This PDF includes a chapter from the following book:

Type-Logical Syntax

© 2020 Massachusetts Institute of Technology

License Terms:

Made available under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International Public License
https://creativecommons.org/licenses/by-nc-nd/4.0/

OA Funding Provided By:

The open access edition of this book was made possible by generous funding from Arcadia—a charitable fund of Lisbet Rausing and Peter Baldwin.

The title-level DOI for this work is:

doi:10.7551/mitpress/11866.001.0001

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

8 Extraction/Ellipsis Interactions

Building on the analysis of ellipsis in chapter 6 and the analysis of extraction in chapter 7, this chapter addresses what we take to be the central question posed by ellipsis phenomena: Do the latter receive an adequate explanation only if we posit covert syntactic structure at the ellipsis site? The great bulk of the literature on ellipsis has always assumed a positive answer to this question, if only implicitly, and the authors of much of the more recent work have taken pains to offer arguments defending that view (e.g., Kennedy and Merchant 2000; Johnson 2001; Kennedy 2003; Merchant 2015; and Thoms 2016).

Kennedy (2003) in particular offers a clear, concise summary of five of the seemingly strongest such arguments focusing on VP ellipsis:

1. The distribution of ellipsis remnants reflects sensitivity to island conditions on configurations which are not visible in the surface string at the ellipsis site.
2. The interpretations available in ellipsis data exhibit Strong Crossover effects that imply the existence of syntactic gaps within the ellipsis site.
3. The anaphora possibilities available in the interpretations of VP ellipsis reflect Binding Condition B effects, implying the existence of pronouns within the ellipsis site.
4. Parasitic gap licensing behavior requires the presence of a syntactic gap within the ellipsed material in order to license a visible gap in an island context within the remnant material.
5. The "attributive comparative" construction permits certain possibilities which appear to violate the "Left Branch Constraint" just in case ellipsis is also involved, a pattern that can be accounted for as an instance of repair, via deletion, of an offending covert structure.

We evaluate these five arguments, concluding that each of them is either empirically deficient or predicated on an undermotivated treatment of the relevant data. In the latter case, there are alternative analyses which are at least as successful in accounting for the facts and require no reference at all to configurational properties of the ellipsed

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

material. Our conclusion is that the various data sets offered in Kennedy (2003) in support of covert structural analyses for ellipsis phenomena do not in fact motivate such analyses.

8.1 Kennedy's Arguments for Covert Structure in Ellipsis

Each of Kennedy's five arguments is based on a syntactic pattern which is either commonly held, or explicitly argued, to require appeal to a specific syntactic configuration for a satisfactory explanation. In all but the last case, the ellipsed clause displays parallel behavior to its unellipsed counterpart, and since that behavior is supposedly motivated by specific structural facts, it would follow that the observed parallelism is eo ipso a sufficient basis to posit covert structure. We begin by outlining the storyline of each of Kennedy's arguments, returning in section 8.3 for a reassessment of them.

8.1.1 Island Effects

Probably the most frequently encountered argument for covert structure in ellipsis is the claim that ellipsis possibilities mirror the (un)acceptability of their unellipsed counterparts with respect to island effects.[1] Kennedy (2003) offers an instance of this argument for VP ellipsis in English based on the paradigm in (363) and (364):

(363) a. Sterling criticized every decision that Lou $\left\{\begin{matrix}\text{did} \\ \text{criticized } t\end{matrix}\right\}$.

b. *Sterling criticized every decision that Doug was upset because Lou $\left\{\begin{matrix}\text{did} \\ \text{criticized } t\end{matrix}\right\}$.

(364) Max refused to buy the shirt that I picked out even though it was less expensive than the one that (*the salesperson complimented him after) he $\left\{\begin{matrix}\text{did} \\ \text{picked out } t\end{matrix}\right\}$.

Extraction in the "full" version of (363a) is impeccable, and so is the ellipsed version. When the material destined for ellipsis appears in an adjunct island, however, the extraction is bad, as is the ellipsed variant, as in (363b). Since the basis for islandhood in Kennedy's framework is assumed to be structural, the parallel behavior of full and ellipsed clauses with respect to islandhood is taken as a strong argument that ellipsis

1. Such arguments, however, typically offer little detailed consideration of the well-formedness of certain species of ellipsis in which island violations do not appear to incur any penalty. This problem for the islandhood argument was in fact already noted in Ross's (1969) watershed paper on sluicing, in which Ross explicitly acknowledges that he has no account to offer for it. One relatively recent strategy for dealing with this problem, pursued, for example, in considerable detail in Barros et al. (2014), is to deny that the covert syntax of the ellipsed material involves any islandhood in the first place. Jacobson (2016), however, offers a persuasive rebuttal to this line of solution. At present it seems fair to say that there is no consensus among researchers advocating covert structure in ellipsis about these difficulties.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

must apply to a structural object which either complies with island conditions and is acceptable or fails to comply and is judged ill-formed. Likewise, we see that in (364), both the full and VP ellipsis versions are fine when the gap is within an extraction-accessible complement but are bad when the ellipsed VP hosting the gap is within an island.

8.1.2 Strong Crossover

The second major argument for covert structure in Kennedy (2003) hinges on the Strong Crossover (SCO) Condition discussed at length in Postal (1971, 2004). SCO blocks the appearance of a pronoun on a filler-gap pathway which is coreferential with, and in some sense structurally superior to, the gap. Thus, (365) does not allow the interpretation 'for which male person is it the case that that person is always criticizing himself?':

(365) *Who$_i$ is he$_i$ always criticizing t_i ?

Unlike islandhood, the kind of gradient effects noted in recent processing-oriented approaches such as Kluender (1998), Hofmeister et al. (2013), and Chaves (2013) have never been reported in connection with SCO judgments. This absolute character seems to give appeal to SCO considerable weight as evidence for covert structure in data such as (366):

(366) Who$_i$ will Mary vote for t_i if he$_{j/*i}$ does ~~vote for t_i~~?

It is not possible to construe the referent of *he* as the candidate in question.[2] But when the pronoun is clearly not to be construed as coreferential with the gap following *for* (i.e., $j \neq i$), the sentence is acceptable. This pattern follows more or less immediately by assuming that there is covert syntactic structure featuring a movement trace within the ellipsis site. By contrast, (according to Kennedy) it is difficult to see how a purely interpretive approach could account for the fact that the sentence is either acceptable or unacceptable depending on the interpretation of the pronoun *he*. The point is that purely interpretive approaches do not posit any sort of trace or gap in the ellipsis site. But without such a gap somewhere in the representation, there is no straightforward way to induce the SCO effect in (366).

---

2. If we were to replace the trace in (366) with a pronoun, the result in (i) will still be bad due to Condition B effects (see section 8.1.3):

(i) Who$_i$ will Mary vote for if he$_{j/*i}$ does ~~vote for him$_i$~~?

Thus, the only alternative is for the object of *for* to be a real gap, which leads to the ill-formedness of (366), on the assumption that a trace is present at the gap site.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

8.1.3 Ellipsis and Binding Condition B

Kennedy's argument from Binding Condition B hinges on data such as (367):

(367) John$_k$ takes care of him$_i$ because he$_{j/*i}$ won't { a. take care of him$_i$ / b. ~~take care of him$_i$~~ }.

Condition B, first stated in Chomsky (1981), blocks coreference between an NP and a pronoun within a certain locality domain. The generalization itself is taken to be relevant in the more recent literature, though there does not seem to be a clear, broadly accepted formulation of this condition in the Minimalist framework. Condition B effects are evident in the unellipsed form, (367a): when the subject of the adjunct clause is interpreted as coreferential with the direct object, the example is bad. On the assumption that (367b) is derived by deleting a VP that has the same syntactic form as its unellipsed counterpart, we predict that the status of the example depends entirely on the adjunct subject pronoun's index. In contrast, on a purely semantic treatment of ellipsis, there is supposedly no reason why a predicate of the form $\lambda y.\mathbf{take\text{-}care\text{-}of}(y)(\mathbf{k})$ cannot take **k** as an argument (with **k** the denotation of the pronoun *he*); after all, the semantics of *He takes care of himself* requires exactly this type of interpretation. Hence, the argument goes, to capture the ill-formedness of (367b) on the coreferential interpretation, there must be an actual pronoun in the representation for the sentence at the level at which Condition B applies.[3]

8.1.4 Ellipsis and Parasitic Gaps

Parasitic gap licensing is standardly taken to be a matter of syntax. (368) illustrates the basis for the common claim that subjects cannot host gaps unless at least one gap also appears in the main VP.

(368) a. *Who$_i$ did close friends of __$_i$ become famous?
b. Who$_i$ do close friends of __$_i$ always defend __$_i$?

The examples in (369) exhibit a slightly more complex pattern. Standard islandhood tests seem to confirm that the gap in the adjunct in (369) is derived by *wh* extraction. As (369a) shows, when no gap appears in the adjunct's own VP, the gap in its clausal subject yields an ill-formed result, as would be expected on the basis of (368).

(369) a. *Otis is a person who$_i$ I admire __$_i$ because close friends of __$_i$ became famous.

3. Kennedy's reasoning here depends on an implicit assumption which is open to question, viz., that the relevant locality condition accounting for Condition B effects is to be formulated in terms of syntactic configurations. Jacobson (2007) has in fact thoroughly argued that this assumption is deeply problematic, showing that such approaches require quite elaborate and indirect mechanisms to overcome a number of serious empirical mispredictions.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

b. Otis is a person who$_i$ I admire because close friends of __$_i$ admire __$_i$/*him$_i$.

In contrast, in (369b), a gap in the adjunct subject position is licensed since it is legally parasitized on the VP internal gap. Here there are two separate extraction chains, with a presumed empty operator binding both the licensing gap following *admire* and the parasitic subject-internal gap. But this derivation is only legal on condition that the gap within the adjunct VP is indeed a trace. If, instead, a null pronoun were the object of *admire*, the results would be unacceptable—something we know because, if the pronoun is overt, as per (369b), the result is bad. Since, however, (369b) (in its gap version) is good, we can be sure that what is missing in this example is a true gap, licensing the parasitic gap in the subject position. From this conclusion, Kennedy infers that in the example (370), a true gap must somehow be located within the ellipsis site:

(370) Otis is a person who$_i$ I admire __$_i$ because close friends of __$_i$ seem to ∅ .
(Engdahl 1985; Kennedy 2003)

That is, ∅ must be ~~[$_{VP}$ admire __$_i$]~~ entailing a covert structure analysis for such examples.

8.1.5 The Argument from Comparatives

There is one more argument that Kennedy (2003) adduces on behalf of covert structure—in effect, a very compressed synopsis of the argument based on properties of comparative (sub)deletion phenomena presented in Kennedy and Merchant (2000). The argument runs as follows: ordinary comparatives, of the sort illustrated in (371a), are arguably best analyzed as instances of *wh* movement involving a null operator, along the lines of (371b):

(371) a. Pico's novel was much more interesting than Brio thought it would be __.
b. Pico's novel was much more interesting than [O_{deg_i} Brio thought it would be t_i]

The movement account of comparative "deletion" is supported by the fact that it predicts the ill-formedness of examples such as (372a), presumably with the structure (372b):

(372) a. *John buys more expensive wine than he buys beer.
b. John buys more expensive wine than [O_{deg_i} he buys [$_{NP}$ t_i beer]]

In (372), the gap is in the "left branch" position (Ross 1967), which is known to disallow extraction:

(373) *Whose did you borrow [$_{NP}$ __ book]?

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

But this seemingly straightforward account of (372a) is at odds with the fact, apparently not noticed prior to Kennedy and Merchant (2000), that pseudogapped analogues of (372a) are impeccable:

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

Without argument, Kennedy and Merchant take the discrepancy between (372a) and (374) to be syntactic in nature. They offer a complex account which in its essence derives the ill-formedness of (372a) from a lexical condition on the highest functional head position (FocP, in their account) in the constituent from which the null degree operator in left-branch position is extracted. This condition makes lexical insertion of an actual head unavailable, with the result that the Spec of this focal projection bears a certain unchecked feature which cannot be interpreted at LF, leading the derivation to crash. If, however, the constituent [$_{\mathrm{DetP}}$ t_i beer] escapes by rightward movement to surface as a pseudogapping remnant, the uninterpretable feature is "trapped" within FocP and deleted along with the rest of the VP, as shown in (375):

(375) . . . than he does [$_{\mathrm{VP}}$ ~~[$_{\mathrm{VP}}$ buy [$_{\mathrm{FocP}}$ O_i t_j]]~~ [$_{\mathrm{DetP}_j}$ t_i beer]]

Thus, on Kennedy and Merchant's account, the contrast between (372a) and (374) is the direct result of the structural deletion operation responsible for ellipsis—a powerful argument, in their view, for treating ellipsis as a reduction of syntactic structures.

Furthermore, a prediction follows directly from their analysis which they argue is confirmed cross-linguistically, namely, that in languages in which left-branch extraction is permitted—and in which, therefore, the lexical condition Kennedy and Merchant posit for English does not hold—we should expect to find attributive comparatives not only in ellipsed forms of the construction but in their unellipsed counterpart as well. They cite Polish and Czech, in which left-branch extraction is legal, in support of this prediction: in both languages, examples comparable to (372a) are altogether unproblematic. In Greek and Bulgarian, on the other hand, which mirror the prohibition in English against left-branch extraction, the judgment patterns in the attributive comparative data parallel the patterns found in English, as indeed predicted on their account.

8.1.6 Summary

The preceding arguments appear to implicate unavoidably the presence of structure at the ellipsis site in VP ellipsis—but, as we argue below, the logic of Kennedy's (and Merchant's) argument hinges crucially on two key assumptions. First, it takes as a given the structural basis of the various diagnostics that are invoked to probe for concealed configuration at VP ellipsis. Second, it assumes that the extracted material in cases such as (376) is linked to a position *within* the missing material following the auxiliary rather than to a complement position directly associated with the auxiliary itself:

(376) I know what John ate for lunch, but I don't know what Bill did.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

In the discussion below, we argue that there is no compelling basis for either of these assumptions. We argue further that there is an alternative, processing-based explanation for the attributive adjective comparative pattern in which the data, including the cross-linguistic facts, have a markedly simpler explanation without reference at all to particular structural conditions.

8.2 Extraction/Ellipsis Interaction in Hybrid TLCG

In this section, we extend the analysis of ellipsis in chapter 6 to the basic cases of ellipsis/extraction interaction and ellipsis in comparatives.

8.2.1 "Extraction from Elided VP" Revisited

Examples like the following seem to have been taken as prima facie evidence for the presence of covert structure, even without consideration of specific structure-dependent effects along the lines Kennedy pursues.

(376) I know what John ate for lunch, but I don't know what Bill did.

For example, referring to data such as (376), Johnson (2001) remarks, "In these cases too the ellipsis site seems to have internal parts." (Elbourne 2008, 216) echoes this assessment, referring the reader to Johnson's article

> for a summary of the controversy about whether theories without normal syntactic structures in the ellipsis sites can deal with examples like these. The upshot is not encouraging, and things seem especially difficult for [the version in Hardt (1999)], according to which there is nothing whatsoever in ellipsis sites.

As we show below, however, our own proposal, which posits no syntactic material at all in the supposed ellipsis site, is vulnerable to none of the objections Johnson and Elbourne raise.[4] Our own analysis is immune to this criticism since it analyzes apparent extraction from an ellipsis site as a genuine syntactic extraction—but not extraction

---

4. A certain caution is necessary in evaluating these objections, however. Elbourne apparently regards the problem with Hardt's analysis, vis-à-vis the extraction data, as the fact that since the latter hinges on a purely semantic recovery process, it ipso facto cannot include syntactic information, such as the presence of a syntactic gap site. But this criticism assumes that Hardt's approach still incorporates a movement-based source for *wh* fillers (or some analogous mechanism for syntactically registering the connectivity relationship). But nothing in Hardt's paper requires such a source; for example, if *wh*-fillers are licensed in place, and the linkage to their gap sites is treated as a matter of interpretation, then data such as (376) do not present a problem for a direct interpretation approach.

The real difficulty that Hardt's approach faces is precisely the kind of structure-dependent patterns which Kennedy adduces in his paper. A purely interpretive account of apparent extraction from VP ellipsis contexts runs into trouble not, *pace* Johnson and Elbourne, because of the extraction itself (which might indeed be only apparent as per the scenario just sketched); rather, the serious challenges arise when the extraction clause displays behavior which in its unellipsed counterparts appears to be strictly syntactic. It is precisely facts such as the parallelism between ellipsed and unellipsed clauses with respect to, for example, Condition

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

from an ellipsed position. Rather, we propose that apparent extraction from an ellipsed VP is in fact extraction from one or another argument of the "transitive" auxiliary which is associated with the general ellipsis operator introduced in chapter 6. That is, examples such as (376) involve not just a semantic object, as in Hardt's analysis, but an actual syntactic extraction from an ordinary overt VP, as we show below. Hence, these constructions are predicted to conform to whatever conditions hold on extraction in general without any concomitant assumption of covert structure corresponding to an "ellipsed" VP.

To flesh out the analysis outlined above, consider first a VP from which material is missing in some unknown, possibly non-peripheral position. Such a VP—which is just what is captured by the description VP↾XP—corresponds to a VP in a filler-gap relationship with an XP filler. Thus, in *I wonder what John said to Mary*, the subconstituent *said __ to Mary* constitutes a VP with a medial NP gap, meeting the description VP↾NP. It follows, then, that in principle, a sentence such as (376) can be licensed by mapping the auxiliary, not to VP but to a VP looking for an NP missing from *somewhere* "inside" it. Such a predicate is in the simplest case a transitive verb which can take an object argument (cf. the analysis of pseudogapping in chapter 6) but could in fact represent any VP from which an NP is missing, as in (377).

(377) John was someone whom I had [$_{\mathrm{VP}\upharpoonright\mathrm{NP}}$ heard [stories about __] for a long time].

Our generalized ellipsis operator in (299) from chapter 6 won't quite be sufficient here, because it is stated in terms of categories of the form VP/\$, which targets categories seeking arguments only to the right and thus excludes VPs with medial gaps. To correct this omission, we amend the definition of the ellipsis operator to give it still greater generality.

We start by demonstrating that a Geach-style proof is available for auxiliaries which will map them to types of the form $(\mathrm{VP}\upharpoonright\mathrm{XP})\upharpoonright(\mathrm{VP}\upharpoonright\mathrm{XP})$:

(378)

$$\dfrac{\dfrac{\dfrac{k;\ \mathcal{O};\ \mathrm{VP/VP} \qquad \dfrac{[\sigma_1; f; \mathrm{VP}\upharpoonright\mathrm{NP}_{\#wh}]^1 \quad [\varphi_2; x; \mathrm{NP}_{\#wh}]^2}{\sigma_1(\varphi_2);\ f(x);\ \mathrm{VP}}}{k\circ\sigma_1(\varphi_2);\ \mathcal{O}(f(x));\ \mathrm{VP}}}{\lambda\varphi_2.k\circ\sigma_1(\varphi_2);\ \lambda x.\mathcal{O}(f(x));\ \mathrm{VP}\upharpoonright\mathrm{NP}_{\#wh}}{\upharpoonright}\mathrm{I}^2}{\lambda\sigma_1\lambda\varphi_2.k\circ\sigma_1(\varphi_2);\ \lambda f\lambda x.\mathcal{O}(f(x));\ (\mathrm{VP}\upharpoonright\mathrm{NP}_{\#wh})\upharpoonright(\mathrm{VP}\upharpoonright\mathrm{NP}_{\#wh})}{\upharpoonright}\mathrm{I}^1$$

B, which Hardt's analysis cannot handle easily because, given the existence of reflexives, a purely semantic approach such as his cannot rule out interpretations in which a subject and object are coreferential.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

The generalized ellipsis operator then takes such “Geached” auxiliaries and maps them to type $\mathrm{VP}{\upharpoonright}\mathrm{XP}$, anaphorically supplying the meaning of the gapped VP:[5]

(379) $\lambda\rho\lambda\varphi_1.\rho(\lambda\varphi_0.\varphi_0)(\varphi_1);\ \lambda\mathscr{F}.\mathscr{F}(R);\ (\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}){\upharpoonright}((\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}){\upharpoonright}(\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}))$
——where R is the semantic term of a sign retrieved from the context whose type is $\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}$

The analysis of (the antecedent clause of) (380) then goes as in (381).

(380) I know what John ate for lunch, but I don’t know what$_i$ Bill did ~~eat~~ __$_i$ ~~for lunch~~.

(381)

$$\dfrac{\dfrac{\dfrac{\dfrac{\text{①}\rightarrow\dfrac{\begin{array}{c}[\varphi_1;x;\mathrm{NP}_{\#wh}]^1\\ \vdots\\ \mathsf{ate}\circ\varphi_1\circ\mathsf{for}\circ\mathsf{lunch};\\ \mathbf{ate}(x)(\mathbf{lunch});\mathrm{VP}\end{array}}{\begin{array}{c}\lambda\varphi_1.\mathsf{ate}\circ\varphi_1\circ\mathsf{for}\circ\mathsf{lunch};\\ \boxed{\lambda x.\mathbf{ate}(x)(\mathbf{lunch})};\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}\end{array}}{\upharpoonright}\mathrm{I}^1 \quad \begin{bmatrix}\varphi_2;\\ u;\\ \mathrm{NP}_{+wh}\end{bmatrix}^2}{\mathsf{ate}\circ\varphi_2\circ\mathsf{for}\circ\mathsf{lunch};\ \mathbf{ate}(u)(\mathbf{lunch});\ \mathrm{VP}} \quad \begin{array}{l}\mathsf{john};\\ \mathbf{j};\\ \mathrm{NP}_{-wh}\end{array}}{\mathsf{john}\circ\mathsf{ate}\circ\varphi_2\circ\mathsf{for}\circ\mathsf{lunch};\ \mathbf{ate}(u)(\mathbf{lunch})(\mathbf{j});\ \mathrm{S}}{\upharpoonright}\mathrm{I}^2}{\lambda\varphi_2.\mathsf{john}\circ\mathsf{ate}\circ\varphi_2\circ\mathsf{for}\circ\mathsf{lunch};\ \lambda u.\mathbf{ate}(u)(\mathbf{lunch})(\mathbf{j});\ \mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}} \quad \begin{array}{l}\lambda\sigma.\mathsf{what}\circ\sigma(\boldsymbol{\varepsilon});\\ \lambda P.\mathbf{what}(P);\\ \mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh})\end{array}}{\mathsf{what}\circ\mathsf{john}\circ\mathsf{ate}\circ\boldsymbol{\varepsilon}\circ\mathsf{for}\circ\mathsf{lunch};\ \mathbf{what}(\lambda u.\mathbf{ate}(u)(\mathbf{lunch})(\mathbf{j}));\ \mathrm{Q}}$$

The derivation of the ellipsis clause involves free instantiation of the $\mathrm{VP}{\upharpoonright}\mathrm{NP}$ variable introduced in the generalized ellipsis operator defined above. We take R to be the grayed-in semantic term obtained in the proof line ①. The first part of the proof for *what Bill did* then takes the following form:

(382)

$$\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{l}\quad\vdots\\ \lambda\sigma\lambda\varphi.\mathsf{did}\circ\sigma(\varphi);\\ \lambda f\lambda x\lambda y.f(x)(y);\\ (\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}){\upharpoonright}(\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh})\end{array} \quad \begin{array}{l}\lambda\rho\lambda\varphi.\rho(\lambda\varphi_0.\varphi_0)(\varphi);\\ \lambda\mathscr{F}.\mathscr{F}(\lambda x.\mathbf{ate}(x)(\mathbf{lunch}));\\ (\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}){\upharpoonright}\\ \quad((\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}){\upharpoonright}(\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}))\end{array}}{\lambda\varphi.\mathsf{did}\circ\varphi;\ \lambda x\lambda y.\mathbf{ate}(x)(\mathbf{lunch})(y);\ \mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}} \quad \begin{bmatrix}\varphi_3;\\ v;\\ \mathrm{NP}_{+wh}\end{bmatrix}^3}{\mathsf{did}\circ\varphi_3;\ \lambda y.\mathbf{ate}(v)(\mathbf{lunch})(y);\ \mathrm{VP}} \quad \begin{array}{l}\mathsf{bill};\\ \mathbf{b};\\ \mathrm{NP}_{-wh}\end{array}}{\mathsf{bill}\circ\mathsf{did}\circ\varphi_3;\ \mathbf{ate}(v)(\mathbf{lunch})(\mathbf{b});\ \mathrm{S}}{\upharpoonright}\mathrm{I}^3}{\lambda\varphi_3.\mathsf{bill}\circ\mathsf{did}\circ\varphi_3;\ \lambda v.\mathbf{ate}(v)(\mathbf{lunch})(\mathbf{b});\ \mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}}$$

5. One might wonder whether this operator needs to be further generalized via a vertical-slash version of $. We will not address this question here.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

The term obtained at the last step of this proof, supplied as an argument to the extraction operator, yields an interpretation identical to the unellipsed embedded question *what bill ate for lunch*.

The above (re)analysis of "extraction out of an elided VP" as extraction of a pseudogapping remnant gives us, in effect, a proof-of-concept argument for rejecting the assumption that covert structures in VP ellipsis necessarily exist in order that a "site of origin" exist for filler-gap linkages that appear to implicate material missing from deleted VPs. While various versions of this approach have been challenged in previous work, we show in section 8.4 that the empirical basis for these objections is quite fragile and problematic when a larger range of relevant data is taken into account.

8.2.2 Ellipsis and Comparatives

The other technical piece that is needed is an analysis of comparatives. For this purpose, we outline here our assumptions about the syntax and semantics of comparatives, implementing in Hybrid TLCG the basic semantic characterization of comparative constructions in standard accounts (see, e.g., Kamp 1975; Cresswell 1976; Hellan 1981; von Stechow 1984; Hendriks 1995a; Kennedy 2005). As we show below, hypothetical reasoning with the vertical slash tied to prosodic λ-abstraction enables a flexible and simple treatment of the complex syntax-semantics interface of the comparative construction, which straightforwardly interacts with an independently motivated analysis of VP ellipsis to yield the right results in cases in which these two phenomena interact. This interaction between comparatives and ellipsis also plays a crucial role in the non-deletion analysis of the "attributive comparative" data we formulate in the next section.

8.2.2.1 The basic syntax-semantics interface of comparatives The semantics of comparatives reflects an (in)equality in the degree of some property or predicate. Thus, in the following examples, what is asserted is a proposition of the form $d_1 > d_2$, where d_1 and d_2 are, in the case of (383a), for example, the degrees to which Mary and Ann are tall respectively.

(383) a. Mary is taller than Ann is.
b. Mary runs faster than John runs.

Other forms of the comparative equate two degrees (e.g., *John runs as fast as Mary runs*) or quantify the difference (e.g., *Mary runs twice as fast as John runs*).

One way to compositionally derive the comparative meaning from the surface form of the sentence in examples like those in (383) is to assume that both the main clause and the *than* clause denote predicates of degree of type $d \to t$. For example, (383a)

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

is derived (either via ellipsis of the adjective *tall* in the *than* clause or by some other means that derives the same semantic effect) from the following source,[6]

(384) (?)Mary is taller than Ann is tall.

with an "LF" that looks like the following:

(385) er-than [λd. Mary is d-tall] [λd. Ann is d-tall]

That is, at the level relevant for semantic interpretation, the comparative operator *er-than* takes two degree descriptions of type $d \to t$ and compares the maximum degrees that satisfy each of these degree descriptions. Here, we assume (following many previous authors) that gradable adjectives take a degree argument (which is often implicit). In Hybrid TLCG, this can be modeled by assuming that adjectives are of type $d \to e \to t$ semantically and Adj↾Deg syntactically.

Putting aside the exact morpho-syntax of comparative forms, the degree operator can be defined as follows:

(386) $\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{er}) \circ \mathsf{than} \circ \sigma_2(\boldsymbol{\varepsilon});$
$\lambda P\lambda Q.\mathbf{max}(P) > \mathbf{max}(Q); \mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{Deg}){\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{Deg})$
where $\mathbf{max} =_{def} \lambda P.\iota d.P(d) \wedge \neg\exists d'[P(d') \wedge d' > d]$

The derivation for (385) then goes as follows:

(387)

$$\dfrac{\lambda\varphi.\mathsf{tall}\circ\varphi;\ \mathbf{tall};\ \mathrm{Adj}{\upharpoonright}\mathrm{Deg} \qquad [\varphi;\ d;\ \mathrm{Deg}]^1}{\mathsf{tall}\circ\varphi;\ \mathbf{tall}(d);\ \mathrm{Adj}}$$

$$\dfrac{\mathsf{is};\ \lambda P.P;\ \mathrm{VP/Adj} \qquad \mathsf{tall}\circ\varphi;\ \mathbf{tall}(d);\ \mathrm{Adj}}{\mathsf{is}\circ\mathsf{tall}\circ\varphi;\ \mathbf{tall}(d);\ \mathrm{VP}}$$

$$\dfrac{\mathsf{mary};\ \mathbf{m};\ \mathrm{NP} \qquad \mathsf{is}\circ\mathsf{tall}\circ\varphi;\ \mathbf{tall}(d);\ \mathrm{VP}}{\mathsf{mary}\circ\mathsf{is}\circ\mathsf{tall}\circ\varphi;\ \mathbf{tall}(d)(\mathbf{m});\ \mathrm{S}}$$

$$\dfrac{\mathsf{mary}\circ\mathsf{is}\circ\mathsf{tall}\circ\varphi;\ \mathbf{tall}(d)(\mathbf{m});\ \mathrm{S}}{\lambda\varphi.\mathsf{mary}\circ\mathsf{is}\circ\mathsf{tall}\circ\varphi;\ \lambda d.\mathbf{tall}(d)(\mathbf{m});\ \mathrm{S}{\upharpoonright}\mathrm{Deg}}{\upharpoonright}\mathrm{I}^1$$

$$\dfrac{\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{er})\circ\mathsf{than}\circ\sigma_2(\boldsymbol{\varepsilon});\ \lambda P\lambda Q.\mathbf{max}(P) > \mathbf{max}(Q);\ \mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{Deg}){\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{Deg}) \qquad \lambda\varphi.\mathsf{mary}\circ\mathsf{is}\circ\mathsf{tall}\circ\varphi;\ \lambda d.\mathbf{tall}(d)(\mathbf{m});\ \mathrm{S}{\upharpoonright}\mathrm{Deg}}{\lambda\sigma_2.\mathsf{mary}\circ\mathsf{is}\circ\mathsf{tall}\circ\mathsf{er}\circ\mathsf{than}\circ\sigma_2(\boldsymbol{\varepsilon});\ \lambda Q.\mathbf{max}(\lambda d.\mathbf{tall}(d)(\mathbf{m})) > \mathbf{max}(Q);\ \mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{Deg}){\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{Deg})}$$

$$\dfrac{\lambda\sigma_2.\mathsf{mary}\circ\mathsf{is}\circ\mathsf{tall}\circ\mathsf{er}\circ\mathsf{than}\circ\sigma_2(\boldsymbol{\varepsilon});\ \lambda Q.\mathbf{max}(\lambda d.\mathbf{tall}(d)(\mathbf{m})) > \mathbf{max}(Q);\ \mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{Deg}){\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{Deg}) \qquad \begin{matrix}\vdots \\ \lambda\varphi.\mathsf{ann}\circ\mathsf{is}\circ\mathsf{tall}\circ\varphi;\ \lambda d.\mathbf{tall}(d)(\mathbf{m});\ \mathrm{S}{\upharpoonright}\mathrm{Deg}\end{matrix}}{\mathsf{mary}\circ\mathsf{is}\circ\mathsf{tall}\circ\mathsf{er}\circ\mathsf{than}\circ\mathsf{ann}\circ\mathsf{is}\circ\mathsf{tall};\ \mathbf{max}(\lambda d.\mathbf{tall}(d)(\mathbf{m})) > \mathbf{max}(\lambda d.\mathbf{tall}(d)(\mathbf{a}));\ \mathrm{S}}$$

Whether (383a) is derived from (384) via VP ellipsis or the missing adjective meaning is instead supplied by the comparative operator is debatable, but since the latter option is needed in the analysis of "attributive comparative deletion" examples, we illustrate this latter option here. This option also makes it possible to encode the comparative

6. (384) may sound less natural than (383a), but this is arguable due to the redundancy of repeating the adjective *tall* in the *than* clause.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

morphology via a functional prosodic term, which is convenient for the treatment of suppletive forms. Specifically we assume that the **st**→**st** function ER takes the string prosody of the base form of an adjective and returns the comparative form of the adjective.

On this assumption, there is a slightly more complicated variant of the comparative operator defined as in (388):

(388) $\lambda\sigma_0\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{ER}(\sigma_0(\boldsymbol{\varepsilon}))) \circ \mathsf{than} \circ \sigma_2(\boldsymbol{\varepsilon});\ \ \lambda f\lambda P\lambda Q.\mathbf{max}(\lambda d.P(f(d))) > \mathbf{max}(\lambda d.Q(f(d)));\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{Adj})\upharpoonright(\mathrm{S}\upharpoonright\mathrm{Adj})\upharpoonright(\mathrm{Adj}\upharpoonright\mathrm{Deg})$

The difference between (386) and (388) is that in (388), the two clauses are missing the entire adjective rather than just the degree argument. Correspondingly, in (388), the comparative operator lowers the comparative form of the adjective in the first clause in the prosodic component. Obtaining the right degree descriptions for the two clauses is straightforward given the semantics of the comparative operator in (388).

With the definition in (388), the derivation for (383a) goes as in (389):

(389)

$$
\dfrac{
\dfrac{
\dfrac{
\begin{array}{l}\lambda\sigma_0\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{ER}(\sigma_0(\boldsymbol{\varepsilon})))\circ\\ \quad\mathsf{than}\circ\sigma_2(\boldsymbol{\varepsilon});\\ \lambda f\lambda P\lambda Q.\mathbf{max}(\lambda d.P(f(d))) >\\ \quad\mathbf{max}(\lambda d.Q(f(d)));\\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{Adj})\upharpoonright(\mathrm{S}\upharpoonright\mathrm{Adj})\upharpoonright(\mathrm{Adj}\upharpoonright\mathrm{Deg})\end{array}
\quad
\begin{array}{l}\lambda\varphi.\mathsf{tall}\circ\varphi;\\ \mathbf{tall};\\ \mathrm{Adj}\upharpoonright\mathrm{Deg}\end{array}
}{
\begin{array}{l}\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{ER}(\mathsf{tall}))\circ\mathsf{than}\circ\sigma_2(\boldsymbol{\varepsilon});\\ \lambda P\lambda Q.\mathbf{max}(\lambda d.P(\mathbf{tall}(d))) >\\ \quad\mathbf{max}(\lambda d.Q(\mathbf{tall}(d)));\\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{Adj})\upharpoonright(\mathrm{S}\upharpoonright\mathrm{Adj})\end{array}
}
\quad
\dfrac{
\dfrac{
\begin{array}{l}\mathsf{mary};\\ \mathbf{m};\\ \mathrm{NP}\end{array}
\quad
\dfrac{
\begin{array}{l}\mathsf{is};\\ \lambda P.P;\\ \mathrm{VP/Adj}\end{array}
\quad
\begin{bmatrix}\varphi;\\ P;\\ \mathrm{Adj}\end{bmatrix}^1
}{\mathsf{is}\circ\varphi;\ P;\ \mathrm{VP}}
}{\mathsf{mary}\circ\mathsf{is}\circ\varphi;\ P(\mathbf{m});\ \mathrm{S}}
}{
\begin{array}{l}\lambda\varphi.\mathsf{mary}\circ\mathsf{is}\circ\varphi;\\ \lambda P.P(\mathbf{m});\mathrm{S}\upharpoonright\mathrm{Adj}\end{array}
}{\upharpoonright\mathrm{I}^1}
}{
\begin{array}{l}\lambda\sigma_2.\mathsf{mary}\circ\mathsf{is}\circ\mathsf{tall}\circ\mathsf{er}\circ\mathsf{than}\circ\sigma_2(\boldsymbol{\varepsilon});\\ \lambda Q.\mathbf{max}(\lambda d.\mathbf{tall}(d)(\mathbf{m})) > \mathbf{max}(\lambda d.Q(\mathbf{tall}(d)));\mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{Adj})\end{array}
}
\quad
\begin{array}{l}\vdots\\ \lambda\varphi.\mathsf{ann}\circ\\ \quad\mathsf{is}\circ\varphi;\\ \lambda P.P(\mathbf{a});\\ \mathrm{S}\upharpoonright\mathrm{Adj}\end{array}
}{
\mathsf{mary}\circ\mathsf{is}\circ\mathsf{tall}\circ\mathsf{er}\circ\mathsf{than}\circ\mathsf{ann}\circ\mathsf{is};\ \mathbf{max}(\lambda d.\mathbf{tall}(d)(\mathbf{m})) > \mathbf{max}(\lambda d.\mathbf{tall}(d)(\mathbf{a}));\ \mathrm{S}
}
$$

8.2.2.2 VP ellipsis in comparatives The analysis of comparatives sketched above interacts straightforwardly with the analysis of VP ellipsis from chapter 6 to yield the right meanings for examples like the following:

(390) John ran faster than Mary did.

Here again, there can be two possible analyses of the missing status of *fast* in the *than* clause. We assume without argument here that the adverb is part of the "elided material." An analysis in which the comparative operator is responsible for the recovery of the adverb meaning (corresponding to the derivation in (389)) is also straightforward in our approach.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

We assume that, like adjectives, adverbs that have a gradable component in meaning take an additional degree argument (syntactically of type Deg and semantically of type d).

(391) $\lambda\varphi.\text{fast}\circ\varphi;\ \mathbf{fast};\ (\mathrm{VP}\backslash\mathrm{VP})\upharpoonright\mathrm{Deg}$

With this lexical entry for the adverb *fast*, the derivation for (390) goes as follows (here, we abbreviate Deg as D):

(392)

$$\dfrac{\begin{array}{l}\lambda\sigma_1\lambda\sigma_2.\sigma_1(\text{er})\circ\\ \quad\text{than}\circ\sigma_2(\boldsymbol{\varepsilon});\\ \lambda P\lambda Q.\mathbf{max}(P)\\ \quad>\mathbf{max}(Q);\\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{D})\upharpoonright(\mathrm{S}\upharpoonright\mathrm{D})\end{array} \quad \dfrac{\dfrac{\begin{array}{l}\text{john};\\ \mathbf{j};\\ \mathrm{NP}\end{array} \quad \dfrac{\begin{bmatrix}\varphi;\\ d;\\ \mathrm{D}\end{bmatrix}^2 \quad \dfrac{\dfrac{\begin{array}{l}\text{ran};\\ \mathbf{ran};\\ \mathrm{VP}\end{array} \quad \dfrac{\begin{array}{l}\lambda\varphi.\text{fast}\circ\varphi;\\ \mathbf{fast};(\mathrm{VP}\backslash\mathrm{VP})\upharpoonright\mathrm{D}\end{array} \quad \begin{bmatrix}\varphi;\\ d;\mathrm{D}\end{bmatrix}^3}{\text{fast}\circ\varphi;\ \mathbf{fast}(d);\ (\mathrm{VP}\backslash\mathrm{VP})}}{\text{ran}\circ\text{fast}\circ\varphi;\ \mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}}\upharpoonright\mathrm{I}^3}{\begin{array}{l}\lambda\varphi.\text{ran}\circ\text{fast}\circ\varphi;\\ \lambda d.\mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}\upharpoonright\mathrm{Deg}\end{array}}}{\text{ran}\circ\text{fast}\circ\varphi;\ \mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}}}{\text{john}\circ\text{ran}\circ\text{fast}\circ\varphi;\ \mathbf{fast}(d)(\mathbf{ran})(\mathbf{j});\ \mathrm{S}}\upharpoonright\mathrm{I}^2}{\begin{array}{l}\lambda\varphi.\text{john}\circ\text{ran}\circ\text{fast}\circ\varphi;\\ \lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{j});\mathrm{S}\upharpoonright\mathrm{D}\end{array}}}{\begin{array}{l}\lambda\sigma_2.\text{john}\circ\text{ran}\circ\text{fast}\circ\text{er}\circ\text{than}\circ\sigma_2(\boldsymbol{\varepsilon});\\ \lambda Q.\mathbf{max}(\lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{j}))>\mathbf{max}(Q);\mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{D})\end{array}}$$

$$\dfrac{\begin{array}{l}\quad\vdots\\ \lambda\sigma_2.\text{john}\circ\text{ran}\circ\\ \quad\text{fast}\circ\text{er}\circ\\ \quad\text{than}\circ\sigma_2(\boldsymbol{\varepsilon});\\ \lambda Q.\mathbf{max}(\lambda d.\\ \quad\mathbf{fast}(d)(\mathbf{ran})(\mathbf{j}))\\ \quad>\mathbf{max}(Q);\\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{D})\end{array} \quad \dfrac{\dfrac{\begin{array}{l}\text{mary};\\ \mathbf{m};\\ \mathrm{NP}\end{array} \quad \dfrac{\begin{bmatrix}\varphi;\\ d;\\ \mathrm{D}\end{bmatrix}^1 \quad \dfrac{\begin{array}{l}\lambda\rho\lambda\varphi.\rho(\lambda\varphi.\varphi)(\varphi);\\ \lambda\mathscr{G}.\mathscr{G}(\lambda d.\mathbf{fast}(d)(\mathbf{ran}));\\ (\mathrm{VP}\upharpoonright\mathrm{D})\upharpoonright((\mathrm{VP}\upharpoonright\mathrm{D})\upharpoonright(\mathrm{VP}\upharpoonright\mathrm{D}))\end{array} \quad \begin{array}{l}\lambda\sigma\lambda\varphi.\text{did}\circ\sigma(\varphi);\\ \lambda\mathscr{F}.\mathscr{F};\\ (\mathrm{VP}\upharpoonright\mathrm{D})\upharpoonright(\mathrm{VP}\upharpoonright\mathrm{D})\end{array}}{\lambda\varphi.\text{did}\circ\varphi;\ \lambda d.\mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}\upharpoonright\mathrm{D}}}{\text{did}\circ\varphi;\ \mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}}}{\text{mary}\circ\text{did}\circ\varphi;\ \mathbf{fast}(d)(\mathbf{ran})(\mathbf{m});\ \mathrm{S}}\upharpoonright\mathrm{I}^1}{\lambda\varphi.\text{mary}\circ\text{did}\circ\varphi;\ \lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{m});\ \mathrm{S}\upharpoonright\mathrm{D}}}{\begin{array}{l}\text{john}\circ\text{ran}\circ\text{fast}\circ\text{er}\circ\text{than}\circ\text{mary}\circ\text{did};\\ \mathbf{max}(\lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{j}))>\mathbf{max}(\lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{m}));\mathrm{S}\end{array}}$$

In somewhat informal terms, on this analysis what is "elided" in the *than* clause in (390) is a VP containing a "gap" position for the degree argument: "ran __ fast." In Hybrid TLCG terms, this will correspond to an expression of type $\mathrm{VP}\upharpoonright\mathrm{Deg}$, with the denotation $\lambda d.\mathbf{fast}(d)(\mathbf{ran})$. Thus, the derivation is parallel to the extraction/ellipsis interaction case in (382) from section 8.2.1, with the only difference being that the missing expression here is of type Deg rather than type NP.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

8.3 Kennedy's Arguments: A Second Look

We now revisit Kennedy's arguments as summarized in section 8.1, in light of the alternative analysis of extraction/ellipsis interaction just sketched, apart from his island arguments (which we take up in chapter 10). In the following discussion we demonstrate that, in every case, the data that Kennedy takes to establish the inevitability of a covert structure analysis for ellipsis phenomena can be accounted for far more straightforwardly on a "direct interpretation" approach. In some cases, Kennedy's assumptions about the structural basis of the diagnostic probes he employs are undermotivated, while in others, the fact that on our approach examples such as (376) do involve an actual extraction renders his argument moot.

8.3.1 Ellipsis and Islandhood

As noted above, we discuss Kennedy's arguments involving islandhood in chapter 10 (section 10.2.4), together with similar arguments for structural analyses of other phenomena. To preview the conclusion, the judgment patterns that Kennedy gives to support his deletion analyses either receive alternative explanations or simply are not robust enough to support the existence of islandhood as a syntactic condition.

8.3.2 Ellipsis and SCO

Kennedy's argument based on SCO, as outlined in section 8.1.2, hinges on the assumption that data such as (366) indeed reflect extraction from an elided VP.

(366) Who$_i$ will Mary vote for t_i if he$_{j/*i}$ does ~~vote for t_i~~?

However, on the alternative analysis presented in section 8.2.1, an analogue of the following representation of the relevant coindexing relations in (366) is licensed:

(393) *wh$_i$ . . . he$_i$ [$_{VP}$ [$_{VP/NP}$ does] t_i]

The point is that what is extracted in (393) could correspond, on this analysis, to an argument of the auxiliary itself. Thus, any account of SCO which motivates ordinary cases such as (365) will predict SCO in pseudogapping + extraction cases of the sort exhibited in (366) on exactly the same basis, with no covert structure.[7]

In other words, it follows that on the extended valence analysis of auxiliaries, the SCO facts that Kennedy takes as prima facie support for covert structure in VP ellipsis

7. Although it is true that, in principle, nothing excludes a transformational counterpart to our pseudogapping-based analysis and that any such analysis undercuts Kennedy's use of SCO effects to defend covert structure, whether or not such a transformational treatment is possible is a complex question. The remnants in pseudogapping examples represent the survivors of a movement + deletion process in transformational approaches. The eligibility of such an expression for further movement depends on a possibly intricate network of further assumptions about iterated movement which in practice appear to vary from author to author.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

are just another expected consequence of auxiliary behavior, fully compatible with a nonstructural treatment of those facts.

We do not formulate an account of SCO in our approach, and acknowledge that the phenomenon originally described in Postal (1971) is difficult to analyze—with the caveat that this holds for all current theories, including the Principles and Parameters approach.

8.3.3 Ellipsis and Condition B

For the analysis of the so-called Binding Condition B, we adopt Jacobson's approach, representing part of a long tradition of non-configurational approaches to anaphoric relations amongst nominal terms. In this work, Jacobson offers a thorough critique of the standard configuration-based approach to "Condition B" effects, concluding that it faces "problems which seem insurmountable" (2007, 218).[8] The key idea in Jacobson's alternative, non-configurational account of Condition B effects is that NPs are divided by a binary-valued feature $\pm p$, with pronouns marked NP_{+p} and all other NPs NP_{-p}. In all lexical entries of the form in (394), all NP (and PP) arguments in any realization of /\$ are specified as $-p$.[9]

(394) k; P; VP/\$

The effect of this restriction is to rule out pronouns from argument positions of verbs with ordinary semantic denotations. On this approach, the only way a lexically specified functional category can take $+p$ arguments is via the application of the following irreflexive operator:[10]

8. The pioneering study of anaphora in Bach and Partee (1980) offered a purely compositional semantic treatment of reflexives which hinges on treating anaphors in a fashion parallel to generalized quantifiers as scope-taking elements. In HPSG, on the other hand, anaphoric possibilities are determined by neither phrase structure configuration nor semantic composition but by the combinatory argument structure specified for heads (Pollard and Sag 1983, 1992, 1994), an approach subsequently adopted by LFG (Bresnan et al. 2016). Even within the general Principles and Parameters framework, there are both older and more recent treatments which do not derive the anaphora possibilities from configurational representations (see, e.g., Reinhart and Reuland 1993; Safir 2004). It is worth noting that efforts to extend the Bach-Partee approach to cover Condition B were not particularly successful until Jacobson's (2007) work; see Jäger (2005, chap. 2) and Morrill (2010, 124–125) for discussion.

9. We take VP to abbreviate NP\S, where NP is underspecified for the value of the feature p. Unlike pronouns in the object position, no special conditions or restrictions apply to subject pronouns, hence the choice of $\pm$ value for such pronouns need not be specified.

10. For expository purposes, we state the operator in (395) in its most restricted form, dealing with only the case where there is a single syntactic argument apart from the subject. A much broader coverage is of course necessary in order to handle cases like the following:

(i) a. *John$_i$ warned Mary about him$_i$.
b. *John talked to Mary$_i$ about her$_i$.
c. *John explained himself$_i$ to him$_i$.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

(395) $\lambda\varphi.\varphi;\ \lambda f \lambda u \lambda v. f(u)(v),\ u \neq v\ ;\ (\mathrm{VP}/\mathrm{NP}_{+p}) \upharpoonleft (\mathrm{VP}/\mathrm{NP}_{-p})$

The grayed-in part $u \neq v$ separated from the truth conditional meaning by a comma is a presupposition introduced by the pronoun-seeking variant of the predicate. It says that the subject and object arguments are forced to pick out different objects in the model. For the semantics of pronouns themselves, we continue to assume from chapter 6 that free (i.e., unbound) pronouns are simply translated as arbitrary variables (cf. Cooper 1979).

Crucially, the operator in (395) is restricted in its domain of application to the set of signs which are specified in the lexicon. We notate this restriction by using the dashed line notion in what follows. Then *John praises him* will be derived as in (396):

(396)

$$\dfrac{\dfrac{\dfrac{\begin{array}{l}\lambda\varphi.\varphi;\\ \lambda f \lambda u \lambda v. f(u)(v), u \neq v;\\ (\mathrm{VP}/\mathrm{NP}_{+p}) \upharpoonleft (\mathrm{VP}/\mathrm{NP}_{-p})\end{array} \quad \begin{array}{l}\mathsf{praises};\\ \mathbf{praise}; \mathrm{VP}/\mathrm{NP}_{-p}\end{array}}{\mathsf{praises};\ \lambda u \lambda v.\mathbf{praise}(u)(v), u \neq v;\ \mathrm{VP}/\mathrm{NP}_{+p}} \quad \mathsf{him};\ z;\ \mathrm{NP}_{+p}}{\mathsf{praises} \circ \mathsf{him};\ \lambda v.\mathbf{praise}(z)(v), z \neq v;\ \mathrm{VP}} \quad \mathsf{john};\ \mathbf{j};\ \mathrm{NP}_{-p}}{\mathsf{john} \circ \mathsf{praises} \circ \mathsf{him};\ \mathbf{praise}(z)(\mathbf{j}), z \neq \mathbf{j};\ \mathrm{S}}$$

The presupposition $z \neq \mathbf{j}$ ensures that the referent of the pronoun is different from John.

This approach extends straightforwardly to bound variable readings of pronouns. We first illustrate how a well-formed example of bound variable anaphora such as (397) can be analyzed in this setup, and then explain the Condition B effects for bound pronouns.

(397) Every editor$_i$ believes that John admires him$_i$.

For the bound variable reading of pronouns, we assume the following lexical entry, which encodes nonlinearity enforced by binding in the lexical specification (note that,

---

What we need in effect is a schematic type specification that applies to a pronoun in any or all argument positions, i.e., stated on an input of the form $\mathrm{VP}/\$/\mathrm{XP}_{-p}/\$$ to yield an output of the form $\mathrm{VP}/\$/\mathrm{XP}_{+p}/\$$. To ensure the correct implementation of this extension, some version of the "wrapping" analysis needs to be assumed (cf. Jacobson 2007), so that the order of the arguments in verbs' lexical entries is isomorphic to the obliqueness hierarchy (of the sort discussed by Pollard and Sag [1992]).

Cases such as the following also call for an extension (also a relatively straightforward one):

(ii) *John$_i$ is proud of him$_i$.

By assuming (following Jacobson 2007) that the $\pm p$ feature percolates from NPs to PPs and by generalizing the irreflexive operator still further so that it applies not just to $\mathrm{VP}/\mathrm{XP}_{-p}$ but also to $\mathrm{AP}/\mathrm{XP}_{-p}$, the ungrammaticality of (ii) follows straightforwardly.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

in the semantic term, two tokens of the same variable w are bound by the same lambda operator):[11]

(398) $\lambda\sigma.\sigma(\mathsf{him});\ \lambda\mathscr{F}\lambda w.\mathscr{F}(w)(w);\ \mathrm{VP}\upharpoonright(\mathrm{VP}\upharpoonright\mathrm{NP}_{+p})$

With this lexical sign for *him*, we obtain the correct interpretation of (397) as follows (here, $\mathbf{V_{editor}}$ is an abbreviation for the term $\lambda Q.\forall x[\mathbf{editor}(x) \rightarrow Q(x)]$):

(399)

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdots \quad \mathsf{admires};\ \lambda u\lambda v.\mathbf{admire}(u)(v), u \neq v;\ \mathrm{VP}/\mathrm{NP}_{+p} \qquad [\varphi_0;\ z;\ \mathrm{NP}_{+p}]^0}{\mathsf{admires} \circ \varphi_0;\ \lambda v.\mathbf{admire}(z), z \neq v;\ \mathrm{VP}} \qquad \mathsf{john};\ \mathbf{j};\ \mathrm{NP}_{-p}}{\mathsf{john} \circ \mathsf{admires} \circ \varphi_0;\ \mathbf{admire}(z)(\mathbf{j}), z \neq \mathbf{j}; \mathrm{S}} \qquad \mathsf{thinks};\ \mathbf{think};\ \mathrm{VP}/\mathrm{S}}{\mathsf{thinks} \circ \mathsf{john} \circ \mathsf{admires} \circ \varphi_0;\ \mathbf{think}(\mathbf{admire}(z)(\mathbf{j})), z \neq \mathbf{j}; \mathrm{VP}}{\upharpoonright}\mathrm{I}^0}{\lambda\varphi_0.\mathsf{thinks} \circ \mathsf{john} \circ \mathsf{admires} \circ \varphi_0;\ \lambda z.\mathbf{think}(\mathbf{admire}(z)(\mathbf{j})), z \neq \mathbf{j};\ \mathrm{VP}\upharpoonright\mathrm{NP}_{+p}} \qquad \lambda\sigma_0.\sigma_0(\mathsf{him});\ \lambda\mathscr{F}\lambda w.\mathscr{F}(w)(w);\ \mathrm{VP}\upharpoonright(\mathrm{VP}\upharpoonright\mathrm{NP}_{+p})}{\mathsf{thinks} \circ \mathsf{john} \circ \mathsf{admires} \circ \mathsf{him};\ \lambda w.\mathbf{think}(\mathbf{admire}(w)(\mathbf{j}))(w), w \neq \mathbf{j}; \mathrm{VP}} \qquad [\varphi_1;\ y;\ \mathrm{NP}]^1}{\varphi_1 \circ \mathsf{thinks} \circ \mathsf{john} \circ \mathsf{admires} \circ \mathsf{him};\ \mathbf{think}(\mathbf{admire}(y)(\mathbf{j}))(y), y \neq \mathbf{j}; \mathrm{S}}{\upharpoonright}\mathrm{I}^1}{\lambda\varphi_1.\varphi_1 \circ \mathsf{thinks} \circ \mathsf{john} \circ \mathsf{admires} \circ \mathsf{him};\ \lambda y.\mathbf{think}(\mathbf{admire}(y)(\mathbf{j}))(y), y \neq \mathbf{j}; \mathrm{S}\upharpoonright\mathrm{NP}} \qquad \lambda\sigma_1.\sigma_1(\mathsf{every} \circ \mathsf{editor});\ \mathbf{V_{editor}};\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{-p})}{\mathsf{every} \circ \mathsf{editor} \circ \mathsf{thinks} \circ \mathsf{john} \circ \mathsf{admires} \circ \mathsf{him};\ \mathbf{V_{editor}}(\lambda y.\mathbf{think}(\mathbf{admire}(y)(\mathbf{j}))(y), y \neq \mathbf{j}); \mathrm{S}}$$

11. One might worry about the duplication of lexical entries for pronouns for the free and bound uses. It is in fact straightforward to unify the two by deriving the bound form from the free form by the following operator (which, like the irreflexive operator, is a lexical operator), with α a variable over case values:

(i) $\lambda\varphi\lambda\sigma.\sigma(\varphi);\ \lambda v\lambda\mathscr{G}\lambda w.\mathscr{G}(w)(w);\ (\mathrm{VP}\upharpoonright(\mathrm{VP}\upharpoonright\mathrm{NP}_{+p,\alpha}))\upharpoonright\mathrm{NP}_{+p,\alpha}$

By applying (i) to the free pronoun (ii), we obtain the bound pronoun entry identical to (398).

(ii) *him*; u; $\mathrm{NP}_{+p,acc}$

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

The bound variable reading for the following example is blocked for essentially the same reason that the coreferential reading is blocked in simpler examples like (396).

(400) *Every editor$_i$ congratulated him$_i$.

The forced application of the irreflexive operator to the verb lexical entry dictated by the pronoun has the effect that the verb's subject and object have disjoint reference. However, this directly contradicts the subject quantifier's binding the object pronoun, thus rendering the sentence uninterpretable on the intended reading. The derivation is shown in (401):

(401)

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdots \atop \begin{array}{l}\textsf{congratulated};\\ \lambda u\lambda v.\textbf{congratulated}(u)(v),\\ \quad u \neq v;\\ \mathrm{VP/NP}_{+p}\end{array} \qquad \begin{bmatrix}\varphi_0;\\ w;\\ \mathrm{NP}_{+p}\end{bmatrix}^0}{\begin{array}{l}\textsf{congratulated} \circ \varphi_0;\\ \lambda v.\textbf{congratulated}(w)(v), w \neq v; \mathrm{VP}\end{array}}}{\begin{array}{l}\lambda\varphi_0.\textsf{congratulated} \circ \varphi_0;\\ \lambda w\lambda v.\textbf{congratulated}(w)(v), w \neq v;\\ \mathrm{VP}{\upharpoonright}\mathrm{NP}_{+p}\end{array}}{\upharpoonright}\mathrm{I}^0 \qquad \begin{array}{l}\textsf{him};\\ \lambda\mathscr{G}\lambda y.\mathscr{G}(y)(y);\\ \mathrm{VP}{\upharpoonright}(\mathrm{VP}{\upharpoonright}\mathrm{NP}_{+p})\end{array}}{\begin{array}{l}\textsf{congratulated} \circ \textsf{him};\\ \lambda y.\textbf{congratulated}(y)(y), y \neq y; \mathrm{VP}\end{array}} \qquad \begin{bmatrix}\varphi_1;\\ x;\\ \mathrm{NP}\end{bmatrix}^1}{\begin{array}{l}\varphi_1 \circ \textsf{congratulated} \circ \textsf{him};\\ \textbf{congratulated}(x)(x), x \neq x; \mathrm{S}\end{array}}}{\begin{array}{l}\lambda\varphi_1.\varphi_1 \circ \textsf{congratulated} \circ \textsf{him};\\ \lambda x.\textbf{congratulated}(x)(x), x \neq x; \mathrm{S}{\upharpoonright}\mathrm{NP}\end{array}}{\upharpoonright}\mathrm{I}^1 \qquad \begin{array}{l}\lambda\sigma_1.\sigma_1(\textsf{every} \circ\\ \quad \textsf{editor});\\ \mathbf{V}_{\textbf{editor}};\\ \mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{-p})\end{array}}{\begin{array}{l}\textsf{every} \circ \textsf{editor} \circ \textsf{congratulated} \circ \textsf{him};\\ \mathbf{V}_{\textbf{editor}}(\lambda x.\textbf{congratulated}(x)(x), x \neq x); \mathrm{S}\end{array}}$$

With this strictly semantic approach to "Condition B" effects in hand, let us return to Kennedy's argument based on such effects. Recall that Kennedy's argument rested on the assumption that in examples such as (402) (= (367) from section 8.1.3) the obligatorily disjoint interpretation of the two pronouns requires a specific covert structure as the supposed basis for this enforced interpretation.

(402) John$_k$ takes care of him$_i$ because he$_{j/*i}$ won't.

But the interaction of our lexically based semantic treatment of "Condition B" effects with the analysis of ellipsis from section 8.2 automatically accounts for the pattern in

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

(402). The derivation is given in (403) (here, we ignore the tense and modal meaning of the future tense auxiliary *won't*; **irr** abbreviates the term $\lambda f \lambda u \lambda v. f(u)(v), u \neq v$).[12]

(403)

$$\dfrac{\dfrac{\text{takes} \circ \text{care} \circ \text{of}; \ \mathbf{irr}(\mathbf{take\text{-}care}); \mathrm{VP/NP}_{+p} \quad \text{him}; w; \mathrm{NP}_{+p}}{\text{takes} \circ \text{care} \circ \text{of} \circ \text{him}; \mathbf{irr}(\mathbf{take\text{-}care})(w); \mathrm{VP}} \quad \text{john}; \mathbf{j}; \mathrm{NP}_{-p}}{\text{john} \circ \text{takes} \circ \text{care} \circ \text{of} \circ \text{him}; \mathbf{irr}(\mathbf{take\text{-}care})(w)(\mathbf{j}); \mathrm{S}}$$

$$\dfrac{\text{won't}; \neg\mathbf{irr}(\mathbf{take\text{-}care})(w); \mathrm{VP} \quad \text{he}; y; \mathrm{NP}_{+p}}{\text{he} \circ \text{won't}; \neg\mathbf{irr}(\mathbf{take\text{-}care})(w)(y); \mathrm{S}}$$

Since $\mathbf{irr}(\mathbf{take\text{-}care})(w)(y) \equiv \mathbf{take\text{-}care}(w)(y), w \neq y$, it is guaranteed that the subject pronoun is disjoint in reference from the object (note that the term $w \neq y$, when evaluated at a given model and a value assignment g, entails that the individuals assigned to w and y by g are distinct).

Examples like (402), therefore, offer no support to the claim that covert syntactic structure must be present in VP ellipsis. The disjoint reference requirement follows equally straightforwardly in a nonstructural analysis of VP ellipsis combined with a nonstructural analysis of "Condition B" effects. It is important to keep in mind that, though these perspectives on ellipsis and binding depart from the standard view, they both receive ample empirical support, since their advantages over the standard view have independently been established in the literature.

8.3.4 Ellipsis and Parasitic Gaps

We now extend our approach to Kennedy's parasitic gap data, demonstrating that, again, appeal to hidden configuration is unnecessary. Consider again (404) and (405) ((369) and (370) from section 8.1.4):

(404) a. *Otis is a person who$_i$ I admire $\_\__i$ because close friends of $\_\__i$ became famous.

b. ?Otis is a person who I admire because close friends of $\_\__i$ admire $\_\__i$/*him$_i$.

(405) Otis is a person who$_i$ I admire $\_\__i$ because close friends of $\_\__i$ seem to $\varnothing$.

12. We take the idiom *takes care of* to be a complex transitive verb, as per (i):

(i) take ∘ care ∘ of; **take-care-of**; VP/NP

The treatment reflected in (i) is strongly supported by the passivization and extraction pattern in (ii).

(ii) a. Our problem seems to have been taken care of.
b. *Care was taken of that problem.
c. ??*Of which problem have you taken care?

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

Here we annotate Kennedy's original datum in (404b) with our own judgment, which, while taking this example to be grammatical, assigns it a rather more diminished status than does Kennedy.

The central issue about parasitic gaps comes down to the following question: Why is (404b) (in the gap version) reasonably acceptable but (404a) bad? It has been widely assumed in the literature that the contrast between (404a) and (404b) is a grammatical one, that is, that the "parasitic" gap in the subject position cannot be licensed except for the presence of a supporting "real gap" in the main VP. But in recent years, this position has begun to give way to an alternative perspective, in which processing-based factors bear the primary responsibility for the contrast between (404a) and (404b), as we discuss in greater detail below (see also section 10.1.3.1).

Note first that, just as in the case of extractions from other putative island positions, subject gap sites are more or less permissible in principle. Thus we find examples such as the following (where the [$_\pi$. . .] notation identifies the bracketed material as a prosodic phrase):

(406) a. There are certain topics which [$_\pi$ jokes about __] [$_\pi$ are completely unacceptable.] (Levine and Sag 2003).

b. [$_\pi$ Of which cars] [$_\pi$ were only the hoods __] [$_\pi$ damaged by the explosion?] (Ross 1967, 252, cited in Chaves 2013)

c. They have eight children [$_\pi$ of whom] [$_\pi$ five __ are still living at home.] (Huddleston et al. 2002, 1093, cited in Chaves 2013)

d. [$_\pi$ That is the lock] [$_\pi$ to which] [$_\pi$ the key __ has been lost.] (Levine and Hukari 2006)

e. [$_\pi$ Which disease] [$_\pi$ will the cure for __] [$_\pi$ never be discovered]? (Chaves 2013, 17)

Chaves (2013) and Chaves and Dery (2019) contain many such examples, which point to the conclusion that subjects are not, in themselves, island contexts. On this view, examples such as (407) are grammatical, but unacceptable for reasons explored in Kluender (2004), Chaves (2013), and Chaves and Dery (2019).

(407) ??*Which people did friends of __ become famous?

But there appears to be an important exception to this pattern. When otherwise reasonably acceptable examples of subject-internal gaps appear as adjunct clauses in parasitic gap constructions (i.e., when the relevant gap is an adjunct parasitic gap) rather than as main clauses, there is a dramatic drop in speakers' assessments of their status.

(408) a. ??*Which disease$_i$ do people rightly fear $\_\__i$ because the cure for $\_\__i$ will never be discovered?

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

b. ??*There are certain topics which$_i$ comedians ignore __$_i$ because jokes about __$_i$ are completely unacceptable.

These examples essentially have the same syntactic structure as (404a).

Thus, we have four classes of data to account for (the contrast between 3 and 4 is what crucially supports Kennedy's argument from parasitic gap–licensing patterns):

1. unsupported *wh* extraction from subjects (sometimes good (407); sometimes bad (406))
2. supported *wh* extraction from subjects (typically fine)
3. embedded unsupported *wh* extraction from subjects ((408); markedly worse than 1)
4. embedded supported *wh* extraction from subjects ((404b); markedly better than 3)

The key explanatory mechanism for 1–2 is offered in Chaves (2013), based on the premise that parsing efficiency considerations have led to a strategy in which, once a filler has been encountered and stored, the processor expects to find no corresponding gap within the subject, but strongly anticipates a gap in the immediately following VP. This is discussed in greater detail in chapter 10. Cases such as (407), on this view, are unacceptable precisely because they violate both of these presumptions. As Chaves notes, prosodic or pragmatic cues can ameliorate these violations and allow the processor to successfully link the filler to gap sites even where neither of the parser's wired-in "first pass" expectations is satisfied (as attested in (406)), but such cues have to be quite prominent to offset the effect of those expectations failing. On the other hand, simple subject parasitic gap cases as in 2 violate only one of the parser's expectations and are thus predicted to be significantly better than 1.

The final issue corresponds to cases in which we have an unsupported and supported parasitic gap in an adjunct clause subject (3 and 4). Here too, we have an answer ready. Pattern 3 is significantly worse than pattern 1 since it violates a parser's expectations in *three* ways. Specifically, the gap is inside an adjunct clause, it is in the subject position, and there is no supporting gap in the following VP. With three independent violations of the processor's parsing strategy, it is not surprising that such examples are severely ill-formed.[13] By contrast, 4 is markedly better than 3 because it violates only two of the three conditions. Thus, with no further development or modification, Chaves's model

13. As noted in Kluender (1998), each successive clausal boundary (or its real-time parsing analogue) represents an additional processing bottleneck which adds significant costs to the parser's efforts to link stored fillers with identifiable gap sites. Subsequent work reported in Kluender (2004) suggests that subject-internal gaps are intrinsically harder to process than gaps in other parts of the sentence, but are not ruled out in the grammar. Kluender (2004) offers well-formed examples (including some first noted in Ross [1967]) in support of this claim but does not offer an explanation for why subsequent gaps within the VP ameliorate the processing difficulties involved.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

already accounts for what we have identified as the key problems posed by subject-internal gaps.

We summarize the situation in the following table. The check mark indicates that the pattern violates the condition in question.

(409)

	Gap in subject position	No gap in the following VP	Embedding in adjunct clause
1	✓	✓	
2	✓		
3	✓	✓	✓
4	✓		✓

Against the background of this account of parasitic gap licensing, the well-formedness and interpretation of (405), which adds ellipsis in the adjunct clause to the mix, now follows automatically, with no further conditions, restrictions, or machinery. Recall that in the case of the simpler "extraction out of an elided VP" example (380), we identified the extraction "site" as the direct object of the transitive auxiliary *did*. Exactly the same analysis is available in the case of (405). Using the gap-multiplying operator introduced at the end of the preceding chapter, we have proofs such as the following:

(410)

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots \atop \lambda\varphi.\mathsf{to}\circ\varphi;\ \lambda x\lambda y.\mathbf{admire}(x)(y);\ \mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh} \qquad \left[\varphi_1;\ u;\ \mathrm{NP}_{+wh}\right]^1}{\mathsf{to}\circ\varphi_1;\ \lambda y.\mathbf{admire}(u)(y);\ \mathrm{VP}} \qquad \mathsf{seem};\ \mathbf{seem};\ \mathrm{VP/VP}}{\mathsf{seem}\circ\mathsf{to}\circ\varphi_1;\ \lambda w.\mathbf{seem}(\mathbf{admire}(u)(w));\ \mathrm{VP}}{\upharpoonright}\mathrm{I}^1}{\lambda\varphi_1.\mathsf{seem}\circ\mathsf{to}\circ\varphi_1;\ \lambda u\lambda w.\mathbf{seem}(\mathbf{admire}(u)(w));\ \mathrm{VP}{\upharpoonright}\mathrm{NP}_{+wh}}$$

$$\dfrac{\dfrac{\dfrac{\vdots}{\lambda\varphi_1.\mathsf{seem}\circ\mathsf{to}\circ\varphi_1;\ \lambda u\lambda w.\mathbf{seem}(\mathbf{admire}(u)(w));\ (\mathrm{NP}_{-wh}\backslash \mathrm{S}){\upharpoonright}\mathrm{NP}_{+wh}} \qquad \begin{array}{l}\lambda\sigma_1\lambda\sigma_2\lambda\varphi.\ \sigma_2(\varphi)\circ\sigma_1(\varphi);\\ \lambda R\lambda g\lambda x.R(x)(g(x));\\ (\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}){\upharpoonright}(\mathrm{NP}_{-wh}{\upharpoonright}\mathrm{NP}_{+wh}){\upharpoonright}((\mathrm{NP}_{-wh}\backslash \mathrm{S}){\upharpoonright}\mathrm{NP}_{+wh})\end{array}}{\lambda\sigma_2\lambda\varphi.\sigma_2(\varphi)\circ\mathsf{seem}\circ\mathsf{to}\circ\varphi;\ \lambda g\lambda x.\mathbf{seem}(\mathbf{admire}(x)(g(x)));\ (\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}){\upharpoonright}(\mathrm{NP}_{-wh}{\upharpoonright}\mathrm{NP}_{+wh})} \qquad \begin{array}{l}\lambda\varphi_3.\mathsf{the}\circ\mathsf{close}\circ\mathsf{friends}\circ\mathsf{of}\circ\varphi_3;\\ \lambda v.\iota(\lambda z.\mathbf{close\text{-}fr}(v)(z));\\ \mathrm{NP}_{-wh}{\upharpoonright}\mathrm{NP}_{+wh}\end{array}}{\lambda\varphi.\mathsf{the}\circ\mathsf{close}\circ\mathsf{friends}\circ\mathsf{of}\circ\varphi\circ\mathsf{seem}\circ\mathsf{to}\circ\varphi;\ \lambda x.\mathbf{seem}(\mathbf{admire}(x)(\iota(\lambda z.\mathbf{close\text{-}fr}(x)(z))));\ \mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}}$$

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

On our analysis, cases such as (405), licensed by the derivation in (410), should have exactly the same status as their overt counterpart; like the latter, (405) displays a gap in subject position and a corresponding gap in the VP, the critical requirement for full acceptability in English when a subject gap is involved. The gap corresponds to an argument of a transitive auxiliary, but this fact in itself does not make any difference to the status of the example. What is crucial is that there be an actual syntactic "gap" involved and that this gap need not be a configurational object (this latter assumption is shared widely among so-called lexicalist theories of syntax such as G/HPSG and most versions of categorial grammar).[14]

Thus, Kennedy's argument for covert structure based on the distribution of parasitic gaps proves unfounded as well. Combining the independently required operator treatments of simple extraction and parasitic gaps with a generalization of the analysis of pseudogapping in chapter 6 automatically yields the parallelism between the "real" and "deleted" host VPs for subject position parasitic gaps attested in (404b) and (405), without recourse to any actual configurational object that undergoes syntactic deletion. At the same time, the combinatorics of our system (without augmentation of syntactic mechanisms blocking islands along the lines proposed, for example, in Morrill [2017]) license subject island violation examples such as (368a) and (404a), whose deviant status follows on independent grounds from the interactions of the factors discussed in detail in Kluender (2004), Chaves (2013), and Chaves and Dery (2019). On this view, the "parasitism" of subject-internal gaps is not a strictly syntactic phenomenon but an emergent effect of real-time parsing strategy.[15]

---

14. In fact, the processing-based account of parasitic gap acceptability essentially leads us to this conclusion directly. The arguments in Chaves (2013), which present a strong case for this processing basis, can be adopted directly on our account on the assumption that the processing of filler-gap linkages is not dependent on the gap being a configurational object. Pickering and Barry (1991, 250) defend just this possibility (see also Pickering 1993), namely,

> that a categorial grammar can recover information that is specified in phrase structural accounts by the use of empty categories, and . . . that a categorial processing model is capable of making the associations that we have argued to be appropriate to the processing of unbounded dependencies.

The rules in the version of categorial grammar they assume (which is based on CCG) are theorems of the proof theory of Hybrid TLCG and other versions of TLCG, and the analysis of extraction in the two (broad) versions of categorial grammar are the same in the relevant respects. Most important, the specific processing mechanism Pickering and Barry hypothesize—a direct association between the filler and a functional category which is specified for an argument of the same type as the filler—can be carried over directly to our analysis of extraction via hypothetical reasoning. Thus, their arguments and conclusion about the direct linkage of fillers with the predicates that select them directly carry over to our approach.

15. However, while we believe that Chaves's processing account of parasitic gap patterns is very much on the right track, this issue is in a sense orthogonal to the question of whether or not such patterns support the existence of covert structure. On our analysis, an ordinary elided VP and an elided VP supposedly containing a gap site have different syntactic types: the former is VP while the latter is $VP{\upharpoonright}XP$. Correspondingly, their semantic types are different. Thus, as long as there is some way (either processing-based, as we have

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

8.3.5 Ellipsis and Attributive Comparatives

Kennedy's (and Merchant's) argument for covert syntactic structure in VP ellipsis in the attributive comparative construction rests on the premise that the contrast between the elided and non-elided counterparts of attributive comparatives such as the following reflects a difference in *syntactic* well-formedness:

(411) a. *John buys more expensive wine than he buys beer.
b. John buys more expensive wine than he does beer.

We disagree with Kennedy and Merchant on the assessment of the relevant data. Specifically, we take both (411a) and (411b) to be licensed by the grammar and assume that the unacceptability of the former is explained by a processing-based factor (which is essentially a form of garden path effect). We address this issue in detail in section 10.2.4.2.

In this section, we formulate an explicit compositional analysis that derives both (411a) and (411b) as syntactically well-formed sentences of English, with the same meaning (after the VP ellipsis in the latter is resolved appropriately). As we show below, the basic analysis of the syntax-semantics interface of comparatives and its interactions with VP ellipsis introduced in section 8.2.1 extends straightforwardly to these somewhat more complex attributive comparative data.

There is one technical complication that needs to be addressed when formulating a compositional analysis of attributive comparatives: in examples like those in (411), the comparative form is in the attributive modifier position of an existentially quantified (or indefinite) NP, and we need to make sure that the **max** operator introduced by the comparative operator scopes lower than this existential quantifier. For this purpose, the definition of the comparative operator needs to be made slightly more complex. We first illustrate the new definition of the comparative operator with the simple example (412) (= (383a) from section 8.2.2), showing that the new definition yields exactly the same truth conditions for this sentence as the older definition.

(412) Mary is taller than Ann is.

We then show that with this new definition of the comparative operator, the attributive comparative examples in (411) can be analyzed in a way fully parallel to the analysis of (412), with the only difference being that the adjective is in the prenominal attributive position rather than in the predicative position.

---

assumed here, or syntactic, as is more traditionally entertained) of accounting for the acceptability contrast between "supported" and "unsupported" subject parasitic gaps, then that account straightforwardly carries over to the VP ellipsis cases. In categorial grammar, if one wanted to encode the parasitic gap licensing patterns in the syntax, the category distinction between VP↾XP and VP already present would provide just enough information to implement such a syntactic condition.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

The new comparative operator has the same syntactic type and prosodic form as the second version of the comparative operator from section 8.2.2 in (388):

(413) $\lambda\sigma_0\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{ER}(\sigma_0(\boldsymbol{\varepsilon}))) \circ \mathsf{than} \circ \sigma_2(\boldsymbol{\varepsilon})$;
$\lambda f \lambda \mathscr{P} \lambda \mathscr{Q} \exists d_1, d_2.\mathscr{P}(\lambda x.\mathbf{max}(\lambda d.f(d)(x)) = d_1) \wedge$
$\mathscr{Q}(\lambda x.\mathbf{max}(\lambda d.f(d)(x)) = d_2) \wedge d_1 > d_2$;
$\mathrm{S}\restriction(\mathrm{S}\restriction\mathrm{Adj})\restriction(\mathrm{S}\restriction\mathrm{Adj})\restriction(\mathrm{Adj}\restriction\mathrm{Deg})$

The difference lies in the semantic component. Instead of directly forming degree predicates with the two clauses whose degree argument positions are abstracted over, the new operator first identifies the maximal degrees that satisfy the relevant degree descriptions and then compares the two degrees thus obtained. The crucial difference from the older definition in (388) is that the new definition forces the relevant degree descriptions (and hence the **max** operator that scopes immediately over these degree descriptions) to be included in the scope of other scopal expressions (if there are any) in the sentence.

Since the syntactic type of the comparative operator remains the same, the structure of the derivation for (412) is identical to (389) from section 8.2.2. The following translation is obtained by replacing the comparative operator in (389) with the new one in (413):

(414) $\exists d_1, d_2.\mathbf{max}(\lambda d.\mathbf{tall}(d)(\mathbf{m})) = d_1 \wedge \mathbf{max}(\lambda d.\mathbf{tall}(d)(\mathbf{a})) = d_2 \wedge d_1 > d_2$

This is semantically equivalent to the older translation except that there is (in this case) redundant existential quantification over the two degrees d_1 and d_2.

As noted above, the only (substantial) difference between the simpler example in (412) and the attributive comparative example in (411a) is that the gradable adjective is in the predicative position in the former, whereas it is in the prenominal attributive modifier position in the latter. We assume that prenominal adjectives are of type N/N (semantically, $et \to et$). The (polymorphic) definition of the comparative operator needs to be slightly adjusted to accommodate this type difference. The version of the comparative operator for the prenominal adjective is as follows:

(415) $\lambda\sigma_0\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{ER}(\sigma_0(\boldsymbol{\varepsilon}))) \circ \mathsf{than} \circ \sigma_2(\boldsymbol{\varepsilon})$;
$\lambda f \lambda \mathscr{P} \lambda \mathscr{Q} \exists d_1, d_2.\mathscr{P}(\lambda P \lambda x.\mathbf{max}(\lambda d.f(d)(P)(x)) = d_1) \wedge$
$\mathscr{Q}(\lambda P \lambda x.\mathbf{max}(\lambda d.f(d)(P)(x)) = d_2) \wedge d_1 > d_2$;
$\mathrm{S}\restriction(\mathrm{S}\restriction(\mathrm{N/N}))\restriction(\mathrm{S}\restriction(\mathrm{N/N}))\restriction((\mathrm{N/N})\restriction\mathrm{Deg})$

With this definition, the derivation for (411a) is straightforward. Just as in the simpler example in (412), the derivation proceeds by abstracting over the position of the adjective in the main clause and the *than* clause to form two clauses that have adjectival gaps (of type $\mathrm{S}\restriction(\mathrm{N/N})$). Two such gapped clauses are then given as argu-

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

ments to the comparative operator. The full derivation is given in (417) (here, $\mathbf{int} = \lambda d \lambda P \lambda x.P(x) \wedge \mathbf{interesting}_{d \to t \to t}(d)(x)$). The final translation obtained in (417) can be unpacked as in (416).

(416)
$$\begin{aligned}
&\exists d_1, d_2.\mathfrak{E}_{(\lambda x.\mathbf{max}(\lambda d.\mathbf{int}(d)(\mathbf{novel})(x))=d_1)}(\lambda x.\mathbf{wrote}(x)(\mathbf{j})) \wedge \\
&\quad \mathfrak{E}_{(\lambda x.\mathbf{max}(\lambda d.\mathbf{int}(d)(\mathbf{play})(x))=d_2)}(\lambda x.\mathbf{wrote}(x)(\mathbf{m})) \wedge d_1 > d_2 \\
&= \exists d_1, d_2.\exists x.[\mathbf{max}(\lambda d.\mathbf{novel}(x) \wedge \mathbf{interesting}(d)(x)) = d_1 \wedge \mathbf{wrote}(x)(\mathbf{j})] \\
&\quad \wedge \exists y.[\mathbf{max}(\lambda d.\mathbf{play}(y) \wedge \mathbf{interesting}(d)(y)) = d_1 \wedge \mathbf{wrote}(y)(\mathbf{m})] \\
&\quad \wedge d_1 > d_2
\end{aligned}$$

This says that there is a novel that John wrote and there is a play that Mary wrote and that the former is more interesting than the latter. This corresponds to the intuitive meaning of the sentence.

(417)

$$\dfrac{\dfrac{\dfrac{\begin{array}{l}\lambda\sigma_0\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{ER}(\sigma_0(\boldsymbol{\varepsilon}))) \circ \mathsf{than} \circ \sigma_2(\boldsymbol{\varepsilon}); \\ \lambda f \lambda \mathscr{P} \lambda \mathscr{Q} \exists d_1, d_2. \\ \quad \mathscr{P}(\lambda P \lambda x.\mathbf{max}(\lambda d.f(d)(P)(x)) = d_1) \wedge \\ \quad \mathscr{Q}(\lambda P \lambda x.\mathbf{max}(\lambda d.f(d)(P)(x)) = d_2) \wedge \\ \quad d_1 > d_2; \\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright(\mathrm{N/N}))\upharpoonright(\mathrm{S}\upharpoonright(\mathrm{N/N}))\upharpoonright((\mathrm{N/N})\upharpoonright\mathrm{Deg})\end{array} \quad \begin{array}{l}\lambda\varphi.\mathsf{interesting} \circ \\ \quad \varphi; \\ \mathbf{int}; \\ (\mathrm{N/N})\upharpoonright\mathrm{Deg}\end{array}}{\begin{array}{l}\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{ER}(\mathsf{interesting})) \circ \mathsf{than} \circ \sigma_2(\boldsymbol{\varepsilon}); \\ \lambda \mathscr{P} \lambda \mathscr{Q} \exists d_1, d_2.\mathscr{P}(\lambda P \lambda x.\mathbf{max}(\lambda d.\mathbf{int}(d)(P)(x)) = d_1) \wedge \\ \quad \mathscr{Q}(\lambda P \lambda x.\mathbf{max}(\lambda d.\mathbf{int}(d)(P)(x)) = d_2) \wedge d_1 > d_2; \\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright(\mathrm{N/N}))\upharpoonright(\mathrm{S}\upharpoonright(\mathrm{N/N}))\end{array}} \quad \begin{array}{l}\vdots \\ \lambda\varphi.\mathsf{john} \circ \\ \quad \mathsf{wrote} \circ \\ \quad \mathsf{a} \circ \varphi \circ \\ \quad \mathsf{novel}; \\ \lambda f.\mathfrak{E}_{f(\mathbf{novel})} \\ \quad (\lambda x.\mathbf{wrote} \\ \quad (x)(\mathbf{j})); \\ \mathrm{S}\upharpoonright(\mathrm{N/N})\end{array}}{\begin{array}{l}\lambda\sigma_2.\mathsf{john} \circ \mathsf{wrote} \circ \mathsf{a} \circ \mathsf{ER}(\mathsf{interesting}) \circ \mathsf{novel} \circ \mathsf{than} \circ \sigma_2(\boldsymbol{\varepsilon}); \\ \lambda \mathscr{Q} \exists d_1, d_2.\mathfrak{E}_{(\lambda x.\mathbf{max}(\lambda d.\mathbf{int}(d)(\mathbf{novel})(x))=d_1)}(\lambda x.\mathbf{wrote}(x)(\mathbf{j})) \wedge \\ \quad \mathscr{Q}(\lambda P \lambda x.\mathbf{max}(\lambda d.\mathbf{int}(d)(P)(x)) = d_2) \wedge d_1 > d_2; \\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright(\mathrm{N/N}))\end{array}} \quad \begin{array}{l}\vdots \\ \lambda\varphi.\mathsf{mary} \circ \\ \quad \mathsf{wrote} \circ \\ \quad \mathsf{a} \circ \varphi \circ \\ \quad \mathsf{play}; \\ \lambda f.\mathfrak{E}_{f(\mathbf{play})} \\ \quad (\lambda x.\mathbf{wrote} \\ \quad (x)(\mathbf{m})); \\ \mathrm{S}\upharpoonright(\mathrm{N/N})\end{array}}{\begin{array}{l}\mathsf{john} \circ \mathsf{wrote} \circ \mathsf{a} \circ \mathsf{ER}(\mathsf{interesting}) \circ \mathsf{novel} \circ \mathsf{than} \circ \mathsf{mary} \circ \mathsf{wrote} \circ \mathsf{a} \circ \mathsf{play}; \\ \exists d_1, d_2.\mathfrak{E}_{(\lambda x.\mathbf{max}(\lambda d.\mathbf{int}(d)(\mathbf{novel})(x))=d_1)}(\lambda x.\mathbf{wrote}(x)(\mathbf{j})) \wedge \\ \quad \mathfrak{E}_{(\lambda x.\mathbf{max}(\lambda d.\mathbf{int}(d)(\mathbf{play})(x))=d_2)}(\lambda x.\mathbf{wrote}(x)(\mathbf{m})) \wedge d_1 > d_2; \mathrm{S}\end{array}}$$

At this point, it should be obvious that (411b) is merely a pseudogapping counterpart of (411a), where the auxiliary *did* in the *than* clause stands in for the main verb *wrote*, with the syntactic type and semantic translation identical to the latter:

(418) did; **wrote**; VP/NP

The sign in (418) can be obtained straightforwardly in the analysis of pseudogapping from chapter 6. The derivation for (411b) can thus be obtained by replacing the lexical verb *wrote* in the *than* clause in (417) with the "pseudogapped" anaphoric auxiliary *did* in (418).

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

Thus, fully compositional analyses for the sentences in (411) can be obtained without recourse to deletion of abstract syntactic representations. Note in particular that our analysis consists solely of independently motivated assumptions about the relevant phenomena (the general syntax-semantics interface of comparatives and VP ellipsis/pseudogapping in the case of the "attributive comparative" construction analyzed in this section).

8.4 Apparent Challenge to the "Remnant-Extraction" Analysis

One issue needs to be addressed before concluding this chapter, in relation to our analysis of "extraction out of elided VPs" essentially as extraction *of* pseudogapping remnants. This type of analytic possibility has been considered previously in the literature. In particular, Johnson (2001) gives an informal sketch of a version of this analysis and rejects it by citing several examples which supposedly display non-parallelism between pseudogapping and extraction from VP ellipsis sites. We show here that the nonparallel behavior displayed by these examples is, contra Johnson, not a consequence of a systematic syntactic difference between the two constructions but rather a direct result of the nonparallelism in the way the specific examples have been constructed. Once these examples are corrected to eliminate the confounding factor that has led to the misinterpretation of the data, the differences between the extraction and pseudogapping cases that Johnson's claims rest on essentially disappear.

8.4.1 Johnson's Argument

The principal argument in Johnson (2001) is that pseudogapping obeys constraints on the ellipsis remnants not found in the apparent VP ellipsis + extraction cases; thus Johnson argues that pseudogapping "cannot elide part of a prepositional phrase . . . nor . . . remove part of a noun phrase," offering the examples in (419)–(420) to support this claim:

(419) a. *Sally will stand near Mag, but he won't Holly.
 b. *While Holly didn't discuss a report about every boy, she did every girl.

(420) I know which woman HOLLY will discuss a report about, but I don't know which woman YOU will.

Since pseudogapping, as in (419), apparently cannot "reach into" PPs to access their NP objects, whereas extraction from elided VP can, as in (420), Johnson argues that such extraction cannot be reduced to pseudogapping plus movement of the remnant.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

However, Johnson's argument is problematic since the alleged generalization that pseudogapping is limited to the direct object position of the verb is factually wrong: there is ample evidence from the literature that counterexemplifies this claim.[16]

8.4.2 Counterexamples from the Literature

There is ample data already attested in the literature displaying the possibility of pseudogapping remnants corresponding to NP-internal positions in the antecedent clause; see Miller (2014) and chapter 6 for many such examples. In particular, examples such as (421), first documented in Levin (1979), were already known long before Johnson (2001):

(421) You can take the lining out of that coat more easily than you can this one.

Examples such as those in (422) seem unproblematic as well.

(422) a. You'll find more illustrated books on golf than you will on ping-pong.
b. I've collected more somewhat embarrassing stories about John than I have about Bill.
c. As a fruitpacker, I can tell you that the key to high production is fruit size. For example, in one hour I can pack many more boxes of GRAPEFRUIT, especially those enormous Rio Reds, than I can ORANGES.

It is somewhat difficult to pin down exactly the conditions governing the felicity of these examples involving pseudogapping from "embedded" positions, but it seems uncontroversial that the relevant factor is pragmatic rather than syntactic. See section 6.3.5, where we have briefly speculated on the possibility of explaining the difference in acceptability among individual examples analogous to what is found in (421) and (422) versus (419) in terms of a complex set of conditions including factors such as prototypicality and inherent relatedness between the remnant and the elided head noun.

---

16. Johnson's argument has other serious empirical problems. In particular, he assumes the claims in Haïk (1987), who implicitly invokes a second kind of supposed contrast:

(i) *Mary talked about everyone that Peter did ∅ about __.

The point of (i), so far as Johnson's argument is concerned, is that, since this example supposedly shows that extraction from within a pseudogapping remnant is prohibited, it is difficult to justify a treatment of the elided VP extraction data which takes them to reflect extraction *of* such a remnant itself. But this argument too fails on factual grounds, for it is not difficult to find far better cases of extraction from pseudogapping "remnants":

(ii) a. I can predict/say who John will vote FOR __ more confidently than who he will AGAINST __.
b. (?)I can say who John will vote FOR __ more easily than I can who he will AGAINST __.

(iia) in particular seems quite acceptable.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

Note in this connection that examples that are syntactically identical in form to (419b) in relevant respects improve significantly by making the elided predicate semantically more natural. For example, *write a report about* is arguably a much more natural predicate than *discuss a report about*. As expected, (423), while not totally impeccable for all speakers, seems to be distinctly better than (419b):

(423) Given his background and experiences, I'd expect that John could write a report about EUROPE much more easily than he could CHINA.

Further corroboration for the current view comes from the fact that when *near* is replaced by *next to* in (419a), the example improves, particularly if the example also takes into account the strong preference for the very marked contrast in comparatives:

(424) I suspect Mary would sit next to John more readily than she would Bill.

This effect is particularly noticeable if, as suggested in (424), the discourse context is one in which people are making a conscious choice, in response to a request, of who they will sit in immediate proximity to. This fact suggests that pragmatic factors of the sort alluded to above indeed crucially affect the felicity of pseudogapping examples.[17]

8.5 A Note on the P-Stranding Generalization

We have focused on the interactions between extraction and VP ellipsis in our discussion, mostly responding to a particular set of claims made by Kennedy (2003) (and Kennedy and Merchant [2000]). The literature on ellipsis and extraction contains another type of argument, originally due to Merchant (2001), that has become quite influential. The argument (in its original form) involves the sluicing construction and is called the "P-stranding (preposition-stranding) generalization." The P-stranding generalization is often invoked as a strong piece of argument in favor of covert structural representations in ellipsis, though there are recent discussions casting considerable doubt on the validity of this empirical generalization.[18] In what follows, we demonstrate that

17. One might still wonder why a contrast exists in Johnson's original examples: while the same predicate *discuss a report about* is used in both (419b) and (420), the former is unacceptable while the latter is acceptable. We speculate that the improved status of (420) has to do with the fact that extraction is involved in this example. In particular, note that in (420), due to the extraction of the complement NP of the preposition *about*, the material that needs to be recovered in the ellipsis clause (i.e., *discuss a report about* __) is a surface syntactic constituent. This arguably would help greatly in the pragmatic ellipsis resolution process as compared to the case (such as (419b)) in which such a condition does not hold.

18. For example, Sag and Nykiel (2011) argue that Polish counterexemplifies this generalization; Merchant (2019) himself refers to the literature on several such cases and suggests the possibility that in these cases "the P-less 'sluices' in fact derive from a copular or reduced cleft-like source," a possibility which he refers to as pseudosluicing. Recent work, however, challenges this alternative derivation of apparent exception: Nykiel (2013) offers a suite of experimental psycholinguistic tests whose results suggest that in general the

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

even if the P-stranding generalization were a cross-linguistically valid generalization, there is a relatively straightforward way of capturing it in the type of approach to ellipsis phenomena we have advocated in this book that is arguably no more stipulative than the popular configurational analysis in the mainstream syntactic literature. Taken together with the recent reconsideration of the empirical status of the generalization itself, the "existence proof" we offer below for a nonstructural account of the P-stranding generalization removes the force of the argument for covert structure that the P-stranding generalization is typically associated with.

8.5.1 The P-Stranding Generalization

Merchant (2001, 92) contains an argument for covert structure in sluicing based on what has come to be known as the P-stranding generalization:

(425) A language *L* will allow preposition stranding under sluicing iff *L* allows preposition stranding under regular *wh*-movement.

The pattern noted in (425) can be illustrated by the contrast between English and Spanish:

(426) a. John was talking to someone, but I don't know who __.
b. *Juan ha hablado con una chica rubia, pero no sé qué chica más.
'John has talked with a blond girl, but I don't know which other girl (John has talked with).' (Rodrigues et al. 2009)

English permits preposition stranding; Spanish does not. This pattern appears to persist in sluicing constructions where the ellipsed material, corresponding to the portion of the antecedent presumably left behind by *wh* movement, contains a preposition stranded by such movement. The empirical basis for the generality of Merchant's claim is a set of data from eighteen languages: six Germanic languages which support P-stranding (both in the regular *wh* movement context and in sluicing) and twelve from a wide variety of languages (including a variety of Indo-Iranian and Semitic languages, as well as Basque) which do not. On the basis of these facts and the descriptive generalization in (425), Merchant (2001, 107) argues that

> the usual mechanisms for case-assignment and determination of targets of *wh*-movement that operate in a given language to regulate the shapes of *wh*-phrases in non-elliptical questions operate in identical ways under sluicing as well. All of these facts strongly suggest that *wh*-movement of the usual sort has taken place, displacing an IP-internal *wh*-phrase to SpecCP. . . . Similar considerations suggest a movement approach to a variety of parallel . . . form-identity effects in stripping, comparatives, fragment answers, [and] the remnants of gapping, which often show case and P-stranding dependencies like their sluicing cousins.

---

acceptability of sluicing with a lone preposition remnant is unrelated to the acceptability of *wh* extraction of NPs from PPs.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

By the same reasoning, the putative generalization should be observable in examples like the following, where (427a) is standardly analyzed as extraction out of elided VPs and (427b) is a case of pseudogapping:

(427) a. I know whom John argued with, but I don't know whom Mary did.
b. I can deal with Mary more easily than I can Sue.

In (427a) the *wh* filler has moved to the left, leaving behind a VP with a trace in it (*argued with* __), which is subsequently deleted on Merchant's analysis. In a language that allows extraction out of VP ellipsis sites but which has a ban on P-stranding, data such as (427a) should never exist, and likewise for analogues of (427b).[19]

8.5.2 English*

For the purpose of illustration we now consider a language English* (instead of a real language such as German), which is exactly like English except that preposition stranding is forbidden. We demonstrate how our general approach to ellipsis, with no additions or modifications, yields the P-stranding generalization for the so-called extraction out of elided VP cases like (427a) (for which we have defended an analysis [essentially] in terms of extraction of pseudogapping remnants). Assume that in English* we have only (428) as the lexical entry for the preposition *to*.

(428) *to*; $\lambda x.x$; PP_{to}/NP_{-wh}

This specification differs crucially from the lexical description of *to* in English in that the syntactic type of the latter will be $PP_{to}/NP_{\#wh}$ (see below), with both NP_{+wh} and NP_{-wh} as possible instantiations for $NP_{\#wh}$.

Given the lexical specifications for prepositions in English*, indirect questions such as *I wonder who John spoke to* cannot be formed. Since only an NP_{-wh} variable can be supplied to the object of *to*, a sentence missing an NP object of a preposition such as *John spoke to* __ can only be derived in type $S{\upharpoonright}NP_{-wh}$. But such a description fails to satisfy the fronted *wh* word's argument description $S{\upharpoonright}NP_{+wh}$ (cf. (344)). And precisely the same will hold in the attempt to derive the corresponding elided VP. Consider (429):

(429) JOHN talked to BILL, but I don't know whom MARY did.

This example is acceptable in English (as long as the proper intonational and contextual cues are given) but would be ill-formed in English*. The antecedent upon which the ellipsis clause depends will have the following derivation:

19. Matters may be a bit more complex in the case of pseudogapping, though. On some analyses, remnant movement in pseudogapping is to the right, leaving open the possibility that in a language with a P-stranding prohibition on leftward movement only, something like (427b) could be legal.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

(430)

$$\dfrac{\begin{bmatrix}\varphi_0;\\ u;\\ \mathrm{NP}_{-wh}\end{bmatrix}^1 \quad \begin{array}{l}\mathsf{to};\\ \lambda\varphi.\varphi;\\ \mathrm{PP}_{to}/\mathrm{NP}_{-wh}\end{array}}{\mathsf{to}\circ\varphi_0;\ u;\ \mathrm{PP}_{to}} \qquad \begin{array}{l}\mathsf{talked};\\ \mathbf{talk};\\ \mathrm{VP}/\mathrm{PP}_{to}\end{array}$$

$$\dfrac{\dfrac{\mathsf{talked}\circ\mathsf{to}\circ\varphi_0;\ \mathbf{talk}(u);\ \mathrm{VP}}{\lambda\varphi_0.\mathsf{talked}\circ\mathsf{to}\circ\varphi_0;\ \lambda u.\mathbf{talk}(u);\ \mathrm{VP}\restriction\mathrm{NP}_{-wh}}{\restriction}\mathrm{I}^1 \quad \begin{bmatrix}\varphi_1;\\ w;\\ \mathrm{NP}_{-wh}\end{bmatrix}^2}{\mathsf{talked}\circ\mathsf{to}\circ\varphi_1;\ \mathbf{talk}(w);\ \mathrm{VP}} \qquad \begin{array}{l}\mathsf{john};\\ \mathbf{j};\\ \mathrm{NP}_{-wh}\end{array}$$

$$\dfrac{\dfrac{\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\varphi_1;\ \mathbf{talk}(w)(\mathbf{j});\ \mathrm{S}}{\lambda\varphi_1.\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\varphi_1;\ \lambda w.\mathbf{talk}(w)(\mathbf{j});\ \mathrm{S}\restriction\mathrm{NP}_{-wh}}{\restriction}\mathrm{I}^2 \quad \begin{array}{l}\mathsf{bill};\\ \mathbf{b};\\ \mathrm{NP}_{-wh}\end{array}}{\mathsf{john}\circ\mathsf{talked}\circ\mathsf{to}\circ\mathsf{bill};\ \mathbf{talk}(\mathbf{b})(\mathbf{j});\ \mathrm{S}}{\restriction}\mathrm{E}$$

(The expression $\lambda\varphi_0.\mathsf{talked}\circ\mathsf{to}\circ\varphi_0;\ \lambda u.\mathbf{talk}(u);\ \mathrm{VP}\restriction\mathrm{NP}_{-wh}$ is grayed in.)

And for the VP ellipsis auxiliary, we obtain:

(431)

$$\begin{array}{l}\lambda\rho\lambda\varphi.\rho(\lambda\varphi_0.\varphi_0)(\varphi);\\ \lambda\mathscr{F}.\mathscr{F}(P);\\ (\mathrm{VP}\restriction\mathrm{NP}_{-wh})\restriction\\ \quad((\mathrm{VP}\restriction\mathrm{NP}_{-wh})\restriction(\mathrm{VP}\restriction\mathrm{NP}_{-wh}))\end{array}$$

$$\textcircled{1}\rightarrow \cdots\cdots\cdots$$

$$\dfrac{\begin{array}{c}\vdots\\ \lambda\sigma\lambda\varphi.\mathsf{did}\circ\sigma(\varphi);\\ \lambda f\lambda x\lambda y.f(x)(y);\\ (\mathrm{VP}\restriction\mathrm{NP}_{-wh})\restriction\\ \quad(\mathrm{VP}\restriction\mathrm{NP}_{-wh})\end{array} \quad \begin{array}{l}\lambda\rho\lambda\varphi.\rho(\lambda\varphi_0.\varphi_0)(\varphi);\\ \lambda\mathscr{F}.\mathscr{F}(\lambda u.\mathbf{talked}(u));\\ (\mathrm{VP}\restriction\mathrm{NP}_{-wh})\restriction\\ \quad((\mathrm{VP}\restriction\mathrm{NP}_{-wh})\restriction(\mathrm{VP}\restriction\mathrm{NP}_{-wh}))\end{array}}{\lambda\varphi.did\circ\varphi;\ \lambda x\lambda y.\mathbf{talk}(x)(y);\ \mathrm{VP}\restriction\mathrm{NP}_{-wh}} \quad \begin{bmatrix}\varphi_3;\\ v;\\ \mathrm{NP}_{-wh}\end{bmatrix}^3$$

$$\dfrac{\dfrac{\mathsf{did}\circ\varphi_3;\ \lambda y.\mathbf{talked}(v)(y);\ \mathrm{VP} \quad \begin{array}{l}\mathsf{mary};\\ \mathbf{m};\\ \mathrm{NP}_{-wh}\end{array}}{\mathsf{mary}\circ\mathsf{did}\circ\varphi_3;\ \mathbf{talk}(v)(\mathbf{m});\ \mathrm{S}}}{\lambda\varphi_3.\mathsf{mary}\circ\mathsf{did}\circ\varphi_3;\ \lambda v.\mathbf{talk}(v)(\mathbf{m});\ \mathrm{S}\restriction\mathrm{NP}_{-wh}}{\restriction}\mathrm{I}^3$$

The ellipsis operator has a syntactic type schematically of the form $X\restriction(X\restriction X)$, and the anaphora resolution condition requires that X match the category of the relevant antecedent whose meaning is recovered in ellipsis resolution (see chapter 6 for details). In the case at hand, the appropriate antecedent is the grayed-in expression in (430), with semantics $\lambda u.\mathbf{talked}(u)$. The free variable P in the ellipsis operator thus gets resolved as this term at the step (which, strictly speaking, is outside of the syntactic derivation) marked as ①. But then, any attempt to compose the sign derived in (431) with the extraction operator will fail:

(432)

$$\dfrac{\begin{array}{c}\vdots\\ \lambda\varphi_3.\mathsf{mary}\circ\mathsf{did}\circ\varphi_3;\ \lambda v.\mathbf{talk}(v)(\mathbf{m});\ \mathrm{S}\restriction\mathrm{NP}_{-wh}\end{array} \quad \begin{array}{l}\lambda\sigma.\mathsf{whom}\circ\sigma(\boldsymbol{\varepsilon});\\ \mathbf{wh}(\mathbf{person});\mathrm{Q}\restriction(\mathrm{S}\restriction\mathrm{NP}_{+wh})\end{array}}{\textbf{FAIL}}$$

In a nutshell, the extraction operator can only compose with a sentence missing an NP_{+wh}, but the conditions imposed on prepositions in English* allow *to* to combine

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

only with NP_{-wh}, leading to a continuation typed $\mathrm{S}\upharpoonright\mathrm{NP}_{-wh}$, an invalid argument for the extraction operator. Such extractions therefore cannot give rise to well-formed VP ellipsis strandings. No special mechanisms are required, and nothing has to be stipulated other than the lexical condition illustrated in (428), which simply expresses the ban on preposition stranding in such languages. In particular, no covert syntactic structure is required as long as the required syntactic information about category type is made available to the anaphoric process.

In English, in contrast, sentences such as (429) are licensable because in place of (428), the lexical entry for *to* is (433):

(433) $to;\ \lambda x.X;\ \mathrm{PP}_{to}/\mathrm{NP}_{\#wh}$

Since *to* can combine with an NP of either polarity for the *wh* feature, we will have a derivation along the following lines:

(434)

$$\dfrac{\begin{bmatrix}\varphi_0;\\ u;\\ \mathrm{NP}_{\#wh}\end{bmatrix}^1 \quad \begin{array}{l}\textsf{to};\\ \lambda\varphi.\varphi;\\ \mathrm{PP}_{to}/\mathrm{NP}_{\#wh}\end{array}}{\textsf{to}\circ\varphi_0;\ u;\ \mathrm{PP}_{to}}$$

$$\dfrac{\begin{array}{l}\textsf{talked};\\ \textbf{talk};\\ \mathrm{VP}/\mathrm{PP}_{to}\end{array} \quad \textsf{to}\circ\varphi_0;\ u;\ \mathrm{PP}_{to}}{\textsf{talked}\circ\textsf{to}\circ\varphi_0;\ \textbf{talk}(u);\ \mathrm{VP}}$$

$$\dfrac{\textsf{talked}\circ\textsf{to}\circ\varphi_0;\ \textbf{talk}(u);\ \mathrm{VP}}{\boxed{\begin{array}{l}\lambda\varphi_0.\textsf{talked}\circ\textsf{to}\circ\varphi_0;\\ \lambda u.\textbf{talk}(u);\mathrm{VP}\upharpoonright\mathrm{NP}_{\#wh}\end{array}}}\upharpoonright\mathrm{I}^1$$

$$\dfrac{\begin{array}{l}\lambda\varphi_0.\textsf{talked}\circ\textsf{to}\circ\varphi_0;\\ \lambda u.\textbf{talk}(u);\mathrm{VP}\upharpoonright\mathrm{NP}_{\#wh}\end{array} \quad \begin{bmatrix}\varphi_1;\\ w;\\ \mathrm{NP}_{\#wh}\end{bmatrix}^2}{\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \textbf{talk}(w);\ \mathrm{VP}}$$

$$\dfrac{\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \textbf{talk}(w);\ \mathrm{VP} \quad \begin{array}{l}\textsf{john};\\ \mathbf{j};\\ \mathrm{NP}_{-wh}\end{array}}{\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \textbf{talk}(w)(\mathbf{j});\ \mathrm{S}}$$

$$\dfrac{\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \textbf{talk}(w)(\mathbf{j});\ \mathrm{S}}{\lambda\varphi_1.\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \lambda w.\textbf{talk}(w)(\mathbf{j});\ \mathrm{S}\upharpoonright\mathrm{NP}_{\#wh}}\upharpoonright\mathrm{I}^2$$

$$\dfrac{\lambda\varphi_1.\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \lambda w.\textbf{talk}(w)(\mathbf{j});\ \mathrm{S}\upharpoonright\mathrm{NP}_{\#wh} \quad \begin{array}{l}\textsf{bill};\\ \mathbf{b};\\ \mathrm{NP}_{-wh}\end{array}}{\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\textsf{bill};\ \textbf{talk}(\mathbf{b})(\mathbf{j});\ \mathrm{S}}$$

The grayed-in line is the critical proof step for anaphoric retrieval. The argument of this expression is $\mathrm{NP}_{\#wh}$. So, we can derive an auxiliary in the entry $\mathrm{VP}\upharpoonright\mathrm{NP}_{\#wh}$ by anaphorically retrieving its meaning from this antecedent as follows:

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

(435)

$$
\dfrac{
  \dfrac{
    \dfrac{
      \dfrac{
        \begin{array}{c}\vdots\\ \lambda\sigma\lambda\varphi.\mathsf{did}\circ\sigma(\varphi);\\ \lambda f\lambda x\lambda y.f(x)(y);\\ (\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}){\upharpoonright}\\ \quad(\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh})\end{array}
        \quad
        \begin{array}{c}\lambda\rho\lambda\varphi.\rho(\lambda\varphi_0.\varphi_0)(\varphi);\\ \lambda\mathscr{F}.\mathscr{F}(P);\\ (\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}){\upharpoonright}\\ \quad((\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}){\upharpoonright}(\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}))\\ \text{①}\rightarrow\cdots\cdots\cdots\cdots\cdots\cdots\\ \lambda\rho\lambda\varphi.\rho(\lambda\varphi_0.\varphi_0)(\varphi);\\ \lambda\mathscr{F}.\mathscr{F}(\lambda u.\mathbf{talked}(u));\\ (\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}){\upharpoonright}\\ \quad((\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}){\upharpoonright}(\mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}))\end{array}
      }{
        \lambda\varphi.\mathsf{did}\circ\varphi;\ \lambda x\lambda y.\mathbf{talk}(x)(y);\ \mathrm{VP}{\upharpoonright}\mathrm{NP}_{\#wh}
      }
      \quad
      \begin{bmatrix}\varphi_2;\\ w;\\ \mathrm{NP}_{\#wh}\end{bmatrix}^2
    }{
      \mathsf{did}\circ\varphi_2;\ \mathbf{talk}(w);\ \mathrm{VP}
    }
    \quad
    \begin{array}{l}\mathsf{mary};\\ \mathbf{m};\\ \mathrm{NP}_{-wh}\end{array}
  }{
    \dfrac{\mathsf{mary}\circ\mathsf{did}\circ\varphi_2;\ \mathbf{talk}(w)(\mathbf{m});\ \mathrm{S}}{\lambda\varphi_2.\mathsf{mary}\circ\mathsf{did}\circ\varphi_2;\ \lambda w.\mathbf{talk}(w)(\mathbf{m});\ \mathrm{S}{\upharpoonright}\mathrm{NP}_{\#wh}}{\upharpoonright}\mathrm{I}^2
  }
}{
  \dfrac{\vdots}{\lambda\varphi_2.\mathsf{mary}\circ\mathsf{did}\circ\varphi_2;\ \lambda w.\mathbf{talk}(w)(\mathbf{m});\ \mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}}
}
$$

The difference between the failed derivation in English* and the unproblematic derivation in English is a direct reflection of the different valence possibilities for prepositions in the two languages, with no reference to any structure ever coming into the picture. As the proof for English* makes clear, the enforced selection of NP_{-wh} by prepositions translates, through the chain of hypothetical reasoning displayed above, into a sign typed $\mathrm{S}{\upharpoonright}\mathrm{NP}_{-wh}$, which is unable to compose with the extraction operator to complete the derivation. Since the $\pm wh$ distinction is an independently motivated distinction that needs to be encoded in the syntactic type of NPs, the valence information is all that is needed to institute the P-stranding generalization for the "extraction out of VP ellipsis" pattern. We now demonstrate that exactly the same purely valence-based account is sufficient for the P-stranding generalization in the case of sluicing as well.

8.5.3 Sluicing

Following the overall strategy in Barker (2013), we posit the following operator for the analysis of sluicing:

(436) $\lambda\rho.\rho(\lambda\varphi.\varphi);\ \lambda\mathscr{W}.\mathscr{W}(P);\ \mathrm{Q}{\upharpoonright}(\mathrm{Q}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}))$
——where P is a property matching a contextually salient sign compatible with the type description $\mathrm{S}{\upharpoonright}\mathrm{NP}_{+wh}$ in the preceding discourse

In simple cases of sluicing, we obtain derivations such as that given in (438) for (437):

(437) John criticized someone, but Mary doesn't know who(m).

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

(438)

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\text{criticized}; \mathbf{criticize}; \mathrm{VP}/\mathrm{NP}_{\#wh} \quad [\varphi_0; x; \mathrm{NP}_{\#wh}]^1}{\text{criticized} \circ \varphi_0; \mathbf{criticize}(x); \mathrm{VP}} \quad \begin{array}{l}\text{john};\\ \mathbf{j}; \mathrm{NP}_{-wh}\end{array}}{\text{john} \circ \text{criticized} \circ \varphi_0; \mathbf{criticize}(x)(\mathbf{j}); \mathrm{S}}\upharpoonright\mathrm{I}^1}{\lambda\varphi_0.\text{john} \circ \text{criticized} \circ \varphi_0; \lambda x.\mathbf{criticize}(x)(\mathbf{j}); \mathrm{S}\upharpoonright\mathrm{NP}_{\#wh}}}{\vdots}}{\lambda\varphi_0.\text{john} \circ \text{criticized} \circ \varphi_0; \lambda x.\mathbf{criticize}(x)(\mathbf{j}); \mathrm{S}\upharpoonright\mathrm{NP}_{-wh}} \quad \begin{array}{l}\lambda\sigma_0.\sigma_0(\text{someone});\\ \exists_{\mathbf{person}};\\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{-wh})\end{array}}{\text{john} \circ \text{criticized} \circ \text{someone}; \exists_{\mathbf{person}}(\lambda x.\mathbf{criticize}(x)(\mathbf{j})); \mathrm{S}}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{l}\lambda\rho.\rho(\lambda\varphi.\varphi);\\ \lambda\mathscr{W}.\mathscr{W}(P); \mathrm{Q}\upharpoonright(\mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh}))\end{array}}{\begin{array}{l}\lambda\rho.\rho(\lambda\varphi.\varphi);\\ \lambda\mathscr{W}.\mathscr{W}(\lambda x.\mathbf{criticize}(x)(\mathbf{j}));\\ \mathrm{Q}\upharpoonright(\mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh}))\end{array}} \quad \begin{array}{l}\lambda\sigma_1.\\ \quad\text{whom} \circ \sigma_1(\boldsymbol{\varepsilon});\\ \mathbf{wh}(\mathbf{person});\\ \mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh})\end{array}}{\begin{array}{l}\text{whom};\\ \mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j})); \mathrm{Q}\end{array}} \quad \begin{array}{l}\text{know};\\ \mathbf{know};\\ \mathrm{VP}/\mathrm{Q}\end{array}}{\begin{array}{l}\text{know} \circ \text{whom};\\ \mathbf{know}(\mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j}))); \mathrm{VP}\end{array}} \quad \begin{array}{l}\text{doesn't};\\ \lambda Q\lambda y.\\ \quad\neg Q.Q(y);\\ \mathrm{VP}/\mathrm{VP}\end{array}}{\begin{array}{l}\text{doesn't} \circ \text{know} \circ \text{whom};\\ \lambda y.\neg\mathbf{know}(\mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j})))(y); \mathrm{VP}\end{array}} \quad \begin{array}{l}\text{mary};\\ \mathbf{m};\\ \mathrm{NP}_{-wh}\end{array}}{\begin{array}{l}\text{mary} \circ \text{doesn't} \circ \text{know} \circ \text{whom};\\ \neg\mathbf{know}(\mathbf{wh}(\mathbf{person})(\lambda x.\mathbf{criticize}(x)(\mathbf{j})))(\mathbf{m}); \mathrm{S}\end{array}}$$

The lexical entry for the *wh* word *whom* and the sluicing operator (436) are taken to be common to English* and English and, more generally, to all languages with *wh* extraction and sluicing regardless of whether they allow stranded prepositions. Again, the sole difference is in the specification of the class of NPs with which prepositions can combine.

Given the entry for *to* in (428), preposition stranding is already automatically blocked in English* sluicing, just as it is for VP ellipsis extraction. For example, consider the following example:

(439) John talked to someone, but I don't know who(m).

In order to create the appropriate expression to serve as the antecedent in the first clause, *to* needs to combines with a variable, but in view of (428), this variable will necessarily be $-wh$. When the hypothesis corresponding to this variable is withdrawn, the result, $\mathrm{S}\upharpoonright\mathrm{NP}_{-wh}$, will be unable to serve as an antecedent for the sluicing operator, which explicitly requires the antecedent to be compatible with type $\mathrm{S}\upharpoonright\mathrm{NP}_{+wh}$.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

(440)

$$\dfrac{\lambda\sigma_1.\sigma_1(\textsf{someone});\ \exists_{\mathbf{person}};\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{-wh}) \qquad \dfrac{\dfrac{\dfrac{\textsf{talked};\ \mathbf{talk};\ \mathrm{VP}/\mathrm{PP}_{to} \qquad \dfrac{\textsf{to};\ \lambda\varphi.\varphi;\ \mathrm{PP}_{to}/\mathrm{NP}_{-wh} \qquad \left[\varphi_1;\ v;\ \mathrm{NP}_{-wh}\right]^1}{\textsf{to}\circ\varphi_1;\ v;\ \mathrm{PP}}}{\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \mathbf{talk}(v);\ \mathrm{VP}} \qquad \textsf{john};\ \mathbf{j};\ \mathrm{NP}}{\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \mathbf{talk}(v)(\mathbf{j});\ \mathrm{S}}\upharpoonright\mathrm{I}^1}{\lambda\varphi_1.\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \lambda v.\mathbf{talk}(v)(\mathbf{j});\ \mathrm{S}\upharpoonright\mathrm{NP}_{-wh}}}{\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\textsf{someone};\ \exists_{\mathbf{person}}(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\ \mathrm{S}}$$

$$\dfrac{\dfrac{\lambda\rho.\rho(\lambda\varphi.\varphi);\ \lambda\mathscr{W}.\mathscr{W}(P);\ \mathrm{Q}\upharpoonright(\mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh}))}{\lambda\rho.\rho(\lambda\varphi.\varphi);\ \lambda\mathscr{W}.\mathscr{W}(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\ \mathrm{Q}\upharpoonright(\mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh}))}\ \mathbf{FAIL} \qquad \lambda\sigma.\textsf{who}\circ\sigma(\boldsymbol{\varepsilon});\ \mathbf{wh}(\mathbf{person});\ \mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh})}{\textsf{who};\ \mathbf{wh}(\mathbf{person})(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\ \mathrm{Q}}$$

By contrast, in English, since the argument of *to* is underspecified for the *wh* feature, the following derivation is available:

(441)

$$\dfrac{\lambda\sigma_1.\sigma_1(\textsf{someone});\ \exists_{\mathbf{person}};\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{-wh}) \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\textsf{talked};\ \mathbf{talk};\ \mathrm{VP}/\mathrm{PP}_{to} \qquad \dfrac{\textsf{to};\ \lambda\varphi.\varphi;\ \mathrm{PP}_{to}/\mathrm{NP}_{\#wh} \quad [\varphi_1;v;\mathrm{NP}_{\#wh}]^1}{\textsf{to}\circ\varphi_1;\ v;\ \mathrm{PP}}}{\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \mathbf{talk}(v);\ \mathrm{VP}} \qquad \textsf{john};\ \mathbf{j};\ \mathrm{NP}_{-wh}}{\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \mathbf{talk}(v)(\mathbf{j});\ \mathrm{S}}\upharpoonright\mathrm{I}^1}{\lambda\varphi_1.\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \lambda v.\mathbf{talk}(v)(\mathbf{j});\ \mathrm{S}\upharpoonright\mathrm{NP}_{\#wh}}}{\vdots}}{\lambda\varphi_1.\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\varphi_1;\ \lambda v.\mathbf{talk}(v)(\mathbf{j});\ \mathrm{S}\upharpoonright\mathrm{NP}_{-wh}}}{\textsf{john}\circ\textsf{talked}\circ\textsf{to}\circ\textsf{someone};\ \exists_{\mathbf{person}}(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\ \mathrm{S}}$$

$$\dfrac{\dfrac{\lambda\rho.\rho(\lambda\varphi.\varphi);\ \lambda\mathscr{W}.\mathscr{W}(P);\mathrm{Q}\upharpoonright(\mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh}))}{\lambda\rho.\rho(\lambda\varphi.\varphi);\ \lambda\mathscr{W}.\mathscr{W}(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\mathrm{Q}\upharpoonright(\mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh}))} \qquad \lambda\sigma.\textsf{whom}\circ\sigma(\boldsymbol{\varepsilon});\ \mathbf{wh}(\mathbf{person});\ \mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP}_{+wh})}{\textsf{whom};\ \mathbf{wh}(\mathbf{person})(\lambda v.\mathbf{talk}(v)(\mathbf{j}));\ \mathrm{Q}}$$

Here, the free variable P in the sluicing operator can be instantiated as a contextually appropriate predicate $\lambda v.\mathbf{talk}(v)(\mathbf{j})$ denoted by the grayed-in expression of type $\mathrm{S}\upharpoonright\mathrm{NP}$ in the antecedent clause, since $\mathrm{S}\upharpoonright\mathrm{NP}_{\#wh}$, which entails $\mathrm{S}\upharpoonright\mathrm{NP}_{+wh}$ (that is, $\mathrm{S}\upharpoonright\mathrm{NP}_{\#wh} \vdash$

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

$S{\upharpoonright}NP_{+wh}$ is a theorem), is clearly compatible with the description $S{\upharpoonright}NP_{+wh}$. It then follows that in English, preposition stranding is possible in sluicing.

Hence the P-stranding generalization in the case of sluicing falls out directly from the simple lexical treatment in (428) and (436), with no need to posit covert configurations.

The technical details of the proofs given above are necessary so that readers can verify for themselves that our proposal does exactly what we claim it does: guarantee that the P-stranding generalization does indeed fall out of our type-logical framework with no appeal whatever to configurational representations characterizing the "missing" material in any ellipsis construction. We hope, however, that the fundamental simplicity of our solution will not become lost in these technical details. The central point is that in order to capture the P-stranding generalization, nothing more need be assumed than the independently needed lexical prohibition on NP_{+wh} arguments to prepositions in non-stranding languages and the independently motivated assumption that the anaphora recovery process in ellipsis is sensitive to the syntactic category of the antecedent expression (cf. chapter 6).

Importantly, access to syntactic category information that our anaphora-based approach crucially exploits is not something that is "added on" to the theory but is a direct consequence of the fundamental architecture of (most versions of) categorial grammar: at each stage of syntactic derivation, the prosody, semantics, and syntactic type of the linguistic expression are fully explicit. The correlation between a particular semantics and specific syntactic type is thus built into the fundamental architecture of the theory. This architecture, however, does not allow access to the "history of derivation" (i.e., the structure of the proof) up to the point that the expression in question is obtained, and it is in this respect that categorial grammar departs most crucially from derivational variants of syntax that in principle allow (unless additional theory-internal assumptions are made) full access to the internal syntactic structure of a linguistic expression.

Both Hybrid TLCG and the P&P analyses of ellipsis essentially rely on specifications of lexical valence to rule out overgeneration that would arise in purely interpretive accounts, and in this respect, at the descriptive level, both are getting at more or less the same insight. However, covert structure analyses add a further component of hierarchical representations projected from these lexically specified argument structure possibilities—representations which, given the foregoing discussion, are not necessary to capture the P-stranding generalization. Basic considerations of parsimony (i.e., Occam's razor) thus seem to rule in favor of the Hybrid TLCG account (unless it can be shown that this approach incurs some hidden or overlooked additional complexity that is not present in the derivational approach) and the view that syntactic information as reflected in the syntactic categories of linguistic expressions is sufficient in the licensing of elliptical constructions, without the need for hidden configurational structure.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022

8.6 Conclusions: Covert Structure and the Burden of Proof

The overall architecture of the transformational framework, despite its successive recasting in at least seemingly quite different versions over the past seven decades, incorporates a syntax-semantics interface which most naturally handles form/meaning discrepancies in terms of hidden structures which contribute crucial components of the required semantic interpretations but which are subsequently suppressed by deletion operations of one or another sort. The invisibility of this deleted material inherently puts a burden of proof on any covert structural analysis: in any theory, the presence of a syntactic configuration inevitably entails certain empirical predictions as the null hypothesis, and the burden of proof is satisfied by demonstrating that, all other things being equal, constructions derived by movement and deletion conform to these predictions. The clear intent of Kennedy (2003) is to make an overwhelming demonstration along these lines.

But there are two deep, interlocking problems with all such attempts. In the first place, such arguments are only effective cross-theoretically if there is broad consensus that the constraints themselves cannot be explained satisfactorily except by appeal to syntactic structure, and for none of the probes for structure invoked by Kennedy, as we have discussed in detail earlier, does such a consensus exist: islands are increasingly widely assumed to represent functional obstacles not native to the combinatorics themselves; Condition B effects are cogently argued by Jacobson (2007) to correspond to semantic irreflexivity; Strong Crossover effects do not appear to have a clear explanation at all. Even more problematic, a number of the key factual claims in Kennedy and Merchant (2000), Kennedy (2003), and some of the important prior literature on which these claims implicitly rest are robustly counterexemplified in our informants' data, in corpora, or in earlier work on pseudogapping and ellipsis generally. It seems fair to conclude, therefore, that none of the arguments in Kennedy (2003) privilege a covert structure treatment of ellipsis over the kind of direct interpretation approach we have argued for in this book.

Downloaded from http://direct.mit.edu/books/book/chapter-pdf/677089/9780262360807_c000700.pdf by guest on 10 August 2022