Ellipsis and syntactic information

Yusuke Kubota
University of Tsukuba

Robert Levine
Ohio State University

August 28, 2017

1 Narrative outline

The following discussion 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 devoted to 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:

(i) the distribution of ellipsis remnants reflects sensitivity to island conditions on configurations which are not visible in surface string at the ellipsis site;

(ii) the interpretations available in ellipsis data comply with Strong Crossover restrictions that imply the existence of syntactic gaps within the ellipsis site;

(iii) the anaphora possibilities available in the interpretations of VP ellipsis data reflect Binding Condition B effects, implying the existence of pronouns within the ellipsis site;

(iv) 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; and

(v) the 'attributive comparative' construction permits certain possibilities which appear to violate the 'Left Branch Constraint' just in case ellipsis is also involved, a pattern inexplicable on a purely interpretive approach but one 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 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.

2 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 to them in Sect. 4 for a reassessment of how strong they actually are.

2.1 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 (1) and (2):

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

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

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

Extraction in the 'full' version of (1a) 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 (1b). 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 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 (2), 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.

2.2 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

[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 amongst researchers advocating covert structure in ellipsis about these difficulties.

appearance of a pronoun on a filler/gap pathway which is coreferential with, and in some sense structurally superior to, the gap. Thus, (3) does not allow the interpretation 'for which male person is it the case that that person is always criticizing himself?':

(3) *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 (4):

(4) 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 (4).

2.3 Ellipsis and Binding Condition B

Kennedy's argument from Binding Condition B hinges on data such as (5):

(5) John$_k$ takes care of him$_i$ because he$_{j/*i}$ won't $\left\{\begin{array}{l}\text{a. take care of him}_i\\ \text{b. }\cancel{\text{take care of him}_i}\end{array}\right\}$.

Condition B, first stated in Chomsky (1981) blocks co-reference 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, (5a): when the subject of the adjunct clause is interpreted as coreferential with the direct object, the example is bad. On the assumption that (5b) 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 $\lambda y.\textbf{take-care-of}(y)(\mathbf{k})$ cannot take $\mathbf{k}$ as an argument (with $\mathbf{k}$ 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 (5b) on the

[2] Note that if we were to replace the trace in (4) with a pronoun, the result in (i) will still be bad, due to Condition B effects (see Sect. 2.3):

(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 (4), on the assumption that a trace is present at the gap site.

coreferential interpretation, there must be an actual pronoun in the representation for the sentence at the level at which Condition B applies.[3]

2.4 Ellipsis and parasitic gaps

Parasitic gap licensing is standardly taken to be a matter of syntax. (6) illustrates the basis for the common claim that subjects cannot host gaps unless at least one gap also appears in the main VP.

(6) a. *Who$_i$ did close friends of $\_\__i$ become famous?
b. Who$_i$ do close friends of $\_\__i$ always defend $\_\__i$?

The examples in (7) exhibit a slightly more complex pattern. Standard islandhood tests seem to confirm that the gap in the adjunct in (7) is derived by *wh* extraction. As (7a) 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 the pattern in (6).

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

In contrast, in (7b), 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 (7b), the result is bad. Since however (7b) (in its gap version) is good, we can therefore 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 (8), a true gap must somehow be located within the ellipsis site:

(8) Otis is a person who$_i$ I admire $\_\__i$ because close friends of $\_\__i$ seem to $\varnothing$.
(Engdahl 1985; Kennedy 2003)

That is, $\varnothing$ must be $\cancel{[_{VP}\text{ admire } \_\__i]}$ entailing a covert structure analysis for such examples.

2.5 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 (9a), are arguably best analyzed as instances of *wh* movement involving a null operator, along the lines of (9b):

[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.

(9) a. Pico's novel was much more interesting than Brio thought it would be __.
b. Pico's novel was much more interesting than [$\mathsf{O}_{\mathsf{deg}_i}$ 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 (10a), presumably with the structure (10b):

(10) a. *John buys more expensive wine than he buys beer.
b. John buys more expensive wine than [$\mathsf{O}_{\mathsf{deg}_i}$ he buys [$_{\text{NP}}$ t_i beer]]

In (10), the gap is in the 'left branch' position (Ross 1967), which is known to disallow extraction:

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

But this seemingly straightforward account of (10a) is at odds with the fact, apparently not noticed prior to Kennedy and Merchant (2000), that pseudogapped analogues of (10a) are impeccable:

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

Without argument, Kennedy and Merchant take the discrepancy between (10a) and (12) to be syntactic in nature. They offer a complex account which in its essence derives the ill-formedness of (10a) 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 [$_{\text{DetP}}$ 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 (13):

(13) ...than he does [$_{\text{VP}}$ $\cancel{[_{\text{VP}} \text{ buy } [_{\text{FocP}}\ \mathsf{O}_i\ t_j\ ]]}$ [$_{\text{DetP}_j}$ t_i beer]]

Thus, on Kennedy and Merchant's account, the contrast between (10a) and (12) 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, viz., 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 (10a) 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.

2.6 Summary

The preceding arguments appear to implicate unavoidably the presence of structure at the ellipsis site in VP ellipsis—but, as we argue in the remainder of this paper, the logic of Kennedy's (and Merchant's) argument hinges crucially on two key assumptions. In the first place, 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 (14) is linked to a position *within* the missing material following the auxiliary, rather than from a complement position directly associated with the auxiliary itself:

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

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.

3 Ellipsis in Hybrid TLCG

3.1 Hybrid TLCG: a sketch of the hybrid calculus

Our alternative analysis of the ellipsis-extraction relationship is built on our treatment of pseudogapping in a particular version of type-logical grammar called Hybrid Type-Logical Categorial Grammar (Hybrid TLCG; Kubota and Levine 2017).[4] Hybrid TLCG is a novel version of type-logical grammar which combines the Lambek calculus (Lambek 1958) and 'Lambda grammar' (Oehrle 1994; Muskens 2007; de Groote 2001), in that its proof theory includes both directional slashes (/ and \) for local combinatorics from the Lambek calculus and a 'non-directional' slash from Lambda grammar that we notate as a vertical slash ($\upharpoonright$). The latter can be thought of as a formal device that generalizes the notion of syntactic movement (both overt and covert) in the mainstream Chomskian syntax. See Kubota and Levine (2016a, Sect. 3.1.3) for some discussion on this point. The inference rules for Hybrid TLCG can be compactly stated as in (15):

[4] An extended gentle introduction to the framework is available in Kubota and Levine (2014); and detailed applications of the approach to a wide range of challenging empirical problems are offered in Kubota and Levine (2015), Kubota and Levine (2016a), and Kubota and Levine (2016b).

(15) **Connective** — **Introduction** — **Elimination**

$/$

$$\dfrac{\dfrac{\begin{array}{cccccc}\vdots & \vdots & \underline{[\varphi; x; A]^1} & \vdots & \vdots \\ \vdots & \vdots & \vdots \quad \vdots & \vdots & \vdots\end{array}}{b \circ \varphi;\ \boldsymbol{\mathcal{F}};\ B}}{b;\ \lambda x.\boldsymbol{\mathcal{F}};\ B/A}\ /\mathrm{I}^n \qquad \dfrac{a;\ \boldsymbol{\mathcal{F}};\ A/B \quad b;\ \boldsymbol{\mathcal{G}};\ B}{a \circ b;\ \boldsymbol{\mathcal{F}}(\boldsymbol{\mathcal{G}});\ A}\ /\mathrm{E}$$

$\backslash$

$$\dfrac{\dfrac{\begin{array}{cccccc}\vdots & \vdots & \underline{[\varphi; x; A]^1} & \vdots & \vdots \\ \vdots & \vdots & \vdots \quad \vdots & \vdots & \vdots\end{array}}{\varphi \circ b;\ \boldsymbol{\mathcal{F}};\ B}}{b;\ \lambda x.\boldsymbol{\mathcal{F}};\ A\backslash B}\ \backslash\mathrm{I}^n \qquad \dfrac{b;\ \boldsymbol{\mathcal{G}};\ B \quad a;\ \boldsymbol{\mathcal{F}};\ B\backslash A}{b \circ a;\ \boldsymbol{\mathcal{F}}(\boldsymbol{\mathcal{G}});\ A}\ \backslash\mathrm{E}$$

$\upharpoonright$

$$\dfrac{\dfrac{\begin{array}{cccccc}\vdots & \vdots & \underline{[\varphi; x; A]^1} & \vdots & \vdots \\ \vdots & \vdots & \vdots \quad \vdots & \vdots & \vdots\end{array}}{b;\ \boldsymbol{\mathcal{F}};\ B}}{\lambda\varphi.b;\ \lambda x.\boldsymbol{\mathcal{F}};\ B\upharpoonright A}\ \upharpoonright\mathrm{I}^n \qquad \dfrac{a;\ \boldsymbol{\mathcal{F}};\ A\upharpoonright B \quad b;\ \boldsymbol{\mathcal{G}};\ B}{a(b);\ \boldsymbol{\mathcal{F}}(\boldsymbol{\mathcal{G}});\ A}\ \upharpoonright\mathrm{E}$$

The directional slashes and the vertical slash differ only in their prosodic properties, reflecting two different modes of prosodic combination. Directionally slashed functional categories always combine via concatenation. Thus, expressions of directionally slashed categories and their arguments are both of prosodic type **st** (for 'string'). Expressions to which vertically slashed types are assigned are, in contrast, missing material in some possibly medial position, so concatenation is not the mode of prosodic combination for such expressions. Instead, as per the $\upharpoonright$ Elimination rule above, the prosody of a sign of type $Y\upharpoonright X$ is a prosodic function, taking the prosody of X as an argument to return some (itself possible functional) prosody.

Expressions involving the vertical slash are either lexically specified or syntactically derived. In the latter case, a type X variable is *hypothesized* at an earlier proof step to give rise to an expression of type Y, followed by an abstraction on X, yielding an expression 'looking for' an argument of type X. This abstraction in effect keeps track of an internal position occupied by a (hypothetical) expression of type X within the expression of type Y, thereby creating the possibility of a mismatch between the position to which some element is prosodically linked and the position in which it is semantically interpreted.

We illustrate this use of the vertical slash in the treatment of extraction, via the extraction operator due to Muskens (2003) defined as follows:[5]

(16) $\lambda\sigma.\mathsf{what} \circ \sigma(\boldsymbol{\epsilon});\ \mathbf{what};\ \mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP})$

Here, $\boldsymbol{\epsilon}$ represents the empty string. (16) exploits the mismatch possibility just alluded to by correlating the interrogative operator with the semantic variable in the gap position, 'marked' as it were, by $\boldsymbol{\epsilon}$, within the sign over which it takes scope. This is illustrated in the following derivation:

[5]Karttunen (1977), for example, represents a fairly standard approach to the interpretation of *wh* questions, though Muskens' treatment of extraction has no particular commitment to his analysis.

(17)

$$\dfrac{①\!\rightarrow \dfrac{\dfrac{\dfrac{\mathsf{ate};\ \mathbf{ate};\ \mathrm{VP/NP} \quad [\varphi_0; x; \mathrm{NP}]^1}{\mathsf{ate} \circ \varphi_0;\ \mathbf{ate}(x);\ \mathrm{VP}}\,/\mathrm{E} \qquad \mathsf{john};\ \mathbf{j};\ \mathrm{NP}}{\mathsf{john} \circ \mathsf{ate} \circ \varphi_0;\ \mathbf{ate}(x)(\mathbf{j});\ \mathrm{S}}\,\backslash\mathrm{E}}{\lambda\varphi_0.\mathsf{john} \circ \mathsf{ate} \circ \varphi_0;\ \lambda x.\mathbf{ate}(x)(\mathbf{j});\ \mathrm{S}\upharpoonright\mathrm{NP}}\,{\upharpoonright}\mathrm{I}^1 \qquad \begin{array}{l}\lambda\sigma.\mathsf{what} \circ \sigma(\boldsymbol{\epsilon});\\ \mathbf{what};\ \mathrm{Q}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP})\end{array}}{\mathsf{what} \circ \mathsf{john} \circ \mathsf{ate} \circ \boldsymbol{\epsilon};\ \mathbf{what}(\lambda x.\mathbf{ate}(x)(\mathbf{j}));\ \mathrm{Q}}\,{\upharpoonright}\mathrm{E}$$

The sign obtained at the final proof step is then given as an argument to, for example, verbs of type VP/Q such as *wonder*. The abstraction step ① is the key step in this proof. It produces a $\mathbf{st} \rightarrow \mathbf{st}$ functor of type $\mathrm{S}\upharpoonright\mathrm{NP}$. This essentially amounts to the proof that *John ate* __ is a sentence missing an NP inside it. Since an S missing an NP is obtained by first hypothetically assuming an NP in the object position and then later withdrawing that hypothesis, this type of proof is called 'hypothetical proof' (or 'hypothetical reasoning'). The *wh*-word, the prosodic manifestation of the interrogative operator, concatenates to the left of this functor after the latter is fed an empty string to fill in the 'gap'. At the same time, the semantic argument corresponding to the apparently missing direct object is abstracted over, defining a property corresponding to the set of things John ate. Thus, the *wh*-word, though separated by an arbitrary distance from the gap site, is associated with a property defined with respect to the semantic argument correlated with that gap site, capturing the fact that *what* is (syntactically) the 'filler' of the gap.

3.2 Ellipsis in Hybrid TLCG

For the analysis of VP ellipsis, we follow the approach presented in Kubota and Levine (2017), whose key idea—building on the previous work of Schachter (1978) and Miller (1990)—is that auxiliaries may in effect appropriate both the syntax and the semantics of the predicates that they combine with. In the simplest case, these predicates correspond to whole VPs along with their property denotations. We implement this analysis via the operator in (18) (where VP is an abbreviation for NP\S):

(18) **VP ellipsis/pseudogapping operator, version 1**
$\lambda\varphi.\varphi;\ \lambda\mathscr{F}.\mathscr{F}(P);\ \mathrm{VP}\upharpoonright(\mathrm{VP/VP})$
——where P is a free variable whose value is identified with the meaning of some linguistic sign in the preceding discourse with category VP

For example, (19) is licensed by first obtaining an elliptical version of *can't* on the basis of the antecedent:

(19) John can sing. Bill can't.

(20)

$$\dfrac{\begin{array}{l}\lambda\varphi.\varphi;\\ \lambda\mathscr{F}.\mathscr{F}(P); \mathrm{VP}\upharpoonright(\mathrm{VP/VP})\end{array} \qquad \begin{array}{l}\mathsf{can't};\\ \lambda Q \lambda x.\neg\Diamond Q(x); \mathrm{VP/VP}\end{array}}{\begin{array}{l}\mathsf{can't};\\ \lambda x.\neg\Diamond P(x); \mathrm{VP}\end{array}}\,{\upharpoonright}\mathrm{E}$$

and then carrying out the derivation in (21):

(21)

$$\dfrac{\begin{matrix}\mathsf{john};\\ \mathbf{j};\mathrm{NP}\end{matrix} \quad \dfrac{\begin{matrix}\mathsf{can};\\ \lambda P\lambda x.\Diamond P(x);\mathrm{VP/VP}\end{matrix} \quad \begin{matrix}\mathsf{sing};\\ \mathbf{sing};\ \mathrm{VP}\end{matrix}}{\mathsf{can}\circ\mathsf{sing};\ \lambda x.\Diamond\mathbf{sing}(x);\ \mathrm{VP}}/\mathrm{E}}{\mathsf{john}\circ\mathsf{can}\circ\mathsf{sing};\ \Diamond\mathbf{sing}(\mathbf{j});\ \mathrm{S}}\backslash\mathrm{E} \qquad \dfrac{\begin{matrix}\mathsf{bill};\\ \mathbf{b};\mathrm{NP}\end{matrix} \quad \begin{matrix}\vdots \quad \vdots\\ \mathsf{can't};\ \lambda x.\neg\Diamond\mathbf{sing}(x);\ \mathrm{VP}\end{matrix}}{\mathsf{bill}\circ\mathsf{can't};\ \neg\Diamond\mathbf{sing}(\mathbf{b});\ \mathrm{S}}\backslash\mathrm{E}$$

In (21), the free variable P in (20) is replaced by the constant **sing**, on the basis of the antecedent VP in the previous sentence. We assume that some anaphora resolution mechanism at the syntax-semantics interface is responsible for this variable instantiation.

For the reasons laid out in Kubota and Levine (2017), we extend this treatment to verbs of any valence, allowing auxiliaries to take the arguments of such verbs as their own valents. This approach—which goes back to Jacobson (1998), Evans (1988) and ultimately Cormack (1984)—entails that we will have cases such as those in (22), familiarly referred to in the ellipsis literature as instances of *pseudogapping*:

(22)
a. JOHN ate more PIZZA than MARY did ICE CREAM. (*did*: VP/NP)
b. Mary will donate more BOOKS to ENVIRONMENTAL charities than she will CDs to the HOMELESS mission. (*can*: VP/PP/NP)
c. I would offer JOHN more BOOKS than I would MARY CDs. (*would*: VP/NP/NP)
d. John will almost certainly bet more EUROS with MARY on the GAME than he will DOLLARS with SUSAN on the RACE. (*will*: VP/NP/PP/PP)

We sketch the analysis of pseudogapping in Kubota and Levine (2017) in what follows. Note first that in Hybrid TLCG, there is a theorem that any linguistic expression of type VP/VP also inhabits the type (VP/\$)/(VP/\$), where the variable /\$ is instantiated as a sequence of directional argument specifications as introduced in Steedman (1985). A simple proof for VP/VP $\vdash$ TV/TV is given in (23) (with k the name of an arbitrary auxiliary and $\mathscr{O}$ its denotation; TV abbreviates VP/NP).

(23)

$$\dfrac{\dfrac{\dfrac{\mathsf{k};\ \mathscr{O};\ \mathrm{VP/VP} \quad \dfrac{[\varphi_2; f; \mathrm{TV}]^2 \quad [\varphi_3; x; \mathrm{NP}]^3}{\varphi_2\circ\varphi_3;\ f(x);\ \mathrm{VP}}/\mathrm{E}}{\mathsf{k}\circ\varphi_2\circ\varphi_3;\ \mathscr{O}(f(x));\ \mathrm{VP}}/\mathrm{E}}{\mathsf{k}\circ\varphi_2;\ \lambda x.\mathscr{O}(f(x));\ \mathrm{TV}}/\mathrm{I}^3}{\mathsf{k};\ \lambda f\lambda x.\mathscr{O}(f(x));\ \mathrm{TV/TV}}/\mathrm{I}^2$$

Parallel proofs are available for all instances of rightward-argument-seeking types.[6]

In order to map an auxiliary typed (VP/\$)/(VP/\$) into a term typed VP/\$, as per the discussion above, we generalize the lexical entry for the VP ellipsis operator along the lines of (24):

(24) **VP ellipsis/pseudogapping operator, version 2**
$\lambda\varphi.\varphi;\ \lambda\mathscr{F}.\mathscr{F}(P);\ (\mathrm{VP/\$})\upharpoonright((\mathrm{VP/\$})/(\mathrm{VP/\$}))$
——where P is a free variable whose value is identified with the meaning of some linguistic sign in the preceding discourse with category VP/\$

[6] The proof in (23) exemplifies the Geach rule, a directional variant of the familiar rule of transitivity of implication $\phi \supset \psi \vdash (\zeta \supset \phi) \supset (\zeta \supset \psi)$ in classical propositional logic.

Simple VP ellipsis is derived by taking \$ in (24) to denote the empty sequence of right-seeking argument specifications (in which case the category of the ellipsis operator is instantiated as $\mathrm{VP}\upharpoonright(\mathrm{VP}/\mathrm{VP})$), and the proof in (21b) will follow automatically from (24) as a specific instance.

The analysis of a basic pseudogapping example like (25) then works in a completely parallel fashion as the analysis of VP ellipsis, except that it is a functor of type $e \to e \to t$ not $e \to t$, which is anaphorically recovered from context:

(25) John should eat the banana. Bill should ~~eat~~ the apple.

(26)

$$\dfrac{\mathsf{john}; \mathbf{j}; \mathrm{NP} \quad \dfrac{\mathsf{should}; \lambda P\lambda x.\Box P(x); \mathrm{VP/VP} \quad \dfrac{\mathsf{eat}; \mathbf{eat}; \mathrm{TV} \quad \mathsf{the} \circ \mathsf{banana}; \textbf{the-b}; \mathrm{NP}}{\mathsf{eat} \circ \mathsf{the} \circ \mathsf{banana}; \mathbf{eat}(\textbf{the-b}); \mathrm{VP}}/\mathrm{E}}{\mathsf{should} \circ \mathsf{eat} \circ \mathsf{the} \circ \mathsf{banana}; \lambda x.\Box\mathbf{eat}(\textbf{the-b})(x); \mathrm{VP}}/\mathrm{E}}{\mathsf{john} \circ \mathsf{should} \circ \mathsf{eat} \circ \mathsf{the} \circ \mathsf{banana}; \Box\mathbf{eat}(\textbf{the-b})(\mathbf{j}); \mathrm{S}}\backslash\mathrm{E}$$

$$\dfrac{\mathsf{bill}; \mathbf{b}; \mathrm{NP} \quad \dfrac{\dfrac{\begin{array}{c}\vdots\\ \lambda\varphi.\varphi; \lambda\mathscr{F}.\mathscr{F}(\mathbf{eat}); \mathrm{TV}\upharpoonright(\mathrm{TV/TV})\end{array} \quad \begin{array}{c}\vdots\\ \mathsf{should}; \lambda f\lambda x\lambda y.\Box f(x)(y); \mathrm{TV/TV}\end{array}}{\mathsf{should}; \lambda x\lambda y.\Box\mathbf{eat}(x)(y); \mathrm{TV}}/\mathrm{E} \quad \mathsf{the} \circ \mathsf{apple}; \textbf{the-a}; \mathrm{NP}}{\mathsf{should} \circ \mathsf{the} \circ \mathsf{apple}; \lambda y.\Box\mathbf{eat}(\textbf{the-a})(y); \mathrm{VP}}/\mathrm{E}}{\mathsf{bill} \circ \mathsf{should} \circ \mathsf{the} \circ \mathsf{apple}; \Box\mathbf{eat}(\textbf{the-a})(\mathbf{b}); \mathrm{S}}\backslash\mathrm{E}$$

Here, the auxiliary is in the derived TV/TV category. The VP ellipsis/pseudogapping operator in (24) takes this auxiliary category as an argument and saturates its TV argument by anaphorically referring to the transitive verb *eat* in the antecedent clause. The key idea here is that, via the general ellipsis operator, an auxiliary may take on the valence of any functional type which, after combining with some (possibly empty) sequence of arguments, will yield a VP.

3.3 'Extraction from elided VP' revisited

With these technical resources in hand, we now turn to the apparent cases of extraction from ellipsed material which form the basis for most of Kennedy's arguments reviewed in Sect. 1, as illustrated in examples such as (14):

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

Such cases 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. For example, referring to data such as (14), Johnson (2001) remarks that 'In these cases too the ellipsis site seems to have internal parts'. Elbourne (2008) 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. By contrast, the theory advocated here has normal syntactic structure in all ellipsis sites.
>
> (Elbourne 2008, 216)

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.[7] Our own analysis is immune to this criticisms since it analyzes apparent extraction from an ellipsis site as a genuine syntactic extraction—but not extraction *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 Kubota and Levine (2017). That is, examples such as (14) 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 $\mathrm{VP}\upharpoonright\mathrm{XP}$—corresponds to a VP in a filler/gap relationship with an XP filler. Thus, in *I wonder what John said to Mary*, *said. . . to Mary* constitutes a VP with a medial NP gap, meeting the description $\mathrm{VP}\upharpoonright\mathrm{NP}$. It follows then that in principle, a sentence such as (14) 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 (25) in Sect. 3.1), but could in fact represent any VP from which an NP is missing, as in, e.g., (27).

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

Our generalized ellipsis operator in (24) won't quite be sufficient here, because it is stated in terms of categories of the form VP/$, which target categories seeking arguments only to the right, and thus exclude 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 $(\mathrm{VP}\upharpoonright\mathrm{XP})\upharpoonright(\mathrm{VP}\upharpoonright\mathrm{XP})$:

[7] A certain caution is necessary in evaluating these objections, however. Elbourne apparently regards the problem with Hardt's analysis, vis-a-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 (14) 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 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.

(28)
$$\dfrac{\dfrac{\dfrac{\mathsf{k};\ \mathscr{O};\ \mathrm{VP/VP} \qquad \dfrac{[\sigma_1; f; \mathrm{VP{\upharpoonright}NP}]^1 \quad [\varphi_2; x; \mathrm{NP}]^2}{\sigma_1(\varphi_2);\ f(x);\ \mathrm{VP}}}{\mathsf{k} \circ \sigma_1(\varphi_2);\ \mathscr{O}(f(x));\ \mathrm{VP}}}{\lambda\varphi_2.\mathsf{k} \circ \sigma_1(\varphi_2);\ \lambda x.\mathscr{O}(f(x));\ \mathrm{VP{\upharpoonright}NP}}{\upharpoonright}\mathrm{I}^2}{\lambda\sigma_1\lambda\varphi_2.\mathsf{k} \circ \sigma_1(\varphi_2);\ \lambda f \lambda x.\mathscr{O}(f(x));\ (\mathrm{VP{\upharpoonright}NP}){\upharpoonright}(\mathrm{VP{\upharpoonright}NP})}{\upharpoonright}\mathrm{I}^1$$

The generalized ellipsis operator then takes such ‘Geached’ auxiliaries and maps them to type VP↾XP, anaphorically supplying the meaning of the gapped VP:[8]

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

The analysis of (the antecedent clause of) (30) then goes as in (31).

(30) I know what John ate for lunch, but I don't know what$_i$ Bill did ~~eat~~ $\_\__i$ ~~for lunch~~.

(31)
$$\dfrac{\dfrac{\dfrac{\dfrac{\textcircled{1}\rightarrow\dfrac{\dfrac{[\varphi_1; x; \mathrm{NP}]^1}{\vdots \quad \vdots}}{\begin{array}{l}\mathsf{ate} \circ \varphi_1 \circ \mathsf{for} \circ \mathsf{lunch};\\ \mathbf{ate}(x)(\mathbf{lunch}); \mathrm{VP}\end{array}}}{\begin{array}{l}\lambda\varphi_1.\mathsf{ate} \circ \varphi_1 \circ \mathsf{for} \circ \mathsf{lunch};\\ \lambda x.\mathbf{ate}(x)(\mathbf{lunch}); \mathrm{VP{\upharpoonright}NP}\end{array}}{\upharpoonright}\mathrm{I}^1 \qquad \begin{bmatrix}\varphi_2;\\ u; \mathrm{NP}\end{bmatrix}^2}{\begin{array}{l}\mathsf{ate} \circ \varphi_2 \circ \mathsf{for} \circ \mathsf{lunch};\\ \mathbf{ate}(u)(\mathbf{lunch});\\ \mathrm{VP}\end{array}} \qquad \begin{array}{l}\mathsf{john};\\ \mathbf{j};\\ \mathrm{NP}\end{array}}{\begin{array}{l}\mathsf{john} \circ \mathsf{ate} \circ \varphi_2 \circ \mathsf{for} \circ \mathsf{lunch};\\ \mathbf{ate}(u)(\mathbf{lunch})(\mathbf{j}); \mathrm{S}\end{array}}}{\begin{array}{l}\lambda\varphi_2.\mathsf{john} \circ \mathsf{ate} \circ \varphi_2 \circ \mathsf{for} \circ \mathsf{lunch};\\ \lambda u.\mathbf{ate}(u)(\mathbf{lunch})(\mathbf{j});\\ \mathrm{S{\upharpoonright}NP}\end{array}}{\upharpoonright}\mathrm{I}^2 \qquad \begin{array}{l}\lambda\sigma.\mathsf{what} \circ \sigma(\boldsymbol{\epsilon});\\ \lambda P.\mathbf{what}(P);\\ \mathrm{Q}{\upharpoonright}(\mathrm{S{\upharpoonright}NP})\end{array}}{\begin{array}{l}\mathsf{what} \circ \mathsf{john} \circ \mathsf{ate} \circ \boldsymbol{\epsilon} \circ \mathsf{for} \circ \mathsf{lunch};\\ \mathbf{what}(\lambda u.\mathbf{ate}(u)(\mathbf{lunch})(\mathbf{j})); \mathrm{Q}\end{array}}$$

The derivation of the ellipsis clause involves free instantiation of the VP↾NP variable introduced in the generalized ellipsis operated defined above. We take R to be the greyed-in semantic term obtained in the proof line ①. The first part of the proof for *what Bill did* then takes the following form:

(32)
$$\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{l}\vdots \quad \vdots\\ \lambda\sigma\lambda\varphi.\mathsf{did} \circ \sigma(\varphi);\\ \lambda f \lambda x \lambda y.f(x)(y);\\ (\mathrm{VP{\upharpoonright}NP}){\upharpoonright}(\mathrm{VP{\upharpoonright}NP})\end{array} \qquad \begin{array}{l}\lambda\rho\lambda\varphi.\rho(\lambda\varphi_0.\varphi_0)(\varphi);\\ \lambda\mathscr{F}.\mathscr{F}(\lambda x.\mathbf{ate}(x)(\mathbf{lunch}));\\ (\mathrm{VP{\upharpoonright}NP}){\upharpoonright}((\mathrm{VP{\upharpoonright}NP}){\upharpoonright}(\mathrm{VP{\upharpoonright}NP}))\end{array}}{\begin{array}{l}\lambda\varphi.\mathsf{did} \circ \varphi;\\ \lambda x \lambda y.\mathbf{ate}(x)(\mathbf{lunch})(y); \mathrm{VP{\upharpoonright}NP}\end{array}} \qquad [\varphi_3; v; \mathrm{NP}]^3}{\mathsf{did} \circ \varphi_3;\ \lambda y.\mathbf{ate}(v)(\mathbf{lunch})(y);\ \mathrm{VP}} \qquad \mathsf{bill};\ \mathbf{b};\ \mathrm{NP}}{\mathsf{bill} \circ \mathsf{did} \circ \varphi_3;\ \mathbf{ate}(v)(\mathbf{lunch})(\mathbf{b});\ \mathrm{S}}}{\lambda\varphi_3.\mathsf{bill} \circ \mathsf{did} \circ \varphi_3;\ \lambda v.\mathbf{ate}(v)(\mathbf{lunch})(\mathbf{b});\ \mathrm{S{\upharpoonright}NP}}{\upharpoonright}\mathrm{I}^3$$

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

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 Appendix 1 that the empirical basis for these objections is quite fragile and problematic when a larger range of relevant data is taken into account.

3.4 Ellipsis and comparatives

The final 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 1995, 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.

3.4.1 The basic syntax-semantics interface of comparatives

The semantics of comparatives reflect 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 (33a), for example, the degrees to which Mary and Ann are tall respectively.

(33) 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 (33) is to assume that both the main clause and the *than* clause denote predicates of degree of type $d \to t$. For example, (33a) 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:[9]

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

[9](34) may sound less natural than (33a), but this is arguable due to the redundancy of repeating the adjective *tall* in the *than* clause.

with an 'LF' that looks like the following:

(35) 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 modelled by assuming that adjectives are of type $d \to e \to t$ semantically and Adj$\upharpoonright$Deg syntactically.

Putting aside the exact morpho-syntax of comparative forms, the degree operator can be defined as follows:[10]

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

The derivation for (35) then goes as follows:

(37)

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

Whether (33a) is derived from (34) 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 morphology via a functional prosodic term, which is convenient for the treatment of suppletive forms. Specifically we assume that the $\mathbf{st} \to \mathbf{st}$ function ER takes the strong 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 (38):

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

[10] This option makes it possible to encode the comparative morphology via a functional prosodic term, which is convenient for the treatment of suppletive forms. Specifically, we assume that the $\mathbf{st} \to \mathbf{st}$ function ER takes the string prosody of the base form of an adjective and returns the comparative form of that adjective.

The difference between (36) and (38) is that in (38), the two clauses are missing the entire adjective rather than just the degree argument. Correspondingly, in (38), 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 (38).

With the definition in (38), the derivation for (33a) goes as in (39):

(39)

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

3.4.2 VP ellipsis in comparatives

The analysis of comparatives sketched above interacts straightforwardly with the analysis of VP ellipsis from Sect. 3.1 to yield the right meanings for examples like the following:

(40) 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 (39)) is also straightforward in our approach. This is left as an exercise for the reader.

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

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

With this lexical entry for the adverb *fast*, the derivation for (40) goes as follows (here, we abbreviate Deg as D):

(42)

$$\dfrac{\lambda\varphi.\mathsf{fast}\circ\varphi;\ \mathbf{fast};\ (\mathrm{VP}\backslash\mathrm{VP})\upharpoonright\mathrm{D} \qquad \begin{bmatrix}\varphi;\\ d;\\ \mathrm{D}\end{bmatrix}^3}{\mathsf{fast}\circ\varphi;\ \mathbf{fast}(d);\ (\mathrm{VP}\backslash\mathrm{VP})}$$

$$\dfrac{\mathsf{ran};\ \mathbf{ran};\ \mathrm{VP} \qquad \mathsf{fast}\circ\varphi;\ \mathbf{fast}(d);\ (\mathrm{VP}\backslash\mathrm{VP})}{\dfrac{\mathsf{ran}\circ\mathsf{fast}\circ\varphi;\ \mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}}{\lambda\varphi.\mathsf{ran}\circ\mathsf{fast}\circ\varphi;\ \lambda d.\mathbf{fast}(d)(\mathbf{ran});\ \boxed{\mathrm{VP}\upharpoonright\mathrm{Deg}}}\upharpoonright\mathrm{I}^3}$$

$$\dfrac{\mathsf{john};\ \mathbf{j};\ \mathrm{NP} \qquad \dfrac{\lambda\varphi.\mathsf{ran}\circ\mathsf{fast}\circ\varphi;\ \lambda d.\mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}\upharpoonright\mathrm{Deg} \qquad \begin{bmatrix}\varphi;\\ d;\\ \mathrm{D}\end{bmatrix}^2}{\mathsf{ran}\circ\mathsf{fast}\circ\varphi;\ \mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}}}{\dfrac{\mathsf{john}\circ\mathsf{ran}\circ\mathsf{fast}\circ\varphi;\ \mathbf{fast}(d)(\mathbf{ran})(\mathbf{j});\ \mathrm{S}}{\lambda\varphi.\mathsf{john}\circ\mathsf{ran}\circ\mathsf{fast}\circ\varphi;\ \lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{j});\ \mathrm{S}\upharpoonright\mathrm{D}}\upharpoonright\mathrm{I}^2}$$

$$\dfrac{\begin{array}{c}\lambda\sigma_1\lambda\sigma_2.\sigma_1(\mathsf{er})\circ\\ \mathsf{than}\circ\sigma_2(\boldsymbol{\epsilon});\\ \lambda P\lambda Q.\mathbf{max}(P)\\ >\mathbf{max}(Q);\\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{D})\upharpoonright(\mathrm{S}\upharpoonright\mathrm{D})\end{array} \qquad \lambda\varphi.\mathsf{john}\circ\mathsf{ran}\circ\mathsf{fast}\circ\varphi;\ \lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{j});\ \mathrm{S}\upharpoonright\mathrm{D}}{\begin{array}{l}\lambda\sigma_2.\mathsf{john}\circ\mathsf{ran}\circ\mathsf{fast}\circ\mathsf{er}\circ\mathsf{than}\circ\sigma_2(\boldsymbol{\epsilon});\\ \lambda Q.\mathbf{max}(\lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{j}))>\mathbf{max}(Q);\\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{D})\end{array}}$$

$$\dfrac{\begin{array}{l}\lambda\rho\lambda\varphi.\rho(\mathsf{vac}_{\mathsf{s}\rightarrow\mathsf{s}})(\varphi);\\ \lambda\mathscr{G}.\mathscr{G}(\lambda d.\mathbf{fast}(d)(\mathbf{ran}));\\ (\mathrm{VP}\upharpoonright\mathrm{D})\upharpoonright((\mathrm{VP}\upharpoonright\mathrm{D})\upharpoonright(\mathrm{VP}\upharpoonright\mathrm{D}))\end{array} \qquad \begin{array}{l}\lambda\sigma\lambda\varphi.\mathsf{did}\circ\sigma(\varphi);\\ \lambda\mathscr{F}.\mathscr{F};\\ (\mathrm{VP}\upharpoonright\mathrm{D})\upharpoonright(\mathrm{VP}\upharpoonright\mathrm{D})\end{array}}{\lambda\varphi.\mathsf{did}\circ\varphi;\ \lambda d.\mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}\upharpoonright\mathrm{D}}$$

$$\dfrac{\mathsf{mary};\ \mathbf{m};\ \mathrm{NP} \qquad \dfrac{\begin{bmatrix}\varphi;\\ d;\\ \mathrm{D}\end{bmatrix}^1 \qquad \lambda\varphi.\mathsf{did}\circ\varphi;\ \lambda d.\mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}\upharpoonright\mathrm{D}}{\mathsf{did}\circ\varphi;\ \mathbf{fast}(d)(\mathbf{ran});\ \mathrm{VP}}}{\dfrac{\mathsf{mary}\circ\mathsf{did}\circ\varphi;\ \mathbf{fast}(d)(\mathbf{ran})(\mathbf{m});\ \mathrm{S}}{\lambda\varphi.\mathsf{mary}\circ\mathsf{did}\circ\varphi;\ \lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{m});\ \mathrm{S}\upharpoonright\mathrm{D}}\upharpoonright\mathrm{I}^1}$$

$$\dfrac{\begin{array}{l}\lambda\sigma_2.\mathsf{john}\circ\mathsf{ran}\circ\mathsf{fast}\circ\mathsf{er}\circ\mathsf{than}\circ\sigma_2(\boldsymbol{\epsilon});\\ \lambda Q.\mathbf{max}(\lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{j}))>\mathbf{max}(Q);\\ \mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{D})\end{array} \qquad \lambda\varphi.\mathsf{mary}\circ\mathsf{did}\circ\varphi;\ \lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{m});\ \mathrm{S}\upharpoonright\mathrm{D}}{\begin{array}{l}\mathsf{john}\circ\mathsf{ran}\circ\mathsf{fast}\circ\mathsf{er}\circ\mathsf{than}\circ\mathsf{mary}\circ\mathsf{did};\\ \mathbf{max}(\lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{j}))>\mathbf{max}(\lambda d.\mathbf{fast}(d)(\mathbf{ran})(\mathbf{m}));\\ \mathrm{S}\end{array}}$$

In somewhat informal terms, on this analysis what is 'elided' in the *than* clause in (40) 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 VP↾Deg, with the denotation $\lambda d.\mathbf{fast}(d)(\mathbf{ran})$. Thus, the derivation is parallel to the extraction-ellipsis interaction case in (32) from Sect. 3.3, with the only difference that the missing expression here is of type Deg rather than type NP.

4 Kennedy's arguments: a second look

We now revisit Kennedy's arguments as summarized in Sect. 2, in light of the alternative analysis of extraction/ellipsis interaction just sketched. In the following discussion we demonstrate that, in every case, the data that Kennedy takes to establish the inevitability of a covert structure analyses for ellipsis phenomena can be accounted for far more straightforwardly on a 'direct interpretation' approach.[11] 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 (14) do involve an actual extraction renders his argument moot. There are, moreover, serious issues about the status of some of the critical data presented in Kennedy and Merchant (2000) and Kennedy (2003) which undercut the force of the argument in these sources based on those data.

4.1 Ellipsis and islandhood

Kennedy's argument from islandhood presumes that data such as (1), reproduced here, are fully representative.

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

[11] These demonstrations, in turn, lend added support to the alternative non-structural treatment of VP ellipsis and pseudogapping in Kubota and Levine (2017), on the basis of the approach illustrated in (30)–(32).

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

It is in fact not difficult to show that they are not:

(43) STERLING criticized every decision (that) Doug was unhappy (that) LOU $\left\{ \begin{array}{ll} \text{a.} & \text{*did} \\ \text{b.} & \text{criticized } t \end{array} \right\}$.

(43a), for example, is uncontroversially at least as bad as the examples in (1b). But the extraction site that, on a covert-structure analysis, would have been ellipsed is not located within an island; as (43b) makes clear, extraction is perfectly well-formed in this context. Exactly the same observation holds for (2) vs. (44):

(2) Max refused to buy the shirt that I picked out even though it was less expensive than the one that (*the salesperson complemented him after) he did (buy t).

(44) *Max refused to buy the shirt that I picked out even though it was less expensive than the one that the salesperson said that he was wise to.

What (2) and (44) have in common is not covariation of ill-formedness with ellipsis in island configurations (note the well-formedness of *That's one shirt which*$_i$ *the salesman said you were very wise to pick out* t_i), but much more simply, the degree of embedding of the ellipsed material. Theorists from widely divergent approaches have noted that such embedding corresponds to semantic/pragmatic difficulties which are in themselves sufficient to guarantee that such examples are perceived as anomalous (see, e.g., Jacobson (2016) on short answer ellipsis and Toosarvandani (2016) on Gapping). In fact, Kennedy himself had, in prior work on the identity conditions holding between antecedent and ellipsed VP (Kennedy 1994), argued that examples such as (43) and (44) are ill-formed because they violate a complex set of essentially semantic requirements in which islandhood considerations have no role at all. And (1)–(2) violate these conditions in precisely the same way. Thus, Kennedy's own account of ACD in Kennedy (1994), by unifying the patterns displayed in (1) on the one hand and (44), on the other, decisively undercuts his own invocation of islandhood—and the concealed structure which it would entail—in his 2003 analysis of VP ellipsis.

Furthermore, the whole concept of adjuncts as islands is dubious. A wide variety of legal examples of extraction from non-finite clausal constructs are given in Truswell (2007), and certain finite adjunct extraction as in (45) seem unexceptionable as well:

(45) There's a major double standard here: STERLING can freely criticize all the decisions that everyone would be REALLY upset if I criticized.

In (45), the *if* clause is just as much an adjunct as the *because* clause in (1b), but extraction from within the former appears altogether unexceptionable. Note however that the acceptability noticeably declines when VP ellipsis is involved:

(46) ?(?)Sterling can apparently criticize every decision that people would be REALLY upset if I did.

Again, it appears that some version of the kind of condition proposed in Kennedy (1994) and Toosarvandani (2016) accounts for the contrast between (45) and (46).

The same conclusion follows from the second class of island examples in Kennedy (2003):

(47) a. Dogs, I understand, but cats, I don't.

b. *Dogs, I understand, but cats$_i$, I don't know a single person who does ~~[$_{VP}$ understand t_i]~~.

In (47b), the gap labeled t_i is located within a relative clause, violating Subjacency both in its Complex NP Constraint manifestation and in its *wh*-Island form. Kennedy's point is that since (47b), on the syntactic deletion account of VP ellipsis, derives from an underlying structure that violates an island constraint, it is predicted to be bad. In fact, the judgment in (47b) is disputed by our informants, according to whom the example is only somewhat awkward at worst. Moreover, these examples are far from representative, and a slightly broader class of data strongly suggests that here too, syntactic islandhood *per se* is orthogonal to the source of any negative judgments reported in (47):[12]

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

Our informants find this example impeccable, despite the fact that the source of the elided VP in the second *and* conjunct would correspond to a putative island violation structure parallel to (47b). Moreover, ATB extraction in coordination constructions does not ameliorate island violations within any conjunct, so that the goodness of (48) cannot simply be due to the fact that ATB extraction is involved. Rather, we apparently need to take seriously the possibility that extraction from relative clauses is *not* in general structurally forbidden. Whatever the reasons may be for the diminished status of (47b) among certain speakers, the semantic parallelism between the two clauses in the *but* conjunct in (48) is apparently all that is needed to make the gap in the elided VP in the right conjunct transparent to linkage with the fronted topic.[13] The pattern just described then can be seen as the reflection of a priming effect which speaks to a processing problem rather than a grammaticality issue *per se*.[14]

To summarize, we have cases such as (1)–(2) and (43)–(44) in which extraction from non-islands at levels of embedding comparable to island contexts are equally unfriendly to ellipsis, and cases such as (48) where purported islands configurations fail to block ellipsis, leaving behind perfectly well-formed examples containing 'illegally extracted' remnants. It then seems fair to conclude that when a wider perspective on the relevant data is taken, the kinds of examples that Kennedy provides from island patterns are far from compelling.

[12] We can go further: when (47b) is replaced by (i), with emphatically stressed *anyone*, all the speakers we have consulted have unhesitatingly identified the result as good:

(i) . . . , but cats$_i$, I don't know ANYONE who$_j$ t_j does ~~[$_{VP}$ understand t_i]~~.

These judgments are exactly in line with Kluender's (1998) demonstration that both center-embedding and Subjacency are ameliorated when lexically heavy NPs at clause boundaries are replaced by pronouns.

[13] While the mechanisms involved in this effect are far from transparent, the fact that both clauses present a strong negation contrast to the first conjunct is surely a crucial factor.

[14] We discuss below other kinds of priming effects with significant impact on acceptability judgments. Note however that the fact that examples such as (48) are overall acceptable to many speakers runs counter to the pattern noted above in relation to examples such as (43)–(44) whereby the acceptability status of apparent extractions from elided VPs declines with the increasing depth of the ellipsis site. The exact conditions that govern the acceptability patterns of extraction out of embedded elided VPs is an important issue that needs to be investigated further in future research.

4.2 Ellipsis and SCO

Kennedy's argument based on SCO, as outlined in Sect. 2.2, hinges on the assumption that data such as (4) indeed reflects extraction from an elided VP.

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

However, on the alternative analysis presented in Sect. 3.3, an analogue of the following representation of the relevant coindexing relations in (4) is licensed:

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

The point is that what is extracted in (49) could correspond, on this analysis, to an argument of the auxiliary itself. Thus, any account of SCO which motivates ordinary cases such as (3) will predict SCO in pseudogapping + extraction cases of the sort exhibited in (4) on exactly the same basis, with no covert structure.[15] 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 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.[16]

4.3 Ellipsis and Condition B

For the analysis of the so-called Binding Condition B, we adopt Jacobson's (2007) 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 'prob-

[15] 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 remnant in pseudogapping examples such as (22) represents 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.

[16] In the Principles and Parameters approach, SCO has standardly been assumed for decades to represent a consequence of Binding Condition C (Chomsky 1981). However, Postal (2004) has persuasively argued at length that Condition C fails to account for SCO in a variety of different syntactic contexts.

A particularly clear case comes from the following contrast:

(i) a. *[Which nurse]$_i$ did you tell [everyone [who knew her$_i$]] that we had voted for t_i .
b. We told everyone who knew her$_i$ that we had voted for Mary$_i$.
c. [Which nurse]$_i$ did you tell John that we had voted for t_i .

The SCO effect in (ia) cannot be reduced to Condition C since (ib), a counterpart of (ia) with an in-situ R-expression in the gap site, is perfectly acceptable.

Postal adduces many other exceptions and contraindications to the source of SCO in Condition C effects. More recent work argues against a Condition C account (e.g, Safir (2014), although it is not clear that Safir's alternative 'Independence Principle' is any better at accounting for cases such as (ib) above). At present, it seems fair to say that there is no particular consensus on what the origins of the SCO patterns are.

lems which seem insurmountable' (p. 218).[17] The key idea in Jacobson's alternative, non-configurational account of Condition B effects is that NPs are divided by a binary-valued feature ±p, with pronouns marked NP[+p] and all other NPs NP[−p]. In all lexical entries of the form in (50), all NP (and PP) arguments in any realization of /$ are specified as [−p].[18]

(50) 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:[19]

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

The greyed-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 assume, following the standard practice, that free (i.e. unbound) pronouns are simply translated as arbitrary variables (cf. Cooper 1979).

Crucially, the operator in (51) 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 by the proof in (52):

[17]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 neither by 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 2000). 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 2014). It is worth noting that efforts to extend the Bach-Partee approach to cover Condition B were not particularly successful; see Jäger (2005), ch. 2 and Morrill (2010, 124–125) for discussion.

[18]We take VP to abbreviate NP\S, where NP is underspecified for the value of p, i.e., NP[−p] NP[+p]. Unlike pronouns in object position, no special conditions or restrictions apply to subject pronouns, hence the choice of ± value for such pronouns need not be specified.

[19]For expository purposes, we state the operator in (51) 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$_i$.

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 VP/$/XP[−p]/$ to yield an output of the form VP/$/XP[+p]/$. 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 [±p] feature percolates from NPs to PPs and by generalizing the irreflexive operator still further so that it applies not just to VP/NP[−p] but AP/NP[−p] as well, the ungrammaticality of (ii) follows straightforwardly.

(52)
$$\cfrac{\cfrac{\cfrac{\begin{array}{l}\lambda\varphi.\varphi;\\ \lambda f\lambda u\lambda v.f(u)(v), u \neq v;\\ (\mathrm{VP/NP[+p]})\upharpoonright(\mathrm{VP/NP[-p]})\end{array} \qquad \begin{array}{l}\mathsf{praises};\\ \mathbf{praise}; \mathrm{VP/NP[-p]}\end{array}}{\mathsf{praises};\ \lambda u\lambda v.\mathbf{praise}(u)(v), u \neq v;\ \mathrm{VP/NP[+p]}} \qquad \mathsf{him};\ z;\ \mathrm{NP[+p]}}{\mathsf{praises} \circ \mathsf{him};\ \lambda v.\mathbf{praise}(z)(v), z \neq v;\ \mathrm{VP}} \qquad \mathsf{john};\ \mathbf{j};\ \mathrm{NP[-p]}}{\mathsf{john} \circ \mathsf{praises} \circ \mathsf{him};\ \mathbf{praise}(z)(\mathbf{j}), z \neq \mathbf{j};\ \mathrm{S}}$$

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 examples of bound variable anaphora such as (53) can be analyzed in this setup, and then explain the Condition B effects for bound pronouns.

(53) Every editor believes that John admires him.

For the bound variable reading of pronouns, we assume the following lexical entry, which encodes non-linearity enforced by binding in the lexical specification (note that, in the semantic term, two tokens of the same variable w is bound by the same lambda operator):[20]

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

With this lexical sign for *him*, we obtain the correct interpretation of (53) as follows (here, $\mathbf{\forall_{editor}}$ is an abbreviation for the term $\lambda Q.\forall x[\mathbf{editor}(x) \rightarrow Q(x)]$):

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

By applying (i) to the free pronoun (ii), we obtain the bound pronoun entry identical to (54).

(ii) $\mathsf{him};\ u;\ \mathrm{NP}_{acc}[+p]$

(55)

$$
\begin{prooftree}
\AxiomC{$\vdots$}
\noLine
\UnaryInfC{$\mathsf{admires};\ \lambda u\lambda v.\mathbf{admire}(u)(v),\ u \neq v; \mathrm{VP/NP[+p]}$}
\AxiomC{$\vdots$}
\noLine
\UnaryInfC{$\varphi_0;\ z;\ \mathrm{NP[+p]}$}
\BinaryInfC{$\mathsf{admires} \circ \varphi_0;\ \lambda v.\mathbf{admire}(z), z \neq v; \mathrm{VP}$}
\AxiomC{$\mathsf{john};\ \mathbf{j}; \mathrm{NP[-p]}$}
\BinaryInfC{$\mathsf{john} \circ \mathsf{admires} \circ \varphi_0;\ \mathbf{admire}(z)(\mathbf{j}), z \neq \mathbf{j}; \mathrm{S}$}
\AxiomC{$\mathsf{thinks};\ \mathbf{think};\ \mathrm{VP/S}$}
\BinaryInfC{$\mathsf{thinks} \circ \mathsf{john} \circ \mathsf{admires} \circ \varphi_0;\ \mathbf{think}(\mathbf{admire}(z)(\mathbf{j})), z \neq \mathbf{j}; \mathrm{VP}$}
\UnaryInfC{$\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]}$}
\AxiomC{$\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]})$}
\BinaryInfC{$\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}$}
\AxiomC{$\varphi_1;\ y;\ \mathrm{NP}$}
\BinaryInfC{$\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}$}
\UnaryInfC{$\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}$}
\AxiomC{$\lambda\sigma_1(\mathsf{every} \circ \mathsf{editor});\ \mathbf{\forall}_{\mathbf{editor}};\ \mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP[-p]})$}
\BinaryInfC{$\mathsf{every} \circ \mathsf{editor} \circ \mathsf{thinks} \circ \mathsf{john} \circ \mathsf{admires} \circ \mathsf{him};\ \mathbf{\forall}_{\mathbf{editor}}(\lambda y.\mathbf{think}(\mathbf{admire}(y)(\mathbf{j}))(y), y \neq \mathbf{j}); \mathrm{S}$}
\end{prooftree}
$$

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 (52) above.

(56) *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 with the subject quantifier's binding the object pronoun, thus rendering the sentence uninterpretable on the intended reading. The derivation is shown in (57):

(57)

$$
\begin{prooftree}
\AxiomC{$\vdots$}
\noLine
\UnaryInfC{$\mathsf{congratulated};\ \lambda u\lambda v.\mathbf{congratulated}(u)(v), u \neq v;\ \mathrm{VP/NP[+p]}$}
\AxiomC{$\vdots$}
\noLine
\UnaryInfC{$\varphi_0;\ w;\ \mathrm{NP[+p]}$}
\BinaryInfC{$\mathsf{congratulated} \circ \varphi_0;\ \lambda v.\mathbf{congratulated}(w)(v), w \neq v; \mathrm{VP}$}
\UnaryInfC{$\lambda\varphi_0.\mathsf{congratulated} \circ \varphi_0;\ \lambda w\lambda v.\mathbf{congratulated}(w)(v), w \neq v;\ \mathrm{VP}{\upharpoonright}\mathrm{NP[+p]}$}
\AxiomC{$\mathsf{him};\ \lambda\mathscr{G}\lambda y.\mathscr{G}(y)(y);\ \mathrm{VP}{\upharpoonright}(\mathrm{VP}{\upharpoonright}\mathrm{NP[+p]})$}
\BinaryInfC{$\mathsf{congratulated} \circ \mathsf{him};\ \lambda y.\mathbf{congratulated}(y)(y), y \neq y; \mathrm{VP}$}
\AxiomC{$\varphi_1;\ x;\ \mathrm{NP}$}
\BinaryInfC{$\varphi_1 \circ \mathsf{congratulated} \circ \mathsf{him};\ \mathbf{congratulated}(x)(x), x \neq x; \mathrm{S}$}
\UnaryInfC{$\lambda\varphi_1.\varphi_1 \circ \mathsf{congratulated} \circ \mathsf{him};\ \lambda x.\mathbf{congratulated}(x)(x), x \neq x; \mathrm{S}{\upharpoonright}\mathrm{NP}$}
\AxiomC{$\lambda\sigma_1.\sigma_1(\mathsf{every} \circ \mathsf{editor});\ \mathbf{\forall}_{\mathbf{editor}};\ \mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}\mathrm{NP[-p]})$}
\BinaryInfC{$\mathsf{every} \circ \mathsf{editor} \circ \mathsf{congratulated} \circ \mathsf{him};\ \mathbf{\forall}_{\mathbf{editor}}(\lambda x.\mathbf{congratulated}(x)(x), x \neq x); \mathrm{S}$}
\end{prooftree}
$$

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 the following (= (5) from Sect. 2.3):

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

the obligatorily disjoint interpretation of the two pronouns requires a specific covert structure as the supposed basis for this enforced interpretation. But the interaction of our lexically-based semantic treatment of 'Condition B' effects with the theory of ellipsis motivated in Kubota and Levine (2017) automatically accounts for the pattern in (58). The derivation is given in (59) (here, we ignore the tense and modal meaning of the future tense auxiliary *won't*; **irr** abbreviates the term $\lambda f \lambda u \lambda v.f(u)(v), u \neq v$).[21]

(59)

$$\dfrac{\dfrac{\begin{array}{l}\mathsf{takes} \circ \mathsf{care} \circ \mathsf{of};\\ \mathbf{irr}(\mathbf{take\text{-}care}); \mathrm{VP/NP[+p]}\end{array} \quad \mathsf{him};\ w;\ \mathrm{NP[+p]}}{\begin{array}{l}\mathsf{takes} \circ \mathsf{care} \circ \mathsf{of} \circ \mathsf{him};\\ \mathbf{irr}(\mathbf{take\text{-}care})(w); \mathrm{VP}\end{array}} \quad \mathsf{john};\ \mathbf{j};\ \mathrm{NP[-p]}}{\begin{array}{l}\mathsf{john} \circ \mathsf{takes} \circ \mathsf{care} \circ \mathsf{of} \circ \mathsf{him};\\ \mathbf{irr}(\mathbf{take\text{-}care})(w)(\mathbf{j}); \mathrm{S}\end{array}}$$

$$\dfrac{\begin{array}{l}\mathsf{won't};\\ \neg\mathbf{irr}(\mathbf{take\text{-}care})(w); \mathrm{VP}\end{array} \quad \mathsf{he};\ y;\ \mathrm{NP[+p]}}{\mathsf{he} \circ \mathsf{won't};\ \neg\mathbf{irr}(\mathbf{take\text{-}care})(w)(y);\ \mathrm{S}}$$

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 (58), 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 non-structural analysis of VP ellipsis combined with a non-structural 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 advantage over the standard view have independently been established in the literature.

4.4 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 (60) and (61) ((7) and (8) from Sect. 2.4):

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

(i) $\mathsf{take} \circ \mathsf{care} \circ \mathsf{of}$; **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?

(60) 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$_i$.

(61) Otis is a person who$_i$ I admire __$_i$ because close friends of __$_i$ seem to ∅ .

Here we annotate Kennedy's original datum in (60b) 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 (60b) (in the gap version) reasonably acceptable but (60a) bad? It has been widely assumed in the literature that the contrast between (60a) and (60b) 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 (60a) and (60b), as we discuss in greater detail below.

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):

(62) 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) contains many such examples, which point to the conclusion that subjects are not, in themselves, island contexts. On this view, example such as (63) are grammatical, but unacceptable for reasons explored in Kluender (2004) and Chaves (2013).

(63) ??*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 decline in speakers' judgments of their status.

(64) a. ??*Which disease$_i$ do people rightly fear __$_i$ because the cure for __$_i$ will never be discovered __$_i$?

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

These examples essentially have the same syntactic structure as (60a).

We thus seek an explanation for three facts which are very likely interlinked:

(i) the typically (but not universally) greatly reduced acceptability of extraction from subjects without a corresponding gap elsewhere in the clause (which we will call an 'unsupported' gap)

(ii) the much improved status of subject extractions when such a corresponding VP-internal gap is linked to the same filler

(iii) the fact that a gap in an adjunct clause which is unsupported within that clause is markedly unacceptable even when exactly the same gap in a root-clause extraction site, as in e.g. (62) is reasonably acceptable

(iv) the much improved status of embedded subject-internal gaps with supporting gaps as compared to (iii)

In addition to this empirical puzzle, there is a problem forced on us from a completely different direction, based on the logical basis of Hybrid TLCG itself. Multiple gaps linked to a single filler, regardless of construction type, represent a possibly fundamental problem for versions of categorial grammar based on linear logic.[22] The resource-boundedness of linear logic—in particular, the unavailability of rules of contraction and weakening—mean that, in general, two tokens of a single variable cannot be introduced at different points in the course of a proof. Thus, an operator can bind exactly one token of a given variable, or two different operators can bind two different variables of the same type, but we are not permitted to introduced a single variable twice within a proof to be eventually bound by a single operator. But parasitic gaps seem to call for precisely such proof steps in order to be licensed.

Leaving aside for the moment the difficulties of the data set displayed above, we can overcome the technical difficulty just noted by simultaneously introducing multiple tokens of a variable which are 'pre-bound' in a lexical operator, along the lines of (65):[23]

(65) $\lambda\sigma_1\lambda\sigma_2\lambda\varphi.\sigma_2(\varphi) \circ \sigma_1(\varphi)$; $\lambda R\lambda g\lambda x.R(x)(g(x))$; $(\mathrm{S}{\upharpoonright}\mathrm{NP}){\upharpoonright}(\mathrm{NP}{\upharpoonright}\mathrm{NP}){\upharpoonright}((\mathrm{NP}\backslash\mathrm{S}){\upharpoonright}\mathrm{NP})$

This operator takes a VP missing an NP and an NP missing an NP and in effect fuses the missing NP argument variables, and likewise for the prosodic variables. The result, supplied

[22] This includes the original Lambek calculus, various versions of 'Abstract Categorial Grammar' (de Groote 2001; Muskens 2003), Linear Categorial Grammar (Pollard 2013) and Hybrid TLCG, but not Combinatory Categorial Grammar (CCG) (Steedman 2000), which does not correspond to any specific logical proof theory. For a treatment of parasitic gaps in CCG, see Steedman (1987).

[23] Note that this operator, on its own, licenses parasitic gaps in the subject position only. In order to license parasitic gaps in adjunct clauses, a second, separate operator is needed. One might object to this type of analysis on the ground that it fails to capture parasitic gaps in the two environments via a uniform mechanism. However, it is questionable whether a uniform analysis is desirable to begin with. Culicover (1999, 179–181) notes that 'there appear to be many languages that have *without*-parasitic gaps and parasitic gaps in adjuncts, but lack parasitic gaps in subjects', giving Spanish and German as two such languages. We therefore take the need for separate operators for adjunct and subject parasitic gaps to be well-supported empirically.

Questions however remain regarding the much larger issue of how to deal with a wider range of multiple gap phenomena in Hybrid TLCG. We can find in English a variety of other multiple gap/single filler constructions, including 'parasitic' gaps in adjuncts, ATB extraction in coordination, the 'symbiotic' gaps pointed out in Levine and Sag (2003), and cases such as (i), where both gaps are in extractable positions within a single VP:

(i) Which people did you show pictures of $\_\__i$ to $\_\__i$.

We leave it for future work to extend our analysis to such cases.

as an argument to the operator introduced in (31), yields a sign with the empty string in two separate positions, corresponding to two tokens of a single variable in the semantic component of the sign, bound by a single abstraction operator.[24] The action of the gap-multiplying operator (65) is illustrated in (66):

(66)

$$\dfrac{\begin{array}{l}\lambda\sigma.\mathsf{who}\circ\sigma(\epsilon);\\ \lambda P\lambda Q\lambda x.\\ \quad Q(x)\wedge P(x);\\ (\mathrm{N}\backslash\mathrm{N})\upharpoonright(\mathrm{S}\upharpoonright\mathrm{NP})\end{array} \qquad \dfrac{\begin{array}{c}\vdots \quad \vdots\\ \begin{array}{l}\lambda\varphi.\mathsf{the}\circ\mathsf{close}\circ\\ \quad\mathsf{friends}\circ\mathsf{of}\circ\varphi;\\ \lambda x.\iota(\lambda y.\mathbf{close\text{-}fr}\\ \quad(x)(y));\mathrm{NP}\upharpoonright\mathrm{NP}\end{array}\end{array} \qquad \dfrac{\begin{array}{l}\lambda\sigma_1\lambda\sigma_2\lambda\varphi.\sigma_2(\varphi)\circ\sigma_1(\varphi);\\ \lambda R\lambda g\lambda x.R(x)(g(x));\\ (\mathrm{S}\upharpoonright\mathrm{NP})\upharpoonright(\mathrm{NP}\upharpoonright\mathrm{NP})\upharpoonright((\mathrm{NP}\backslash\mathrm{S})\upharpoonright\mathrm{NP})\end{array} \qquad \begin{array}{c}\vdots \quad \vdots\\ \begin{array}{l}\lambda\varphi.\mathsf{admire}\circ\varphi;\\ \mathbf{admire};\\ (\mathrm{NP}\backslash\mathrm{S})\upharpoonright\mathrm{NP}\end{array}\end{array}}{\begin{array}{l}\lambda\sigma_2\lambda\varphi.\sigma_2(\varphi)\circ\mathsf{admire}\circ\varphi;\\ \lambda g\lambda x.\mathbf{admire}(x)(g(x));(\mathrm{S}\upharpoonright\mathrm{NP})\upharpoonright(\mathrm{NP}\upharpoonright\mathrm{NP})\end{array}}}{\begin{array}{l}\lambda\varphi.\mathsf{the}\circ\mathsf{close}\circ\mathsf{friends}\circ\mathsf{of}\circ\varphi\circ\mathsf{admire}\circ\varphi;\\ \lambda x.\mathbf{admire}(x)(\iota(\lambda y.\mathbf{close\text{-}fr}(x)(y)));\mathrm{S}\upharpoonright\mathrm{NP}\end{array}}}{\begin{array}{l}\mathsf{who}\circ\mathsf{the}\circ\mathsf{close}\circ\mathsf{friends}\circ\mathsf{of}\circ\mathsf{admire};\\ \lambda Q\lambda x.Q(x)\wedge\mathbf{admire}(x)(\iota(\lambda y.\mathbf{close\text{-}fr}(x)(y)));\mathrm{N}\backslash\mathrm{N}\end{array}}$$

With this much in hand, we can now turn to the issue of how the combinatoric system interacts with processing-based interface conditions. We have five classes of data to account for (the contrast between (iv) and (v) is what crucially supports Kennedy's argument from parasitic gap-licensing patterns):

(i) unsupported *wh*-extraction from subjects (sometimes good (63); sometimes bad (62))

(ii) supported *wh*-extraction from subjects (typically fine)

(iii) embedded unsupported *wh*-extraction from subjects ((64); markedly worse than (i))

(iv) embedded supported *wh*-extraction from subjects ((60b); markedly better than (iii))

The key explanatory mechanism for (i)–(ii) 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. Cases such as (63), 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 (62)), 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 (ii) (exemplified by (65)) violate only one of the parser's expectations, and are thus predicted to be significantly better than (i).

The final issue corresponds to cases in which we have an unsupported and supported parasitic gap in an adjunct clause subject, i.e. (iii) and (iv). Here too, we have an answer ready to hand. (iii) is significantly worse than (i) since they violate 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

[24]A similar strategy for dealing with non-linearity in natural language is pursued in the cross-categorial analysis of coordination via 'generalized conjunction' by Partee and Rooth (1983) and analyses which build the distribution of reflexive pronouns into the semantics of such pronouns along the lines of Bach and Partee (1980).

the processor's parsing strategy, it is not surprising that such examples are severely ill-formed.[25] By contrast, (iv) is markedly better than (iii) because it violates only two of the three conditions. Thus, with no further development or modification, Chaves' model 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.

(67)

	gap in subject position	no gap in the following VP	embedding in adjunct clause
(i)	✓	✓	
(ii)	✓		
(iii)	✓	✓	✓
(iv)	✓		✓

Against the background of this account of parasitic gap licensing, the well-formedness and interpretation of (61), 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 (30), we identified the extraction 'site' as the direct object of the transitive auxiliary *did*. Exactly the same analysis is available in the case of (61). We have

(68)

$$
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\begin{array}{l}\vdots\quad\vdots\\ \lambda\varphi.\mathsf{to}\circ\varphi;\\ \lambda x\lambda y.\mathbf{admire}(x)(y);\\ \mathrm{VP}{\restriction}\mathrm{NP}\end{array}
\quad
\begin{bmatrix}\varphi_1;\\ u;\\ \mathrm{NP}\end{bmatrix}^1
}{
\begin{array}{l}\mathsf{to}\circ\varphi_1;\\ \lambda y.\mathbf{admire}(u)(y);\\ \mathrm{VP}\end{array}
}
\quad
\begin{array}{l}\mathsf{seem};\\ \mathbf{seem};\\ \mathrm{VP/VP}\end{array}
}{
\dfrac{
\begin{array}{l}\mathsf{seem}\circ\mathsf{to}\circ\varphi_1;\\ \lambda w.\mathbf{seem}(\mathbf{admire}(u)(w));\\ \mathrm{VP}\end{array}
}{
\begin{array}{l}\lambda\varphi_1.\mathsf{seem}\circ\mathsf{to}\circ\varphi_1;\\ \lambda u\lambda w.\mathbf{seem}(\mathbf{admire}(u)(w));\\ \mathrm{VP}{\restriction}\mathrm{NP}\end{array}
}{\restriction}\mathrm{I}^1
}
\quad
\begin{array}{l}\lambda\sigma_1\lambda\sigma_2\lambda\varphi.\sigma_2(\varphi)\circ\sigma_1(\varphi);\\ \lambda R\lambda g\lambda x.R(x)(g(x));\\ (\mathrm{S}{\restriction}\mathrm{NP}){\restriction}(\mathrm{NP}{\restriction}\mathrm{NP}){\restriction}(\mathrm{VP}{\restriction}\mathrm{NP})\end{array}
}{
\begin{array}{l}\lambda\sigma_2\lambda\varphi.\sigma_2(\varphi)\circ\mathsf{seem}\circ\mathsf{to}\circ\varphi;\\ \lambda g\lambda x.\mathbf{seem}(\mathbf{admire}(x)(g(x)));\\ (\mathrm{S}{\restriction}\mathrm{NP}){\restriction}(\mathrm{NP}{\restriction}\mathrm{NP})\end{array}
}
\quad
\begin{array}{l}\lambda\varphi_3.\mathsf{the}\circ\mathsf{close}\circ\mathsf{friends}\circ\mathsf{of}\circ\varphi_3;\\ \lambda v.\iota(\lambda z.\mathbf{close\text{-}fr}(v)(z));\\ \mathrm{NP}{\restriction}\mathrm{NP}\end{array}
}{
\begin{array}{l}\lambda\varphi.\mathsf{the}\circ\mathsf{close}\circ\mathsf{friends}\circ\mathsf{of}\circ\varphi\circ\mathsf{seem}\circ\mathsf{to}\circ\varphi;\\ \lambda x.\mathbf{seem}(\mathbf{admire}(x)(\iota(\lambda z.\mathbf{close\text{-}fr}(x)(z))));\\ \mathrm{S}{\restriction}\mathrm{NP}\end{array}
}
$$

On our analysis, cases such as (61), licensed by the derivation in (68), should have exactly the same status as their overt counterpart; like the latter, (61) displays a gap in subject position

[25] 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.

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).[26]

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 Kubota and Levine (2017) automatically yields the parallelism between the 'real' and 'deleted' host VPs for subject position parasitic gaps attested in (60b) and (61), 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)) licenses subject island violation examples such as (6a) and (60a), whose deviant status follows on independent grounds from the interactions of the factors discussed in detail in Kluender (2004) and Chaves (2013). On this view, the 'parasitism' of subject-internal gaps is not a strictly syntactic phenomenon but an emergent effect of real-time parsing strategy.[27]

4.5 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

[26] 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 type-logical grammar, 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.

[27] It should be noted however that while we believe that Chaves' 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↾XP. Correspondingly, their semantic types are different. Thus, as long as there is some way (either processing-based, as we have 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.

a difference in *syntactic* well-formedness:

(69) 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 (69a) and (69b) 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). In Appendix 1, we offer an extensive argument showing that this view is empirically superior to Kennedy and Merchant's view.

Our goal in this section, therefore, is to derive both (69a) and (69b) 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 Sect. 3.3 extends straightforwardly to these somewhat more complex attributive comparative data.

The attributive comparative construction actually comes with an additional complication: in examples like those in (69), 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 (70) (= (33a) from Sect. 3.4.1), showing that the new definition yields exactly the same truth conditions for this sentence as the older definition.

(70) Mary is taller than Ann is.

We then show that with this new definition of the comparative operator, the attributive comparative examples in (69) can be analyzed in a way fully parallel to the analysis of (70), with the only difference that the adjective is in the prenominal attributive position rather than in the predicative position.

The new comparative operator has the same syntactic type and prosodic form as the second version of the comparative operator from Sect. 3.4.1 in (38):

(71)
$$\lambda\sigma_0\lambda\sigma_1\lambda\sigma_2.\sigma_1(\sigma_0(\mathsf{er}))\circ\mathsf{than}\circ\sigma_2(\epsilon);$$
$$\lambda f\lambda\mathscr{P}\lambda\mathscr{Q}.\exists d_1,d_2.\mathscr{P}(\lambda x.\mathbf{max}(\lambda d.f(d)(x))=d_1)\wedge\mathscr{Q}(\lambda x.\mathbf{max}(\lambda d.f(d)(x))=d_2)\wedge d_1>d_2;$$
$$\mathrm{S}\upharpoonright(\mathrm{S}\upharpoonright\mathrm{Adj})\upharpoonright(\mathrm{S}\upharpoonright\mathrm{Adj})\upharpoonright(\mathrm{Adj}\upharpoonright\mathrm{Deg})$$

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 compares the two degrees thus obtained. The crucial difference from the older definition in (38) 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 (70) is identical to (39) from Sect. 3.4.1. The following translation is obtained by replacing the comparative operator in (39) with the new one in (71):

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

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 (70) and the attributive comparative example in (69a) 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:

(73)
$$\begin{array}{l}
\lambda\sigma_0\lambda\sigma_1\lambda\sigma_2.\sigma_1(\sigma_0(\mathsf{er})) \circ \mathsf{than} \circ \sigma_2(\boldsymbol{\epsilon});\\
\lambda f \lambda \mathscr{P} \lambda \mathscr{Q}.\exists d_1, d_2.\mathscr{P}(\lambda P \lambda x.\mathbf{max}(\lambda d.f(d)(P)(x)) = d_1) \wedge\\
\quad \mathscr{Q}(\lambda P \lambda x.\mathbf{max}(\lambda d.f(d)(P)(x)) = d_2) \wedge d_1 > d_2;\\
\mathrm{S}{\upharpoonright}(\mathrm{S}{\upharpoonright}(\mathrm{N/N})){\upharpoonright}(\mathrm{S}{\upharpoonright}(\mathrm{N/N})){\upharpoonright}((\mathrm{N/N}){\upharpoonright}\mathrm{Deg})
\end{array}$$

With this definition, the derivation for (69a) is straightforward. Just as in the simpler example in (70), 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 $\mathrm{S}{\upharpoonright}(\mathrm{N/N})$). Two such gapped clauses are then given as arguments to the comparative operator. The full derivation is given in (74) (here, $\mathbf{int} = \lambda d \lambda P \lambda x.P(x) \wedge \mathbf{interesting}_{d \to e \to t}(d)(x)$):

(74)

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

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

The final translation obtained can be unpacked as follows:

(75)
$$\begin{array}{l}
\exists d_1, d_2.\mathbf{\exists}_{(\lambda x.\mathbf{max}(\lambda d.\mathbf{int}(d)(\mathbf{novel})(x))=d_1)}(\lambda x.\mathbf{wrote}(x)(\mathbf{j})) \wedge\\
\quad \mathbf{\exists}_{(\lambda x.\mathbf{max}(\lambda d.\mathbf{int}(d)(\mathbf{play})(x))=d_2)}(\lambda x.\mathbf{wrote}(x)(\mathbf{m})) \wedge d_1 > d_2\\
= \exists d_1, d_2.\exists x.[\mathbf{max}(\lambda d.\mathbf{novel}(x) \wedge \mathbf{interesting}(d)(x)) = d_1 \wedge \mathbf{wrote}(x)(\mathbf{j})]\\
\quad \wedge \exists y.[\mathbf{max}(\lambda d.\mathbf{play}(y) \wedge \mathbf{interesting}(d)(y)) = d_1 \wedge \mathbf{wrote}(y)(\mathbf{m})] \wedge d_1 > d_2
\end{array}$$

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

At this point, it should be obvious that (69b) is merely a pseudogapping counterpart of (69a), 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:

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

The sign in (76) can be obtained straightforwardly in the analysis of pseudogapping sketched in Sect. 3 above. The derivation for (69b) can thus be obtained by replacing the lexical verb *wrote* in the *than* clause in (74) with the 'pseudogapped' anaphoric auxiliary *did* in (76).

Thus, fully compositional analyses for the sentences in (69) 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). The remaining issue is how to explain the degraded status of the non-elliptical variant (such as (69a)). We offer a non-syntactic, processing-oriented explanation for this pattern in Appendix 1, and argue that this non-syntactic explanation is empirically superior to Kennedy and Merchant's (2000) syntactic account of the same pattern.

5 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 informant's data, in corpora or in earlier work on pseudogapping and ellipsis generally (see Appendix 2 below). 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 proposed in Kubota and Levine (2017).

Appendix 1: The status of the attributive comparative data

As noted in Sect. 4.5, the analysis of attributive comparatives we offered above makes very different empirical predictions from the proposal in Kennedy and Merchant's (2000) work. Specifically, contrary to Kennedy and Merchant (and much of the literature following it, such as Bacskai-Atkari 2014 and LaCara 2016), our approach takes examples such as (10) to be syntactically well-formed.

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

While such examples are not impeccable, we have observed that they are not altogether ruled out, and much better examples can be found. For example, a substantial number of our informants report that examples such as (77) are simply not ill-formed at all:

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

More generally, we have noted considerable speaker variability in judgments on attributive adjective subdeletion without pseudogapping. It is true that on the whole, the pseudogapped versions are more readily acceptable, but the non-elliptical versions are noticeably improved by, for example, using higher-frequency lexica (thus, for example, replacing *writes* with *composes* materially degrades (77)).

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

Our view here is essentially in line with the 'expectation-based' processing model proposed in Chaves (2013), which assumes that the parser's performance is at least in part guided by application of the rules and constraints of the grammar.[28] To articulate this non-syntactic alternative account a bit further, in languages such as English that do not allow for the option of left branch extraction for overt material, cases in which material corresponding to an NP left branch must be 'imported' from elsewhere in the comparative structure will be initially overlooked by the parser, based on the lack of such syntactic/semantic linkages in the simpler (and more basic) cases involving overt filler-gap linkages. Instead, the parser will compose what appear to be self-evidently well-formed sentences such as *he buys beer* in (10). But once such a parse is obtained for the *than*-clause, it fails to be composable with the antecedent clause in the comparative, and this will at best force the kind of real-time backtracking that

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

registers as diminished acceptability, or even outright ill-formedness (see, e.g., Du and Yu (2012) for a computational model of processing breakdown under backtracking in garden path sentences, and Chaves (2013) for an account linking such effects to extraction possibilities).[29]

But this requirement—that the parser follow the line of least resistance, so to speak—is counteracted in exactly the class of cases in which such comparatives are unproblematic, namely, cases involving ellipsis in the *than*-clauses such as (78):

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

As already noted, what makes these examples different from their non-elliptical counterparts is that they involve an additional process of ellipsis, and there is reason to believe that this additional component has a nontrivial consequence for the real-time processing of the sentence. Recall from the above that in pseudogapping clauses, obtaining a coherent interpretation critically depends on instantiating the free variable P introduced by the ellipsis operator based on discourse context. The real-time recovery of an interpretation from the antecedent clause entailed by this anaphora resolution requirement will, by its very nature, force the processor to inspect the antecedent and ellipsis clauses simultaneously, as it were, upon encountering the 'transitive' auxiliary in cases such as (12). This interruption keeps the otherwise automatic default parsing routine from being completed. Moreover, this ellipsis resolution process reinforces the Parallelism relation between the main clause and the *than*-clause, thereby facilitating the construal of the whole sentence as a comparative construction. The extra anaphora resolution step triggered by ellipsis therefore circumvents the error-plus-backtracking sequence which the default parsing strategy leads to.

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

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

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

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

[29]If, as we have argued here, the unacceptability of (10) follows from a processing-oriented garden path effect, one might wonder why the same garden path effect does not arise in simpler comparative sentences such as (33b) from Sect. 3, whose *than*-clause is similarly parsable as a stand-alone sentence. On our analysis, the discrepancy between (33b) and cases such as (10) reflects the fact that, in the latter, at the end of parsing the *than* clause, the modifier which must be interpreted into the parse in order to yield a semantically coherent interpretation corresponds to a position in which the parser expects to find such a modifying term (as in *runs fast*). Moreover, the sentence-final position is where such a modifier is routinely interpreted in overt extraction cases, as in, e.g., *How fast did John run __?* Based on such extraction possibilities elsewhere in the language, the parser need do nothing more than identify the position following *run* as a site into which a modifier introduced earlier in the sentence can be interpreted.

The default processing routine does not yield a well-formed result on its own, and it seems reasonable to suppose that the parser, in its attempt to make sense of this otherwise uninterpretable cognate object expression, is able to make early use of the contrastive parallelism supplied by the comparative construction (which can be further facilitated by appropriate intonational cues). The amelioration effect observed in (80) provides strong support for the current processing-based view over Kennedy and Merchant's syntactic account: on the latter type of approach, since (80) does not involve any deletion operation, it is predicted to be just as unacceptable as (10), and its improved acceptability remains a total mystery.[30]

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

---

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

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

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

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

According to them, (iib) is a distinct reading corresponding to an LF (different from the one for (iia)) that is derived by a deletion operation (similar to the one involved in their analysis of attributive comparatives) that deletes not just the verb but also a prenominal adjective. However, the assumption that Gapping is derived by deletion directly contradicts the standard assumption about the analysis of Gapping in the Principle and Parameters approach: Johnson (2000, 2009) and Vicente (2010) have argued extensively that Gapping should *not* be analyzed via deletion.

We believe that a much simpler account of the apparent ambiguity of (i) is more plausible. In our view, the two 'readings' of (i) paraphrased in (ii) reflect vagueness (or underspecified interpretation) rather than true ambiguity. Specifically, (iia) is the only reading that the grammar licenses for (i) (which is immediately available in any standard analysis of Gapping). The sentence is of course compatible with a situation in which the speaker intends to convey (iib), especially in the right kind of context (for example, in a discussion of literary successes of one's friends), but the additional meaning that is felt to be present in such contexts is simply inferred as a conversational implicature (Relevance), presumably facilitated by the constructional property of Gapping that it induces a strong Parallelism discourse relation on the two conjuncts.

Appendix 2: Apparent challenge to the 'remnant-extraction' analysis of VP ellipsis/extraction interaction

The analysis of apparent extraction from elided VPs as extraction of pseudogapping remnants given above in Sect. 3.3 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 is 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.

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 (81)–(82) to support this claim:

(81) 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.

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

Since pseudogapping, as in (81), apparently cannot 'reach into' PPs to access their NP objects, whereas extraction from elided VP can, as in (82), Johnson argues that such extraction cannot be reduced to pseudogapping plus movement of the remnant.

However, Johnson's argument above 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 counterexemplify this claim.[31]

[31] 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.

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 Kubota and Levine (2017) for many such examples. In particular, examples such as (83), first documented in Levin (1979), were already known long before Johnson (2001):

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

Examples such as those in (84) seem unproblematic as well.

(84) 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. Kubota and Levine (2017) speculate on the possibility of explaining the difference between the acceptable examples such as (83) and (84) and the unacceptable ones such as (81) in terms of a complex set of conditions including factors such as prototypicality and inherent relatedness between the remnant and the elided head noun.

Note in this connection that examples that are syntactically identical in form to (81b) 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, (85), while not totally impeccable for all speakers, seems to be distinctly better than (81b):

(85) 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 (81a), the example improves, particularly if the example also takes into account the strong preference for the very marked contrast in comparatives:

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

This effect is particularly noticeable if, as suggested in (86), 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. Articulating this pragmatic felicity condition more precisely is a major remaining task in the research on pseudogapping.[32]

[32]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 (81b) and (82), the former is unacceptable while the latter is acceptable. We speculate that the improved status of (82) has to do with the fact that extraction is involved in this

Appendix 3: Comparison with other approaches to ellipsis in categorial grammar

Ellipsis has recently received a considerable amount of attention in categorial grammar research. Interestingly, there is widespread agreement in this research community not only on the suitability of a nonstructural approach to ellipsis, but—at least in broad terms—on the technical prerequisites necessary for such an approach. Two recent examples illustrate the point.

Barker (2013) offers a detailed examination of ellipsis phenomena, in particular sluicing, which is very similar to ours in several respects. Like us, Barker takes ellipsis to be an instance of anaphora which recovers a predicate of a particular type. Barker's analysis is moreover couched in a version of Type-Logical Grammar which (at a broad conceptual level) is very similar to Hybrid TLCG in employing two distinct kinds of connectives, one for ordinary combinatorics and the other for (overt and covert) 'movement' (it should, however, be kept in mind that the two systems differ significantly in the technical implementation of this general idea). Barker also uses what is, in effect, an operator that links a missing VP argument (in the case of VP ellipsis—which straightforwardly extends to the sluicing case) to the overt VP in the antecedent clause, which essentially does the same kind of work as the condition on the instantiation of the free variable in our ellipsis operator. That is, the denotation of the missing VP is required to be associated with a sign of the same *syntactic* type (modulo some minor features) in the antecedent clause.

Our proposal is also related to Jacobson's (2014; 2016) approach to ellipsis. Jacobson's (2014) analysis of VP ellipsis and pseudogapping essentially shares the same analytic ideas as ours, with the major difference being that the former is couched in a version of CCG whereas the latter is couched in a version of Type-Logical Grammar. Jacobson treats anaphoric reference via a special kind of functional type; thus, auxiliaries can optionally undergo a special rule whose effect is to take VP/VP to $\mathrm{VP}^{\mathrm{VP}}$—i.e., a term which, when a VP antecedent (indicated by the superscript) is anaphorically recovered, has the status of a VP. Just as in our approach, the meaning of the 'missing VP' is recovered in reference to some linguistic expression available (in most cases) in the previous discourse. Jacobson takes a constructional schema licensing the antecedent clause and the ellipsis clause together as a unit to be responsible for this anaphora recovery. Just like our ellipsis operator, this constructional schema is sensitive to the syntactic category of the antecedent. Jacobson (2016) extends this approach to the so-called fragment answer phenomenon, providing a detailed argument against deletion-based analyses in the literature.[33]

---

example. In particular, note that in (82), 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 (81b)) in which such a condition does not hold.

[33] We agree with Jacobson in her broader conclusion that fragment answers provide further corroboration to the idea that the syntactic information relevant in ellipsis phenomena is limited to that encoded in the syntactic category of the antecedent. However, implementing this analytic idea by means of a constructional schema (along the lines we have alluded to briefly above—see Jacobson's own papers for details) raises some non-trivial issues. And in a certain sense, the larger consequences of this type of approach seem to be even more serious in the case of question-answer pairs (which, by their very nature, are almost always an inter-speaker phenomenon) than in VP ellipsis and pseudogapping. It seems likely to become difficult to maintain this approach in cases such as the following, where the question and the answer are discontinuous (by multiple

In the conclusion of Kubota and Levine (2017), we commented on the prospects of analyses of ellipsis phenomena in categorial grammar in a positive tone, noting that it may offer an explanation for the 'partial' syntactic sensitivity of ellipsis phenomena noted independently in the wider literature on ellipsis (Chung 2013; Yoshida et al. 2015). We have acknowledged there that our optimism of course needs to be further backed up by empirical evidence. The conclusions independently reached by Barker and Jacobson in their respective studies of different types of ellipsis phenomena, together with our conclusions about the extraction/ellipsis interactions in the current paper, seem to offer some corroboration that the categorial grammar perspective is, in broad terms, indeed on the right track.

References

Bach, Emmon and Barbara Partee. 1980. Anaphora and semantic structure. In K. Kreiman and A. Ojeda, eds., *Papers From the Parasession on Pronouns and Anaphora*, 1–28. Chicago University, Chicago, Illinois: Chicago Linguistic Society.

Bacskai-Atkari, Julia. 2014. *The Syntax of Comparative Constructions: Operators, ellipsis phenomena and functional left peripheries*. Ph.D. thesis, Universitätsverlag Potsdam.

Barker, Chris. 2013. Scopability and sluicing. *Linguistics and Philosophy* 36(3):187–223.

Barros, Matthew, Patrick D. Elliott, and Gary Thoms. 2014. There is no island repair. ms., available at `ling.auf.net/lingbuzz/002100/current.pdf`.

Bresnan, Joan. 2000. *Lexical-Functional Syntax*. Oxford: Blackwell.

Chaves, Rui. 2013. An expectation-based account of subject islands and parasitism. *Journal of Linguistics* 49:285–327.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris Publications.

Chung, Sandra. 2013. Syntactic identity in sluicing: How much, and why. *Linguistic Inquiry* 44(1):1–44.

Cooper, Robin. 1979. The interpretation of pronouns. In F. Heny and H. S. Schnelle, eds., *Selections from the Third Groningen Round Table*, vol. 10 of *Syntax and Semantics*, 61–92. New York: Academic Press.

Cormack, Annabel. 1984. VP anaphora: variables and scope. In F. Landman and F. Veltman, eds., *Varieties of Formal Semantics*, 81–102. Dordrecht: Foris Publications.

Cresswell, Max J. 1976. The semantics of degree. In B. H. Partee, ed., *Montague Grammar*, 261–292. New York: Academic Press.

Culicover, Peter. 1999. *Syntactic Nuts: Hard Cases, Syntactic Theory and Language Acquisition*. Oxford: OUP.

---

sentences):

(i) **Q:** What did John have for lunch?
A: Hmmm. Good question. John's a vegetarian, right? Well, surprise, surprise—a burger and fries!

There are alternative approaches such as Yasavul (2017), which treat question-answer pairs in terms of a dynamic updating of discourse context. In such an approach, the challenge essentially comes down to devising some way of making the pragmatic process of discourse update partly sensitive to the syntactic category of the relevant antecedent expression. If this can be implemented in some way or other, a dynamic approach may offer the necessary flexibility to handle more complex case like (i) while retaining the essential insight of Jacobson's approach.

de Groote, Philippe. 2001. Towards abstract categorial grammars. In *Association for Computational Linguistics, 39th Annual Meeting and 10th Conference of the European Chapter*, 148–155.

Du, Jiali and Pingfang Yu. 2012. Predicting garden path sentences based on natural language understanding systems. *International Journal of Advanced Computer Science and Applications* 3(11):1–6.

Elbourne, Paul. 2008. Ellipsis sites as definite descriptions. *Linguistic Inquiry* 39(2):191–220.

Engdahl, Elisabet. 1985. Parasitic gaps, resumptive pronouns, and subject extractions. *Linguistics* 23:3–44.

Evans, Frederic. 1988. Binding into anaphoric verb phrases. In *Proceedings of ESCOL (East Coast Conference on Linguistics)*. Working Papers in Linguistics, Ohio State University.

Haïk, Isabelle. 1987. Bound VPs that need to be. *Linguistics and Philosophy* 10(4):503–530.

Hardt, Daniel. 1999. Dynamic interpretation of verb phrase ellipsis. *Linguistics and Philosophy* 22:187–221.

Hellan, Lars. 1981. *Towards an Integrated Analysis of Comparatives*. Tübingen: Narr.

Hendriks, Petra. 1995. *Comparatives and Categorial Grammar*. Ph.D. thesis, University of Groningen.

Hofmeister, Philip, Inbal Arnon, T. Florian Jaeger, Ivan Sag, and Neal Snider. 2013. The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes* 28:48–87.

Huddleston, Rodney, Geoffrey K. Pullum, and Peter Peterson. 2002. Relative clause constructions and unbounded dependencies. In R. Huddleston and G. K. Pullum, eds., *The Cambridge Grammar of the English Language*, 1031–1096. Cambridge, England: CUP.

Jacobson, Pauline. 1998. Antecedent contained deletion and pied-piping: Evidence for a variable-free semantics. In D. Strolovitch and A. Lawson, eds., *Proceedings from Semantics and Linguistic Theory VIII*, 74–91. Ithaca, New York: Cornell University.

Jacobson, Pauline. 2007. Direct compositionality and variable-free semantics: The case of "Principle B" effects. In C. Barker and P. Jacobson, eds., *Direct Compositionality*, 191–236. Oxford: OUP.

Jacobson, Pauline. 2014. Ellipsis in categorial grammar. In J. van Craenenbroeck and T. Temmerman, eds., *Handbook of Ellipsis*. Oxford: OUP.

Jacobson, Pauline. 2016. The short answer: implications for direct compositionality (and vice versa). *Language* 92(2):331–375.

Jäger, Gerhard. 2005. *Anaphora and Type-Logical Grammar*. Berlin: Springer.

Johnson, Kyle. 2000. Few dogs eat Whiskas or cats Alpo. In K. Kusumoto and E. Villalta, eds., *University of Massachusetts Occasional Papers*, vol. 23, 47–60. Amherst, MA: GLSA Publications.

Johnson, Kyle. 2001. What VP ellipsis can do, and what it can't, but not why. In M. Baltin and C. Collins, eds., *The handbook of contemporary syntactic theory*, 439–479. Oxford: Blackwell Publishers.

Johnson, Kyle. 2009. Gapping is not (VP-) ellipsis. *Linguistic Inquiry* 40(2):289–328.

Kamp, Hans. 1975. Two theories of adjectives. In E. Keenan, ed., *Formal Semantics of Natural Language*, 123–155. Cambridge: CUP.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy*

1(1):3–44.

Kennedy, Christopher. 1994. Argument contained ellipsis. UCSC Linguistics Reseach Center report LRC-94–03, reprinted in Kyle Johnson (ed.) 2008. *Topics in Ellipsis* (CUP), pp. 95–131.

Kennedy, Christopher. 2003. Ellipsis and syntactic representation. In K. Schwabe and S. Winkler, eds., *The Interfaces: Deriving and Interpreting Omitted Structures*, 29–53. Amsterdam: John Benjamins.

Kennedy, Christopher. 2005. Comparatives, semantics of. In *Encyclopedia of Language and Linguistics (2nd edition)*, 690–694. Oxford, UK: Elsevier.

Kennedy, Chris and Jason Merchant. 2000. Attributive comparative deletion. *Natural Language and Linguistic Theory* 18(1):89–146.

Kluender, Robert. 1998. On the distinction between strong and weak islands: A processing perspective. In P. Culicover and L. McNally, eds., *The Limits of Syntax*, vol. 29 of *Syntax and Semantics*, 241–279. San Diego: Academic Press.

Kluender, Robert. 2004. Are subject islands subject to a processing account? In *Proceedings of WCCFL 23*, 101–125. Somerville, MA: Cascadilla Press.

Kubota, Yusuke and Robert Levine. 2014. Hybrid Type-Logical Categorial Grammar. MS., University of Tsukuba and Ohio State University, available at `http://ling.auf.net/lingbuzz/002313`.

Kubota, Yusuke and Robert Levine. 2015. Against ellipsis: Arguments for the direct licensing of 'non-canonical' coordinations. *Linguistics and Philosophy* 38(6):521–576.

Kubota, Yusuke and Robert Levine. 2016a. Gapping as hypothetical reasoning. *Natural Language and Linguistic Theory* 34(1):107–156.

Kubota, Yusuke and Robert Levine. 2016b. The syntax-semantics interface of 'respective' predication: A unified analysis in Hybrid Type-Logical Categorial Grammar. *Natural Language and Linguistic Theory* 34(3):911–973.

Kubota, Yusuke and Robert Levine. 2017. Pseudogapping as pseudo-VP ellipsis. *Linguistic Inquiry* 48(2):213–257.

LaCara, Nicholas J. 2016. *Anaphora, Inversion and Focus*. Ph.D. thesis, University of Massachusetts, Amherst.

Lambek, Joachim. 1958. The mathematics of sentence structure. *American Mathematical Monthly* 65(3):154–170.

Lau, Ellen, Clare Stroud, Silke Plesch, and Colin Phillips. 2006. The role of structural prediction in rapid syntactic analysis. *Brain and Language* 98(1):74–88.

Levin, Nancy S. 1979. *Main Verb Ellipsis in Spoken English*. Ph.D. thesis, Ohio State University.

Levine, Robert D. and Thomas E. Hukari. 2006. *The Unity of Unbounded Dependency Constructions*. Stanford, California: CSLI Publications.

Levine, Robert D. and Ivan Sag. 2003. Wh-nonmovement. *Gengo Kenkyu: Journal of the Linguistic Society of Japan* 123:171–219.

Merchant, Jason. 2015. On ineffable predicates: Bilingual Greek-English code-switching under ellipsis. *Lingua* 166:199–213.

Michel, Daniel. 2014. *Individual Cognitive Measures and Working Memory Accounts of Syntactic Island Phenomena*. Ph.D. thesis, University of California, San Diego.

Miller, Philip. 1990. Pseudogapping and *do so* substitution. In *Proceedings of CLS 26*, vol. 1, 293–305. Chicago: Chicago Linguistic Society.

Miller, Philip. 2014. A corpus study of pseudogapping and its theoretical consequences. In C. Piñón, ed., *Empirical Issues in Syntax and Semantics*, vol. 10, 73–90.

Morrill, Glyn. 2010. *Categorial Grammar: Logical Syntax, Semantics, and Processing*. Oxford: OUP.

Morrill, Glyn. 2017. Grammar logicised: relativisation. *Linguistics and Philosophy* 40(2):119–163.

Muskens, Reinhard. 2003. Language, lambdas, and logic. In G.-J. Kruijff and R. Oehrle, eds., *Resource Sensitivity in Binding and Anaphora*, 23–54. Dordrecht: Kluwer.

Muskens, Reinhard. 2007. Separating syntax and combinatorics in categorial grammar. *Research on Language and Computation* 5(3):267–285.

Oehrle, Richard T. 1994. Term-labeled categorial type systems. *Linguistics and Philosophy* 17(6):633–678.

Partee, Barbara and Mats Rooth. 1983. Generalized conjunction and type ambiguity. In R. Bäuerle, C. Schwarze, and A. von Stechow, eds., *Meaning, Use, and Interpretation of Language*, 361–383. Berlin: Walter de Gruyter.

Pickering, Martin. 1993. Direct association and sentence processing: A reply to Gorrell and to Gibson and Hickok. *Language and Cognitive Processes* 8:163–196.

Pickering, Martin and Guy Barry. 1991. Sentence processing without empty categories. *Language and Cognitive Processes* 6:229–264.

Pollard, Carl. 2013. Linear categorial grammar. ms., Lecture Notes at ESSLLI 2013.

Pollard, Carl and Ivan Sag. 1983. Reflexives and reciprocals in English: An alternative to the binding theory. In *Proceedings, Second West Coast Conference on Formal Linguistics*, 189–203.

Pollard, Carl and Ivan A. Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23(2):261–303.

Pollard, Carl J. and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. Studies in Contemporary Linguistics. Chicago and London: University of Chicago Press.

Postal, Paul M. 1971. *Cross-over Phenomena*. New York: Holt, Rinehart & Winston.

Postal, Paul M. 2004. *Skeptical Linguistic Essays*. Oxford: OUP.

Reinhart, Tanya and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24(4):657–720.

Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Ph.D. thesis, MIT. Reproduced by the Indiana University Linguistics Club.

Ross, John Robert. 1969. Guess who? In R. I. Binnick, A. Davison, G. M. Green, and J. L. Morgan, eds., *Chicago Linguistic Society*, 252–286. Chicago, Illinois.

Safir, Kenneth. 2014. *The Syntax of Anaphora*. Oxford: OUP.

Schachter, Paul. 1978. English propredicates. *Linguistic Analysis* 4(3):187–224.

Steedman, Mark. 1985. Dependency and coordination in the grammar of Dutch and English. *Language* 61(3):523–568.

Steedman, Mark. 1987. Combinatory grammars and parasitic gaps. *Natural Language and Linguistic Theory* 5(3):403–439.

Steedman, Mark. 2000. *The Syntactic Process*. Cambridge, Mass.: MIT Press.

Thoms, Gary. 2016. Short answers in Scottish Gaelic and their theoretical implications.

Natural Language and Linguistic Theory 34(1):351–391.

Toosarvandani, Maziar. 2016. Embedding the antecedent in gapping: Low coordination and the role of parallelism no access. *Linguistic Inquiry* 47(2):381–390.

Truswell, Robert. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117(8):1355–1377.

Vicente, Luis. 2010. A note on the movement analysis of gapping. *Linguistic Inquiry* 41(3):509–517.

von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3:1–77.

Wagers, Matthew W. and Colin Phillips. 2009. Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics* 45(2):395–433.

Yasavul, Murat. 2017. *Questions and Answers in K'iche'*. Ph.D. thesis, Ohio State University.

Yoshida, Masaya, Tim Hunter, and Michael Frazier. 2015. Parasitic gaps licensed by elided syntactic structure. *Natural Language and Linguistic Theory* 33(4):1439–1471.