ANYONE} \\ \text{a single person} \end{array} \right\}$ who DOES.

This strongly suggests that here, too, syntactic islandhood *per se* is orthogonal to the source of any negative judgments reported in (567b). For speakers who accept (570), the semantic parallelism between the two clauses after *but* is apparently all that is needed to make the gap in the elided VP transparent to linkage with the fronted topic. While the mechanisms involved in this effect are still far from clear, the fact that both clauses present a strong negation contrast to the first conjunct is surely a crucial factor. The pattern just described, then, can be seen as the reflection of a priming effect which speaks to a processing problem rather than a grammaticality issue *per se*.

We therefore have reason to reject the claim that islandhood is relevant for the analysis of the "extraction out of VP ellipsis" type of data (and relatedly, pseudogapping; cf. section 10.2). Rather, it appears to be some currently unknown set of factors related to the complexity of the *antecedent* that bears most critically on the possibilities for instantiating the free variable in the ellipsis operator introduced in chapter 6. This aspect of the syntax-semantics interface has not, so far as we are aware, been the focus of psycholinguistic investigation, although recent work by Kim and Runner (2018) has begun to explore the interaction of generalized conditions on discourse felicity with what they argue is a specific "structural" restriction on VP ellipsis.[20] We conjecture that a full understanding of the patterns in ellipsis is crucially dependent on future work targeting the kind of phenomena noted in this section. In particular, it is possible that the sorts of complexity noted by Kluender and others, summarized earlier, is implicated in the psycholinguistics of anaphora in ellipsis, though the specific ways in which such factors play a role are likely to differ in different types of phenomena. We leave further investigation of this issue for future work.

20. Kim and Runner's conclusion—that there is a specific syntactic identity condition on VP ellipsis whose violation degrades output regardless of any further deficits due to functional infelicities—is completely consonant with the analysis we provide above, at least in the rather broad terms in which it is framed. Their use of the description "structural" in this context does not, so far as we can tell, necessarily implicate facts about branching tree structure, and could instead be interpreted as restrictions on the syntactic type of the antecedent, as per the particular restriction in (330).

10.2.4.2 Attributive comparatives Kennedy (2003), following Kennedy and Merchant (2000), takes the degraded status of examples like (571) to reflect the Left Branch Condition, formalized as a syntactic condition that is sensitive to the form of *structural* representations in the prosodic domain:

(571) John buys more expensive wine than he buys beer.

As noted in section 8.3.5, we disagree with Kennedy and Merchant on the assessment of the empirical status of data like (571). Specifically, contrary to Kennedy and Merchant (and much of the literature following it, such as Bacskai-Atkari [2014] and LaCara [2016]), we take examples such as (571) to be syntactically well-formed. While such examples are not impeccable, we have observed that they are not altogether ruled out, and much better examples can be found. For example, a substantial number of our informants report that examples such as (572) (which is structurally parallel to (571)) are simply not ill-formed at all:

(572) John writes better novels than he/Mary writes plays.

More generally, we have noted considerable speaker variability in judgments on attributive adjective subdeletion without pseudogapping. It is true that on the whole, the pseudogapped versions are more readily acceptable, but the non-elliptical versions are noticeably improved by, for example, using higher-frequency lexica (thus replacing *writes* with *composes* materially degrades (572)).

In place of Kennedy and Merchant's syntactic account, we suggest that a much simpler mechanism is at work here: a garden path effect created in large part by the fact that in the non-elliptical version of the sentence, the *than* clause is already parsable *without* the "comparative deletion" to the left of the direct object. Unlike syntactic ill-formedness, which gives rise to either-or judgments, garden path effects typically result in a much more complex set of speaker responses which often depend on non-structural factors, exactly of the sort found in the responses we have received from our informants.

Our view here is essentially in line with the "expectation-based" processing model proposed in Chaves (2013), which assumes that the parser's performance is at least in part guided by application of the rules and constraints of the grammar.[21] To articulate this non-syntactic alternative account a bit further, in languages such as English that do not allow for the option of left branch extraction for overt material, parses in which material corresponding to an NP left branch must be "imported" from elsewhere in the comparative structure will be initially overlooked by the parser, based on the lack of

21. Similar models of human sentence processing have been proposed in Phillips (2006) and Wagers and Phillips (2009) as well, although the specific claims about the role of grammatical constraints in the processing model that these latter works advocate have been challenged by Michel (2014).

such syntactic-semantic linkages in the simpler (and more basic) cases involving overt filler-gap linkages. Instead, the parser will compose what appear to be self-evidently well-formed sentences such as *he buys beer* in (571). But once such a parse is obtained for the *than* clause, it fails to be composable with the antecedent clause in the comparative, and this will at best force the kind of real-time backtracking that registers as diminished acceptability or even outright ill-formedness (see, e.g., Du and Yu [2012] for a computational model of processing breakdown under backtracking in garden path sentences, and Chaves [2013] for an account linking such effects to extraction possibilities).[22]

But this requirement—that the parser follow the line of least resistance, so to speak—is counteracted in exactly the class of cases in which such comparatives are unproblematic, namely, cases involving ellipsis in the *than* clauses such as (573):

(573) John buys more expensive wine than he does beer.

As already noted, what makes these examples different from their non-elliptical counterparts is that they involve an additional process of ellipsis, and there is reason to believe that this additional component has a nontrivial consequence for the real-time processing of the sentence. Recall from the discussion in chapter 8 that in pseudo-gapping clauses, obtaining a coherent interpretation critically depends on instantiating the free variable *P* introduced by the ellipsis operator based on discourse context. The real-time recovery of an interpretation from the antecedent clause entailed by this anaphora resolution requirement will, by its very nature, force the processor to inspect the antecedent and ellipsis clauses simultaneously, as it were, upon encountering the "transitive" auxiliary in cases such as (573). This interruption keeps the otherwise automatic default parsing routine—and the garden path effect it gives rise to—from being completed. Moreover, this ellipsis resolution process reinforces the Parallelism relation between the main clause and the *than* clause, thereby facilitating the construal of the whole sentence as a comparative construction. The extra anaphora resolution step

22. One might wonder why the same garden path effect does not arise in simpler comparative sentences such as the following, whose *than* clause is similarly parsable as a stand-alone sentence:

(i) Mary runs faster than John runs.

While a complete account is beyond the scope of the present work, the difference between examples like (i) and cases of attributive subdeletion such as (571) seems to follow from a relatively simple syntactic difference in the position of the gap site: in (i) (unlike (571)), at the end of parsing the *than* clause, the modifier which must be interpreted into the parse in order to yield a semantically coherent interpretation corresponds to a position in which the parser expects to find such a modifying term (as in *runs fast*). Moreover, the sentence-final position is where such a modifier is routinely interpreted in overt extraction cases, as in, e.g., *How fast did John run __?* Based on such extraction possibilities elsewhere in the language, the parser need do nothing more than identify the position following *run* as a site into which a modifier introduced earlier in the sentence can be interpreted.

triggered by ellipsis therefore circumvents the error-plus-backtracking sequence which the default parsing strategy leads to.

There is a corollary to this processing-based explanation of the improved status of pseudogapped variants of attributive comparatives: the same amelioration effect should be observed in any other possible form of the *than* clause in which the default parse routine fails to produce a well-formed stand-alone interpretation. This prediction is indeed corroborated by data involving the cognate object construction. Note first that the cognate object construction is usually not well-formed without a modifier:

(574) *John will probably die a death as a result of all this.

But as part of an attributive comparative, this specific construction is well-formed (here, small caps indicate a lower level of contrastive stress and full caps a maximum degree of contrastive stress):

(575) I feel like I'm SLEEPING a much more painful SLEEP these days than I'll eventually DIE a DEATH at the end of my life.

The default processing routine does not yield a well-formed result on its own, and it seems reasonable to suppose that the parser, in its attempt to make sense of this otherwise uninterpretable cognate object expression, is able to make early use of the contrastive parallelism supplied by the comparative construction (which can be further facilitated by appropriate intonational cues). The amelioration effect observed in (575) provides strong support for the current processing-based view over Kennedy and Merchant's syntactic account: on the latter type of approach, since (575) does not involve any deletion operation, it is predicted to be just as unacceptable as (571), and its improved acceptability remains a total mystery.[23]

---

23. Another case that shows the importance of considering plausible alternatives to syntactic accounts of speakers' judgments comes from Kennedy and Merchant's (2000) discussion of examples such as (i).

(i) John wrote a successful novel, and Bill, a play.

Kennedy and Merchant argue that this sentence has two readings, corresponding to (iia) and (iib):

(ii) a. John wrote a successful novel. Bill wrote a play.
b. John wrote a successful novel. Bill wrote a successful play.

According to them, (iib) is a distinct reading corresponding to an LF (different from the one for (iia)) that is derived by a deletion operation (similar to the one involved in their analysis of attributive comparatives) that deletes not just the verb but also a prenominal adjective. However, the assumption that Gapping is derived by deletion directly contradicts the standard assumption about the analysis of Gapping in the Principle and Parameters approach: Johnson (2000, 2009) and Vicente (2010) have argued extensively that Gapping should *not* be analyzed via deletion. We believe that a much simpler account of the apparent ambiguity of (i) is more plausible. In our view, the two readings of (i) paraphrased in (ii) reflect vagueness (or underspecified interpretation) rather than true ambiguity. Specifically, (iia) is the only reading that the grammar licenses for (i) (which is immediately available in any standard analysis of Gapping). The sentence is of course compatible with a situation in which the speaker intends to convey (iib), especially in the right kind of

Seen in this light, the cross-linguistic evidence that Kennedy and Merchant adduce in support of their association between left branch extractability and non-elliptical attributive adjective subdeletion need have no syntactic implications at all. The observed cross-linguistic correlation would follow at least equally well from the garden path alternative account we have proposed above, in tandem with a parser guided in its search possibilities by whichever constructions are syntactically available in a language. Specifically, a parser taking into account the grammatical option of a left branch extraction can be counted on to include this possibility in how it handles the processing of the *than* clause in the comparative, rather than automatically running a default processing routine that creates a garden path effect. In other words, the parser's expectation includes the possibility of covert modification of nominal heads in attributive comparative constructions just in case the language allows for the option of overt left branch extraction (as in Polish and Czech). Thus, since the same effect is predicted just as readily on the current processing-based alternative, this additional "evidence" does not distinguish between Kennedy and Merchant's syntactic account and the present processing-based account.

10.2.5 Right-Node Raising and Islandhood

One point that deserves special attention in connection with island effects is the immunity of the shared material in RNR constructions to island effects. Examples such as (576) are fully representative of the island-exempt status of RNR:

(576) John had the conviction that Bill was the collector who had bought __, and Mary held to the fervent belief that Ann was the dealer who had sold __, the last orange Lamborghini ever sold in North America.

The apparent gap in both conjuncts is located within a relative clause that is itself part of a definite complex NP. Note that a leftward extraction example involving (mostly) the same lexica sound distinctly less natural:

(577) ??Which car did John have the conviction that Bill was the collector who had bought __?

The island insensitivity of RNR has long been recognized as theoretically problematic for mainstream approaches to islandhood which take RNR to be a case of rightward movement (see Wexler and Culicover [1980] for an early discussion of the troublesome issues raised by a movement analysis of RNR for syntactic theory).

---

context (for example, in a discussion of literary successes of one's friends), but the additional meaning that is felt to be present in such contexts is simply inferred as a conversational implicature (Relevance), presumably facilitated by the constructional property of Gapping that induces a strong Parallelism discourse relation on the two conjuncts.

One obvious line of response to the problem posed by (576)–(577) would be rejection of the movement analysis of RNR, and this line has in fact been widely taken in the literature. But in the case of approaches such as ours, matters are a bit more complex. Not only is movement (in the sense of a structure-changing operation applying at the level of configurational syntactic representations) not available in the first place in our framework, but our analysis of both extraction and RNR is based on the same "mechanism"—hypothetical reasoning—in the two families of construction (albeit involving different types of connectives). If islandhood is taken to be a syntactic effect, one might attempt to argue that this identity commits us to the position that both classes of phenomena should display a certain ill-formedness when the apparently missing material—regardless of its relative position to the "gap site"—is associated with an ostensible gap in an island context.

Such a challenge would, however, be seriously misguided. One of our core assumptions, backed by the considerable and increasing body of evidence for the essentially non-syntactic nature of island effects of the kind surveyed earlier, is that islandhood is the by-product of complex interactions among functional factors. In accord with this stance, we hold that there is indeed a crucial asymmetry between extraction and RNR, involving not the position of the "filler" material simpliciter, but rather whether the speaker/hearer encounters the gap or the filler first. Based on the arguments, evidence, and results adduced above from the work of Kluender, Michel, and Chaves, we conclude that the difference between (576) and (577) is exactly what we would expect on the assumption that islandhood is often attributable to difficulties for the parser in the course of real-time processing.

In the case of leftward extraction, the speaker is confronted with material which cannot be interpreted with respect to its immediate local context; it must be "read into" a gap which the parser expects somewhere in the upcoming material, but whose location it has no clues to. All material between the filler and the intended gap site must therefore be inspected, entailing the investment of considerable cognitive effort, particularly when distractors such as referentially "heavy" nominals are encountered along the way. Under these conditions, the parser's task is extremely challenging, and we expect it to be easily derailed as the number of subsidiary processing tasks to be gotten out of the way before the gap can be securely identified increases past a certain point. But in RNR constructions, the parser encounters the gap first. There is no difficulty identifying the context into which any missing material subsequently encountered must be supplied; that context is immediately evident at the point where some argument sought by a selecting functor type is absent. The parser is now looking for a string corresponding to some type X which has not locally combined with some other functor seeking an X argument. All that it need do is process incoming material until it encounters such an argument, in which case its search ends immediately and a complete interpretation can

be composed. The problems that the parser faces in the two respective construction types are therefore of quite different orders, with RNR representing by far the simpler processing task.

Thus, we take it that the puzzle of island insensitivity of RNR in the classical literature is only an artificial puzzle, one arising out of a certain set of assumptions. These assumptions may have seemed plausible back in the 1980s, but given the more nuanced and sophisticated view of island effects in the current literature, they are difficult to maintain. Under the more plausible set of assumptions about the nature of island constraints, the insensitivity of RNR to island effects is in fact consistent with what kinds of effects arise in what kinds of syntactic constructions. While the discussion here is rather sketchy, we take this result to further corroborate the extragrammatical view on island constraints as a whole.

10.3 A Note on Semantic Islands

Our discussion up to this point has been restricted to syntactic islands, but it has been recognized for a long time that there are islands for scope-taking as well, known as *scope islands*. It is sometimes alleged that universal quantifiers are restricted to scope within the closest finite clauses containing them, but this is not true; there are sentences such as (578):

(578) John believes that everyone in this room is trying to get him fired.

There is a reading for (578) in which for it to be true and felicitous, John does not necessarily need to know that any of the people in question is actually in the room. Nonetheless, genuine semantic islandhood effects do seem to exist. Typical examples are given in (579).

(579) a. Fido has a bone that is in every corner of the house. (Rodman 1976)
b. If every woman in this room gave birth to John, then he has a nice mother. (Winter 2001)

(579a) has only the strange reading according to which there is a single bone possessed by Fido that is in every corner of the house (compare: *Fido has a bone in every corner of the house*), suggesting that the universal quantifier inside a relative clause cannot scope out of the relative clause. Similarly, if the quantifier were able to scope out of the *if* clause in (579b), the sentence should have a sensible interpretation 'for every woman in the room, it is the case that if she gave birth to John, then John has a nice mother.' However, (579b) has only the nonsensical reading entertaining the possibility that John had multiple mothers giving birth to him.

One point that should be stressed before we speculate on any possible deeper sources for the anomaly in (579) is that the traditional argument which takes these data to pro-

vide evidence for the representational treatment of scope via syntactic movement at LF (see, e.g., Winter [2001, 81–84] and Ruys and Winter [2010] for relatively recent versions of this type of argument) loses much of its force given that, as abundantly documented above, so far as *syntactic* island constraints go, it is not difficult to find acceptable examples of even supposedly strong islands. Consider, for example, (580), where an acceptable violation of the Subjacency condition is contrasted with a covert-movement counterpart, for which a similar "island violation" reading seems much more difficult to obtain:

(580) a. Euthanasia is [$_{NP}$ a topic [$_{S}$ which I can never find [$_{NP}$ anyone [$_{S}$ I can argue with __ about __]]]].

b. #John is [$_{NP}$ a reporter [$_{S}$ who wrote about [$_{NP}$ a prosecutor [$_{S}$ that claimed that everyone in this room is guilty of perjury]]]].

The intended reading for (580b) is that for every person now in the room with the speaker, some prosecutor(s) claimed that that person is guilty of perjury and John wrote a story about each such prosecutor. But this reading in unavailable. The only accessible interpretation is that John wrote a story or stories about a prosecutor who claimed that the group of people currently in the room were all guilty of perjury.

Given the robustness of (at least some of) the semantic island effects, it may appear that the only way to capture them properly would be to encode them directly within the grammar (which is possible to do in our approach, for example, by building on the work by Pogodalla and Pompigne [2012] briefly discussed in chapter 7). However, we feel that at this point such a move may still be a bit too hasty. After all, syntactic and semantic processing pertain to different components of grammar and deal with somewhat different types of abstract representations of linguistic knowledge. Thus, it would be hardly surprising if it turned out that the same type of dependency/configuration presents a much more robust processing difficulty in one domain than the other. There is some pioneering work investigating the real-time semantic processing of scope islands (see, for example, Wurmbrand 2018) in the recent literature. Hopefully, this new line of work will delineate more clearly the proper division of labor between the combinatoric and processing components regarding the treatment of the semantic island effect, along lines similar to that done for syntactic island effects in the line of work reviewed in the present chapter. We therefore regard the source of apparent scope islands as an entirely open question whose answer is likely to emerge only well in the future.