Ellipsis, economy and the (non)uniformity of traces*

Troy Messick
University of Connecticut

Gary Thoms
Newcastle University

July 21, 2015

1 Introduction

A number of recent articles have analyzed the interactions between ellipsis and movement in terms of *MaxElide*, an economy constraint which ensures that the biggest deletable constituent is elided within a given domain (e.g., Merchant 2008, Takahashi and Fox 2005, Hartman 2011). In this remark, we show that eliminating MaxElide allows us to provide an empirically superior account of restrictions on extraction from ellipsis. We show that a number of the core effects attributed to MaxElide can be explained in terms of parallelism alone; then, picking up on the proposal in Fox and Lasnik (2003) that movement may proceed in one fell swoop in certain elliptical contexts, we argue that the core MaxElide effects follow from more general conditions on the economy of derivations, with ellipsis derivations sometimes requiring fewer steps of movement than others. The picture this leaves us with is one with no need for the ellipsis-specific economy constraint MaxElide. Importantly, the analysis we arrive at makes reference only to traces left by A′-movement and head movement, with no place for A-movement. We argue that this fits with previous work on reconstruction, which has shown that A-movement often seems not to leave a trace (Chomsky 1995, Lasnik 1998, Fox 1999a, Takahashi and Hulsey 2009). Our analysis makes crucial use of a syntactic parallelism condition like the one proposed by Fiengo and May (1994) and Griffiths and Lipták (2014), since the semantic condition proves insufficiently restrictive.

2 Previous accounts

We begin by outlining the important details of Takahashi and Fox (2005) account of MaxElide and Hartman (2011)'s extension of it.[1] The core data which analyses of MaxElide aim to account for is in (1), where VP-ellipsis seems to be blocked when sluicing would also be possible in the same configuration.

(1) Mary was kissing someone, but I don't know who (*she was).

Takahashi and Fox (2005) proposes an account in terms of *ellipsis parallelism*, an identity condition on ellipsis, and *MaxElide*, a constraint which favours deletion of the largest constituent possible; crucially, the domain of application of MaxElide is within a given domain defined by the parallelism constraint, called a *parallelism domain* (PD). The ellipsis parallelism condition is given in (2)-(3), and MaxElide in (4).

*To appear in *Linguistic Inquiry*. For helpful discussion of the topics presented here, we would like to thank Matt Barros, Željko Bošković, Patrick Elliot, Jon Gajewski, Craig Sailor and Susi Wurmbrand. Thanks also to two anonymous reviewers for helpful comments and suggestions. All remaining errors are ours.

[1] Merchant (2008) has a slightly different formulation of MaxElide than Takahashi and Fox (2005) and Hartman (2011); however many of the problems outline here for Takahashi and Fox (2005) and Hartman (2011) are inherited by Merchant's analysis as well.

(2) For ellipsis of EC [elided constituent] to be licensed, there must exist a constituent, which reflexively dominates EC, and satisfies the parallelism condition in (3). [Call this constituent the parallelism domain (PD).]

(3) *Parallelism*: PD satisfies the parallelism condition if PD is semantically identical to another constituent AC, modulo focus-marked constituents.

(4) MAXELIDE: elide the biggest deletable constituent reflexively dominated by the PD.

Together with the assumption that *wh*-traces are interpreted as bound variables, this set of constraints derives (1): the variable in the object position ensures that the smallest possible PD is the constituent immediately dominated by its binder, since a free variable could not be semantically identical to the corresponding element in AC,[2] and applying MaxElide in this domain will ensure that the biggest deletable constituent is elided; in this PD, both VP-ellipsis and sluicing (TP-ellipsis) are possible in principle, so MaxElide chooses sluicing and VP-ellipsis is blocked. This is schematized for (1) in (5).

(5) Mary was kissing someone, but I don't know who (*she was).
someone [λy. Mary was [$_{\text{VP}}$ kissing y]] ... who [λx she was [$_{\text{VP}}$ kissing x]]
$\underbrace{\qquad\qquad}_{AC}$ $\underbrace{\qquad\qquad}_{EC}$

Thus on this analysis VP-ellipsis is blocked because it is in competition with sluicing.

There are a few important aspects of this system that we need to outline before proceeding. First, the competition-based account predicts that VP-ellipsis should be possible when sluicing is ruled out. Takahashi and Fox (2005) note that this seems to be correct, since examples like (1) can be rendered grammatical if either the subject or the auxiliary is focussed, as shown by (6); assuming that focus cannot be deleted, due to its prosodic prominence, this renders TP-ellipsis impossible and thus application of MaxElide predicts the availability of VP-ellipsis.

(6) a. I don't know who JOHN will kiss, but I know who SUSAN will.
b. Mary doesn't know who we can invite, but she knows who we CANNOT.

Second, MaxElide only predicts there to be competition between different ellipsis options in a given PD, so it does not force ellipsis of a larger constituent whenever it is possible. Thus examples like (7) are predicted to be good even though ellipsis of the matrix VP is possible; this is because the lower VP is itself a PD, and so (2)-(3) are satisfied when it is elided.

(7) Mary [$_{\text{VP}}$ said you would [$_{\text{VP}}$ leave]], and Sue also [$_{\text{VP}}$ said you would [$_{\text{VP}}$ ~~leave~~]].

For the most part the conditions in (2)-(4) only have an effect in cases like (1) where there is a variable which is bound from outside a constituent which is a potential target for ellipsis, a configuration which Takahashi and Fox (2005) call *rebinding*. Whether or not MaxElide applies in a given domain thus depends on the distribution of variables and their binders, and this in turn depends on which movement dependencies are posited to leave variables which could refine rebinding configurations.[3]

[2] As Takahashi and Fox (2005) and Hartman (2011) note, this requires assuming that something like Heim's (1997) "no meaningless coindexation" ban holds to prevent the variables in EC and AC being accidentally assigned the same index.

[3] Takahashi and Fox (2005) attempt to account for certain restrictions on the availability of sloppy readings of pronouns in terms of MaxElide, the data they attempt to account for are given below in (i).

(i) a. John λx. said that Mary hit him$_x$, and Bill λy. did ~~say that she hit him$_y$~~ too. ✓strict / ✓sloppy
b. John λx. said that Mary hit him$_x$, and Bill λy. said she did ~~hit him$_y$~~ too. ✓strict / * sloppy

Takahashi and Fox (2005) assume that in order for the pronoun in the ellipsis site to be interpreted as sloppy, it must be bound by the λ-abstraction that composes with the subject (i.e., *John* in the AC and in *Bill* in the EC), thus this creates a re-binding configuration just as wh-movement does in (1). MaxElide would then choose the largest ellipsis target, hence ruling out the sloppy interpretation in (ib). This type of analysis predicts that if intervening focus blocked ellipsis of the larger VP, then MaxElide

Hartman (2011) extends this system by using the distribution of MaxElide effects – that is, where sluicing blocks VP-ellipsis – to diagnose the distribution of rebinding configurations, and considering the wider data set he concludes that traces of all kinds of movement must leave variables which count for the calculation of parallelism. The key data comes from *wh*-adverbial questions. Building on Schuyler (2001) and Merchant (2008), Hartman observes that MaxElide effects are not observed with embedded *wh*-adverbial questions like (8), but they are observed in *matrix wh*-adverbial questions like (9), and he provides evidence from dialectal variation in English to indicate that the crucial difference between (8)-(9) is the T-to-C movement in the latter. But embedded *wh*-adverbial questions do show MaxElide effects when the adverbial is extracted from within the elided VP; that is, (10b), involving VP-ellipsis, is unacceptable on the reading where the question is about the time of leaving, where the adverbial is extracted from the lowest clause. The sluicing example in (10a), on the other hand, allows this reading.

(8) You say you'll pay me back, but you haven't told me when (you will).

(9) We know Anna is going to resign. The only question is: when (*will she)?

(10) a. John said Mary would leave, but I forget when. ✓matrix reading / ✓embedded reading
b. John said Mary would leave, but I forget when he did. ✓matrix reading / *embedded reading

The data in (10a) follows from the fact that the *wh*-adverbial is extracted from the VP which is the target for ellipsis; this makes the VP a rebinding configuration much like (1), and so the smallest possible PD is the one containing the binder of the *wh*-adverbial in Spec,CP. A simplified schematic is given in (11), with the base position of the *wh*-adverbial given as a TP-adjunction position.

(11) $[_{CP}$ when $\underline{\lambda x\ [_{TP}$ John $[_{VP}$ said $[_{CP}\ [_{TP}$ x $[_{TP}$ he $[_{T'}$ would $[_{VP}$ leave $]]]]]]]]$

Applying MaxElide here will derive sluicing and block VP-ellipsis, just like it does for *wh*-object questions, which also involve extraction from VP.

The account of (8)-(9) is more interesting. For (8), Hartman proposes the analysis of the EC as in (12): the *wh*-adverbial is generated as an adjunct to TP, and so the binder of the trace of the raised subject demarcates a PD in which MaxElide can apply to derive VP-ellipsis (the smallest possible PD is underlined). The analysis for (9) (given in (13)) is broadly similar but with one crucial difference: the auxiliary moves from T to C and leaves a variable in TP which is rebound from C′. The interleaving of binding paths leads to a situation where the smallest PD is the one demarcated by the binder of the *wh*-trace. Application of MaxElide to this domain derives sluicing, as a result of which VP-ellipsis is blocked.

(12) $[_{CP}$ when $\lambda x\ [_{TP}$ x $[_{TP}$ you $\underline{\lambda y\ [_{T'}$ will $[_{VP}$ y pay me back $]]]]]$

(13) $[_{CP}$ when $\underline{\lambda x\ [_{C'}$ will $\lambda y\ [_{TP}$ x $[_{TP}$ she $\lambda z\ [_{T'}$ y $[_{VP}$ z resign $]]]]]]$

Importantly, all of the different trace types are implicated in (13): without representation of the A-trace, the VP would be a potential PD and applying MaxElide in this domain would incorrectly derive VP-ellipsis as an option for (9); without representation of the verb movement trace (12) and (13) would be indistinguishable and again we would incorrectly predict VP-ellipsis to be possible in (13) just like it is in (12).

---

would be forced to choose ellipsis of the smaller VP and the sloppy interpretation should be possible in (ib) just as we saw the ameliorating effect of focus in (6) above. As (ii) shows, this prediction is incorrect, the sloppy interpretation is still impossible here, suggesting that whatever is blocking the sloppy interpretation, it is not MaxElide (see also Hardt 2006 for similar examples that lack a sloppy interpretation).

(ii) John said that Mary hit him, and Bill said that she DIDN'T ~~hit him~~. ✓strict / * sloppy

Hardt (2006) and Grant (2008) present even more evidence against a MaxElide account of the contrast between (ia) and (ib), for these reasons, we limit ourselves to discussion of configurations involving movement and leave the constraints on the availability of sloppy interpretations as a topic of future research.

3 Further restrictions on extraction from VP-ellipsis

Hartman's article closes by reflecting on a puzzle for the analysis from the role of intervening focus. Recall that VP-ellipsis was said to be grammatical in the examples in (6) because the presence of focus in the IP-domain ruled out sluicing as an ellipsis option, and so application of MaxElide to the PD (which encompassed the scope of the operator in Spec,CP) yielded deletion of VP, since this was the largest constituent which could be licitly deleted. The prediction of Hartman's account is that this ought to hold with all the other cases where sluicing would otherwise block VP-ellipsis. Hartman shows that this prediction is borne out with matrix *wh*-adverbial questions; this is shown in (14), with (14b) being an only slightly altered version of (9).

(14) a. Mary woke up at 7:00. When did JOHN?
b. If Anna isn't going to resign today, then when WILL she?

However, as Hartman himself notes, this account runs into trouble with matrix *wh*-object questions,[4] as these seem to show no amelioration effect with intervening focus, as shown by (15) (note that examples like this also disallow VP-ellipsis without intervening focus).

(15) a. Mary will kiss Bill. Who will JOHN *(kiss)?
b. If you aren't drinking water, then what ARE you *(drinking)?

Hartman then observes that matrix *wh*-adverbial questions show the same behaviour when they are extracted from VP (as with (10) above): intervening focus in the IP-domain seems not to save the VP-ellipsis option with the low construal, as shown by (16).

(16) a. A: John will ask Mary to leave at 5.
B: When will TOM? ✓matrix / *embedded
b. If John won't ask Mary to leave at 5, then when WILL he? ✓matrix / *embedded

This seems to indicate that matrix extractions from VP cannot be ameliorated by intervening focus, contrary to the predictions of Hartman's analysis of these data.

Here we add a number of observations which further compound the problems for Hartman's MaxElide account. First, we observe that the problem with matrix *wh*-object questions seen in (15) has the character of a parallelism violation, yet this is not expected on Hartman's approach. To begin with, we can see from (17) that VP-ellipsis *is* possible with matrix *wh*-object questions when there is parallel *wh*-movement and head movement in the AC:

(17) a. A: What's he told you?
B: What HASN'T he?[5]
b. Who will Bill kiss, and who will JOHN?

This tells us that there is no fundamental incompatibility between matrix *wh*-object extraction and VP-ellipsis, and one interpretation of the facts is that the problem with (15) is that parallelism is violated, and that this is alleviated by overt parallel extraction in (17). That this is a plausible analysis can be seen by considering the schematic in (18), which represents (15a); here we assume that focussed DPs undergo

[4] An anonymous reviewer notes that they do not find Hartman's (15a) wholly ungrammatical, and that other speakers consulted felt similarly. We believe that this may be due to a potential ambiguity with this example (also noted by the same reviewer), where *who* is the subject, *John* the object and the verb removed by pseudogapping. This ought to be controlled for with examples like (i), which we find worse than (15a):

(i) Mary is eating cake. What is JOHN *(eating)?

Note also that the same problem does not trouble (15b).

[5] Based on attested example at http://www.fanfiction.net/s/4163642/16/Death, accessed June 17th 2014.

Quantifier Raising (QR, Chomsky 1976, Krifka 2006), and in this case we take it that the object *John* QRs to Spec,CP to take scope in parallel to the *wh*-phrase in the EC.

(18) AC: [$_{CP}$ John λx [$_{C'}$ [$_{TP}$ Mary λz [$_{T'}$ will [$_{VP}$ z kiss x]]]]]
EC: [$_{CP}$ who $\underline{\lambda x\ [_{C'}\ will\ \lambda y\ [_{TP}\ John\ \lambda z\ [_{T'}\ y\ [_{VP}\ z\ kiss\ x\ ]]]]]}$

We can see here that the putative PD in EC contains a lambda operator which binds the variable left by T-to-C movement, but this is not matched by a similar binding relation in AC; as such, AC and EC are not strictly parallel with respect to the position of variables and their binders. But this is of no consequence for Hartman's theory as formulated, since he adopts a semantic parallelism condition which would not distinguish the underlined constituents as required. Note however that we would be able to explain (15) as parallelism violations if we were to cast parallelism as an LF-isomorphism constraint, such as the one proposed by Griffiths and Lipták (2014)

One might propose to repair Hartman's theory by bolstering it with an LF-isomorphism-based parallelism constraint. However this would then lead to the incorrect prediction that VP-ellipsis with matrix *wh*-adverbial questions would also be ruled out as parallelism violations. Consider again the schematic for (9), this time presented side-by-side with the AC. Here we represent the correlate of the *wh*-phrase in AC as a covert indefinite roughly equivalent to "at some time," which raises covertly to CP to take scope in parallel to the *wh*-phrase (see Chung et al. 1995, Merchant 2001 for discussion of implicit indefinite correlates in sluicing).

(19) AC: [$_{CP}$ at-some-time λx [$_{C'}$ [$_{TP}$ x [$_{TP}$ she λz [$_{T'}$ will [$_{VP}$ z resign]]]]]]
EC: [$_{CP}$ when $\underline{\lambda x\ [_{C'}\ will\ \lambda y\ [_{TP}\ x\ [_{TP}\ she\ \lambda z\ [_{T'}\ y\ [_{VP}\ z\ resign\ ]]]]]]}$

As with (18), here we can see that the underlined PD in EC is not identical to that in AC, since there is no T-to-C movement in AC. If parallelism were indeed syntactic, and the schematic here correctly represented the LF for (9), then these examples would be predicted to be ungrammatical, contrary to fact. In what follows we will argue that it's the schematic in (19) that is incorrect, while the syntactic parallelism constraint has an empirical advantage over the semantic one when it comes to explaining the restrictions on extraction from VP-ellipsis we have observed. That is, we will argue that it is indeed correct to analyse the matrix *wh*-object question examples as parallelism violations, while the ungrammaticality of matrix *wh*-adverbial questions like (9) (where there is no intervening focus to save the VP-ellipsis option) is to be explained by different means. This takes us in the direction of a non-uniform analysis of restrictions on extraction from VP-ellipsis, with those cases where intervening focus has no effect on the grammaticality of VP-ellipsis having a parallelism-based analysis and those cases where intervening focus does have an effect being explained in another way. This would group the original cases from Takahashi and Fox (2005) in (6) with the matrix *wh*-adverbials in (9) and the cases in (17) above, while separating them from cases like (15)-(16). For terminological convenience, we will henceforth refer to the class of cases where intervening focus renders the VP-ellipsis option in a rebinding configuration grammatical as *salvageable*, picking up on the loose intuition that focus somehow "salvages" the VP-ellipsis option.[6] We will therefore call the other class *unsalvageable*.

A second problem comes from consideration of the role of successive-cyclic movement. Hartman assumes (2011, 374, fn.11) that each step of successive-cyclic movement creates a new binder, and he argues that the fact that there is no competition between high and low VP-ellipsis in raising clauses like (20) provides evidence for successive-cyclic A-movement through the embedded TP, as this would ensure that there is a PD in the infinitival complement, as in (21) (p.384).

(20) John is likely to attend, and Bill is (likely to), also.

[6]We use the term "salvage" to distinguish this effect from *repair*, which has a technical sense in the literature which we wish to avoid. In the account that follows, as in that of Merchant (2008), Takahashi and Fox (2005) and Hartman (2011), focus does not actually repair a "broken" extraction, but rather it rules out a competing, more economical option

(21) Bill λx is likely [$_{TP}$ x λx′ to [$_{VP}$ x′ attend]]

While this works for A-movement, it runs into a number of problems when it is applied to A′-movement (which has been the primary source of evidence for successive-cyclicity in the literature to date). For instance, if we assume that long-distance *wh*-movement passes through *v*P and CP (Chomsky 1986, 2000, Fox 1999a, van Urk and Richards 2015), then we predict that VP-ellipsis ought to be possible in an embedded clause whenever there is extraction from that clause, since the lambda operator introduced by successive-cyclic movement through the embedded CP would create a PD in which the application of MaxElide would derive VP-ellipsis. (22) shows that the prediction is incorrect for object extractions, and (23) provides an illustrative LF, where the PD in the embedded clause is underlined; since sluicing is not an option in the intermediate position, as shown in (24), applying MaxElide in this PD would incorrectly derive VP-ellipsis.

(22) *John said you spoke to someone, but I don't know who he said you did.

(23) [$_{CP}$ wh λx ... [$_{CP}$ x λx′ ... [$_{VP}$ V x′]]]

(24) *John said you spoke to someone, but I don't know who he said [$_{CP}$ t ~~[you spoke to t]~~]

Interestingly, the same effect can be seen with embedded subject extractions, (25). This is doubly surprising for the MaxElide approach, since not only does *wh*-movement through the embedded Spec,CP create a PD, A-movement of the subject to the embedded Spec,TP ought to do so as well, as (26) illustrates.

(25) *John thinks one of the teachers is leaving, but I don't know which one he thinks is.

(26) [$_{CP}$ which one λx he thinks [$_{CP}$ x λx′ [$_{TP}$ x′ λx″ is [$_{VP}$ x″ leaving]]]]

This indicates that sluicing blocks VP-ellipsis in a wider set of situations than can be defined in terms of PDs as in the MaxElide approach.

Related to this, a third problem is that various kinds of non-local extractions from VP-ellipsis are unsalvageable. First, consider extraction from embedded clauses. Lasnik and Park (2013, 240) observe that VP-ellipsis is not possible when there is long-distance extraction from a clause contained within the ellipis site, as shown by (27a); note that this is superficially similar to the regular object extractions cases like (6), where intervening focus has a salvaging effect. (28) shows that such extractions from VP-ellipsis are in fact possible when there is overt parallel extraction from AC.

(27) a. *Abby said they heard about a Balkan language, but I don't know what kind of language BEN did.
b. *John thinks you should kiss SARAH, but I don't know who BILL does.

(28) I know who JOHN thinks you should kiss, but I don't know who BILL does.

(29) shows that non-parallel extraction of *wh*-adverbials from embedded clauses is unsalvageable as well, with (30) again showing that overt parallel extraction cases are different. (As expected, the matrix readings for the adjuncts are still available here.)

(29) JOHN said Mary would leave at noon, but I forget when BILL did. ✓matrix / *embedded

(30) I know when JOHN said Mary left, but I don't know when BILL did. ✓matrix / ✓embedded

The generalization here seems to be that non-parallel extraction from VP-ellipsis is only salvageable if it does not cross a clause boundary. This is not quite right, as non-parallel extraction is salvageable from at least some non-finite clausal complements, such as control complements.

(31) ?John WILL try to kiss MARY, but I don't know who he WON'T.

Thus the generalization may seem to be that non-parallel extraction from *finite* clause is unsalvageable.

Interestingly, though, non-parallel extraction from VP-ellipsis is subject to a number of other restrictions which do not involve finite clause boundaries. For instance, Lasnik and Park (2013) observes that subextraction from a DP in the object position in VP-ellipsis is unsalvageable, even when the object is contained in the highest VP in the ellipsis site, (32); (33) again shows that overt parallel extraction is different.

(32) *ABBY heard a lecture about a Balkan language, but I don't know what kind of language BEN did.

(33) What did ABBY hear a lecture about, and what did BEN?

Similarly, non-parallel extraction from VP-ellipsis seems to be unsalvageable with certain kinds of *wh*-phrases, such as degree *wh*-phrases like *how upset*. Once more, the extraction from VP-ellipsis is much better if there is overt parallel extraction in AC (cf. Baltin 2011).

(34) *John became very upset, but I don't know how upset BILL did.

(35) ?I know how upset JOHN became, but I don't know how upset MARY did.

Thus salvageable non-parallel extraction from VP-ellipsis seems to be restricted to a small set of local DP-extractions. This does not follow from the MaxElide account as things stand, and it indicates that we need to rethink exactly what the key factors are which make extraction from VP-ellipsis so limited.

Finally, we observe that there are also cases not involving A′-extraction where the MaxElide theory incorrectly predicts that ellipsis of a larger category should block ellipsis of a smaller one. In particular, there are those simple cases where VP-ellipsis can optionally include non-finite auxiliaries:

(36) a. John has been singing, and Mary has (been), too.
b. John shouldn't be drinking, and Mary shouldn't (be), either.

In both of these cases the MaxElide-based account incorrectly predicts that ellipsis of the larger constituent containing non-finite *be* should block ellipsis of the smaller constituent. To see why this is the case, consider the following simplified schematic of the ellipsis clause in (36a) (ignoring movement of auxiliaries, which is immaterial here). As before, the smallest possible PD is underlined; applying MaxElide to this domain would derive VP-ellipsis of everything up to and including non-finite *been*, with the option of retaining *been* being blocked.

(37) $[_{TP}$ Mary $\underline{\lambda x\ [_{T'}\ \text{has}\ [_{vP}\ \text{been}\ [_{VP}\ x\ \text{singing}\ ]]]]}$

One might expect that this can be accounted for by breaking down the A-chain formed by movement of the subject into a number of short intermediate steps, in the spirit of Hartman's account of the optionality of VP-ellipsis with raising structures like (20). Thus one might propose that the correct structure is not (37) but (38), where A-movement passes through all the projections in the inflectional layer, including some other projection below *been*, say VoiceP; applying MaxElide to the underlined PD would derive the option for smaller ellipsis, as desired.

(38) $[_{TP}$ Mary $\lambda x\ [_{T'}\ \text{has}\ [_{vP}\ x\ \lambda x'\ [_{v'}\ \text{been}\ [_{VOICEP}\ x'\ \underline{\lambda x''\ [_{VP}\ x''\ \text{singing}\ ]]]]]]$

While this would work for these cases, it would undermine the analysis of the crucial *wh*-adverbial data in (9), as schematized in (12) above. That is, allowing adjunction to intermediate projections would mean that it ought to be possible to analyse (12) as (39), and this would make the underlined portion a potential PD, incorrectly predicting VP-ellipsis to be an option in (9).[7]

[7]We follow Hartman here in representing the base position of the subject as Spec,VP here and in other schematics above as well. Things would be complicated further if we were to follow much recent work in representing the base position of the subject as the specifier of a separate VP-shell projection like *v*P, as this would also lead us to a situation where the lower VP-shell would not contain a rebound variable, again incorrectly predicting the availability of VP-ellipsis (at least if it turned out that VP-ellipsis were not to be reanalysed as *v*P-ellipsis).

(39) $[_{CP}$ when λx $[_{C'}$ will λy $[_{TP}$ x $[_{TP}$ you λz $[_{T'}$ y $[_{VOICEP}$ z $\lambda z'$ $[_{VP}$ z′ pay me back]]]]]]]

Thus it is not possible to provide both sets of facts with a unified analysis on this account.

As we see it, the crux of the matter with (36) is the proposal that A-traces count for the calculation of MaxElide, as this leads us to expect a far greater number of rebinding configurations. If we remove this component of the analysis, the data in (36) no longer present an immediate problem, although then we are left without an account of the matrix *wh*-adverbial cases in (9). However we already noted above that the analysis of (9) needs to be rethought anyway, since it seems to involve a systematic violation of the parallelism condition, and the fact that VP-ellipsis is in fact salvageable in related examples like (14) tells us that this is a problematic analysis; that is, a structure that violates parallelism ought not to be salvageable, since parallelism is an inviolable constraint. If A-traces did not count for the calculation of parallelism, then this problem would go away, because then the projections below TP would form a PD and thus the lack of parallelism in the CP-domain would be irrelevant in cases like (14), where what is elided is just the VP, as shown by the revised schematic in (40). That is, the elided VP would be parallel to its antecedent, and the fact that the two structures differ in the higher domain would not matter.

(40) $[_{CP}$ when λx $[_{C'}$ WILL λy $[_{TP}$ x $[_{TP}$ you $[_{T'}$ y $[_{VP}$ pay me back]]]]]]

While this might be a step in the right direction, it would still leave mysterious the difference between matrix and embedded *wh*-adverbial questions. With those cases, it would seem to be T-to-C movement which is crucial in conditioning whether or not sluicing blocks VP-ellipsis, but this does not follow from the MaxElide account once A-movement is taken out of the picture. The question, then, is how to account for this effect of T-to-C movement on ellipsis options without A-traces.

To summarize, we have seen that the MaxElide account has a number of technical and empirical problems. Two important empirical results have emerged from this critical discussion. The first result is that non-parallel extraction from VP-ellipsis is highly restricted, with very many cases of VP-ellipsis not being salvaged when intervening focus rules out the sluicing option, contrary to the predictions of MaxElide. This indicates that a number of the effects normally attributed to MaxElide may in fact be better analysed as involving some other hard constraint on ellipsis. The fact that these extractions from VP-ellipsis *are* possible with overt *parallel* extraction indicates that the relevant hard constraint may be parallelism. The second result is that allowing A-traces to count for the calculation of MaxElide gets us into trouble with simple cases of VP-ellipsis; removing A-movement from the picture removes this problem, and while it undermines aspects of the analysis of matrix *wh*-adverbials, these were ripe for an alternative analysis anyway, since we had seen that the previous account had its own problems.

4 An alternative account of MaxElide effects

As the previous sections established, there are many problems with MaxElide. It was shown that in many cases extraction from VP-ellipsis was restricted much more than MaxElide would predict. In this section, we put forth a new analysis that can handle all of the problematic data, according to which the observed restrictions on extraction from VP-ellipsis have different sources.

Our analysis divides the data into two groups: the salvageable cases, where focussing the subject or an auxiliary saves the VP-ellipsis options, and the unsalvageable cases, where focus has no salvaging effect (while overt parallel extraction is still possible). For the unsalvageable cases, we provide an analysis in terms of *parallelism*, which we take to be a hard constraint on ellipsis. An account of this kind was already offered in the previous section, where we pointed out that the ungrammaticality of matrix *wh*-object questions like those in (15) follows from the fact that AC and EC mismatch with respect to the position of binders of head movement. We show that mismatches with respect to movement paths created by A′-movement also lead to parallelism failures. Once this large class of cases is taken out of the picture, only a small set of salvageable extractions from VP-ellipsis remains: local object extractions like in (6) and matrix

wh-adverbial questions like (14). We analyse these in terms of generalized derivational economy. We argue that ellipsis bleeds certain movement steps, with the result that these derivations are shorter than competing derivations; this results in a preference for some elliptical derivations over others, and this preference interacts with hard constraints to derive the core MaxElide effects. The picture this leaves us with is one without the ellipsis specific constraint of MaxElide. We will see that it also leaves us with a picture where only A′-movement and head movement count for the calculation of parallelism.

4.1 Parallelism and extraction from VP

In the previous section, we proposed a solution to Hartman's puzzle concerning matrix *wh*-object extraction out of VP-ellipsis. Hartman admits the fact that this class of extraction is unsalvageable under his MaxElide analysis is unexpected, however, we showed that extraction *is* possible when there is overt parallel *wh*-movement and head movement in the AC. This led us to conjecture that what is blocking VP-ellipsis with matrix object extraction is the inviolable parallelism constraint on ellipsis. In this section we present a similar analysis for other cases of unsalvageable VP-ellipsis extractions, but before doing so, we must address the notion of parallelism that we adopt for this analysis to go through. As an anonymous reviewer points out, the notion of parallelism needed must take into account the LF structures of the AC and EC, not just their denotations. Parallelism as defined in section 2, does not do this. There are, however, a number definitions of parallelism that do take into account structure (see e.g., Fiengo and May 1994, Fox 1999b, Fox 2000, Fox and Lasnik 2003, Thoms to appear c). For concreteness, let's adopt the definition of parallelism from Griffiths and Lipták (2014) stated below.[8]

(41) *Scopal Parallelism in Ellipsis*
Variables in the antecedent and elided clause must be bound from parallel positions.

With this definition of parallelism, we can now account for all of the unsalvageable extractions from VP-ellipsis.

Recall from the last section that extraction from VP-ellipsis is highly restricted. Lasnik and Park (2013) noted that long distance object extraction from a finite clause is impossible even with intervening focus contra the prediction of MaxElide. The relevant example is repeated below.

(42) *Abby said they heard about a Balkan language, but I don't know what kind of language BEN did.

We noted previously that the same restriction does not hold for control clauses. Long distance extraction out of VP-ellipsis with a control complement is possible, as the below example shows.

(43) ?John WILL try to kiss MARY, but I don't know who he WON'T.

This asymmetry between finite and control clauses does not follow from the application of the MaxElide constraint as things stand: VP is a possible target for deletion in both cases, and so MaxElide would predict both (42) and (43) to be grammatical. However we can show that these restrictions follow from the parallelism requirement on ellipsis, independent of the application of MaxElide or any other such constraint. Our proposal is that in most cases, the crucial factor is the form and position of the XP corresponding to the extracted *wh*-phrase, which we call the *correlate*. This analysis is broadly in line with the proposal in Thoms

[8]This definition is a simplification of what is needed, as there are well known problems with such strict views of parallelism, such as the data in (i). In (ia) and (ib) a sloppy reading is available despite the fact that the DPs in the antecedent and elided clauses occupy different positions in (ib).

(i) a. John's boss fired him and Bill's boss did too.
b. The guy John works for fired him and Bill's boss did too.

Rooth (1992) uses this data to argue against a structural view of parallelism; however, we believe that such examples could be made to fit within a structural view of parallelism if we allow for *accommodation* of an new antecedent as in Fox (1999b) and Thoms (to appear a).

(to appear c), where similar logic is applied to an analysis of very similar restrictions on pseudogapping (see also Griffiths and Lipták 2014).

Recall from the previous sections that parallelism requires that the binding relations found at LF in the EC must match those found in the AC. Assuming successive-cyclic *wh*-movement through *v*P and CP, the binding relations for the EC in (43) would be those in (44).

(44) $[_{CP}$ who λx $[_{TP}$ he $[_{T'}$ won't $[_{vP}$ x λx′ $[_{v'}$ try $[_{TP}$ PRO $[_{T'}$ to $[_{vP}$ x′ λx″$[_{v'}$ kiss x″]]]]]]]]]

On the assumption mentioned above that focused DPs undergo quantifier raising (QR) (Chomsky 1976, Krifka 2006), parallelism is satisfied in this case by QR of the focused DP *Mary* in AC. The LF for the AC is given below.

(45) $[_{CP}$ MARY λx $[_{TP}$ John $[_{T'}$ will $[_{vP}$ x λx′ $[_{v'}$ try $[_{TP}$ PRO $[_{T'}$ to $[_{vP}$ x′ λx″$[_{v'}$ kiss x″]]]]]]]]]

Turning to (42), again assuming successive cyclic movement through VP and CP, the binding relations for the EC of (42) are shown in (46).

(46) $[_{CP}$ what kind of language λx $[_{TP}$ BEN $[_{T'}$ did $[_{vP}$ x λx′ $[_{v'}$ say $[_{CP}$ x′ λx″ $[_{TP}$ they $[_{T'}$ T $[_{vP}$ x″ λx‴$[_{v'}$ heard about x‴]]]]]]]]]

In order for parallelism to be satisfied, matching binding relations must be created at LF for the AC. However these binding relations cannot be created via QR of the correlate DP in the AC, because finite clauses are typically barriers to QR (May 1985, Johnson 2000, Fox 2000).[9] The inability of QR to cross a finite clause boundary is demonstrated by the fact that inverse scope cannot be obtained between the matrix subject and the embedded object in (47a). Compare this to an example with a control infinitive in (47b) which allows for the inverse scope reading, indicating that QR is possible out of the infinitive.

(47) a. A (#different) student said that Sue read every book. *$\forall \gg \exists$
b. A (different) student tried to read every book. $\forall \gg \exists$

Since the correlate in (42) cannot escape the finite clause by QR, it cannot create binding relations which are parallel to those in the EC, and so parallelism is violated when EC is elided and hence (42) is ungrammatical. The same explanation applies to the cases of *wh*-adverbial extraction in (29) as well, where the only difference is that the correlate is adjunct like *at noon* rather than an object. In all cases, we correctly predict that intervening focus has no effect on the ungrammaticality of VP-ellipsis, since they all involve violations of parallelism and this is an inviolable constraint.[10]

This proposal raises an immediate concern regarding the status of cases of sluicing with focussed correlates, such as (48), which is broadly similar to the VP-ellipsis case in (42).

(48) Abby said they heard about a Balkan language, but I don't know what KIND of Balkan language.

The problem is this: if the correlate must undergo QR out of the finite clause to create parallel binding relations with extraction from VP-ellipsis, then applying the same logic to sluicing ought to rule out (48) as a parallelism violation as well, since the correlate will be just as clausebound in the antecedent to sluicing

[9] Principled exceptions to this generalization are noted in Farkas and Giannakidou (1996), Kennedy (1997) and Kayne (1998), but they do not undermine the data discussed here, which control for these exceptions. We consider some of these exceptions in what follows.

[10] As a reviewer notes, long-distance object extractions are *not* unsalvageable when VP-ellipsis targets the *lower* VP.

(i) I know you said John spoke to someone, but I don't know who you said MARY did.

Our analysis does in fact predict that VP-ellipsis will be salvageable in the *embedded* clause, since with successive-cyclic movement the embedded clause of a long-distance extraction will look broadly similar to regular local object extraction: in the ellipsis clause there will be a lambda binder in the Spec,CP local to the VP-ellipsis site, and in the antecedent clause the indefinite correlate will QR to the embedded Spec,CP, thus satisfying parallelism.

as it is in the antecedent to VP-ellipsis. Indeed, as noted by Fox and Lasnik (2003) the parallelism problem goes beyond sluicing with contrastively focussed correlates: even in simple cases of sluicing like (49) where the correlate is a wide-scoping indefinite, the punctuated path of movement created by *wh*-movement will create binding relations in EC which will not be identical to those created by the indefinite, which Fox and Lasnik (2003) and others assume takes wide scope by virtue of an in-situ scoping mechanism like choice functions (Reinhart 1997). This is shown in the diagrams in (50) (Fox and Lasnik 2003, 149-150), where the lack of parallelism is laid bare.

(49) Fred said that I talked to a certain girl, but I don't know which girl.

(50) AC: $\exists\ \lambda f'$ [Fred [said [that I [talked to f'(girl)]]]]
EC: which g ~~girl~~ $\lambda g'$ [Fred [g' λ g'' said [g'' λ g''' that I [g''' λ g'''' talked to g''''(girl)]]]]

Fox and Lasnik (2003) propose that this problem disappears if we assume that the *wh*-movement in sluicing can proceed in one fell swoop from its base position to the landing site in Spec,CP, with no stop-offs at intermediate landing sites like Spec,*v*P or the embedded Spec,CP. On this analysis, the in-situ scoping correlate and the *wh*-operator create binding configurations which do indeed respect parallelism, as (51) shows.

(51) AC: $\exists\ \lambda f'$ [Fred [said [that I [talked to f'(girl)]]]]
EC: which g ~~girl~~ $\lambda g'$ [Fred [said [that I [talked to g'(girl)]]]]

Fox and Lasnik propose that the claim that *wh*-movement in sluicing can proceed in one fell swoop is justified by the well-known difference between sluicing and VP-ellipsis with respect to island amelioration (Ross 1969, Merchant 2008),[11] and Fox and Pesetsky (2005) make the same claim in the context of a discussion of the interaction of cyclic spellout and linearization. We therefore follow Fox and Lasnik (2003) and Fox and Pesetsky (2005) in assuming that this is possible here. We will see in section 4.2 that this plays an important role in our account of the remaining MaxElide effects, so insofar as the analysis holds together, it provides further support for this claim about *wh*-movement in sluicing.

Along with the assumption that contrastive foci are like indefinites and *wh*-in-situ in being able to take scope by in-situ mechanisms as well as by QR (Wold 1996, Reich 2004, Krifka 2006), this allows us to account for the difference between VP-ellipsis and sluicing with respect to extraction from finite clauses. With non-parallel extraction from VP-ellipsis, the path of *wh*-movement from EC is punctuated (just as it is without ellipsis), stopping off at intermediate adjunction positions like Spec,*v*P. In order for parallelism to be satisfied, then, there must be parallel binding relations in AC, and the only way to do this is for the correlate to undergo QR; it is not sufficient for the correlate to take scope by in-situ mechanisms in this case, because if it does there will be a mismatch of the kind seen in (50). As a consequence, non-parallel extraction from VP-ellipsis is tied to QR: extraction from VP-ellipsis will only be possible in those situations where the correlate can be extracted from the antecedent VP by QR. Note, however, that no such restrictions are expected to hold of overt *parallel* extraction from VP-ellipsis – that is, those cases where there was overt parallel *wh*-extraction from the antecedent as well – since the parallel extraction would of course ensure that parallelism was satisfied: the *wh*-phrases in AC and EC will stop off at the same landing sites and will therefore create fully parallel movement paths. And of course no such restriction holds of sluicing either, since the *wh*-movement in sluicing may proceed in one fell swoop to the final Spec,CP, and so the correlate can scope wide by in-situ mechanisms without violating parallelism.

The claim that non-parallel extraction from VP-ellipsis is tied to QR makes a number of predictions, to which we turn now. One prediction concerns finite clausal complements: if it were possible for an argument to scope out of the embedded finite clause, then long extraction would also be available for those arguments, since matching binding relations could be established in the AC by QR of the correlate. Kayne (1998) observes that when the matrix subject binds the embedded subject of an embedded finite clause, the

[11] Whether the facts from island (in)sensitivity support this analysis is debatable: see Abels (2011) and Barros et al. (2014) for critical discussion of the notion of island repair.

embedded object can scope out of the finite clause sas shown in (52).[12]

(52) [At least one of these men]$_i$ thinks he$_i$ is in love with each one of these women. $\forall \gg \exists$

The prediction of our account is that non-parallel extraction from VP-ellipsis should be possible in these configurations. (53) shows that this prediction is borne out: (53) is grammatical with extraction from VP-ellipsis, in striking contrast to the ungrammatical long extraction in (42) above.

(53) John$_i$ said he$_i$ kissed MARY but I don't know who BILL$_k$ did ~~say he$_k$ kissed t~~.

Turning to control complements, recall from (47b) that QR is normally possible out of control infinitives, however it has been noted to us by Susi Wurmbrand (p.c., attributing the observation to Benjamin Bruening) that QR appears to be blocked when the infinitive is extraposed, as demonstrated by the lack of inverse scope in (54) (cf. (47b) where inverse scope is available).

(54) Some European country tried, in the 20th century, to invade every African nation. *$\forall \gg \exists$

Since QR out of the complement is impossible in such configurations, we predict that non-parallel extraction will also be impossible. This prediction is again borne out. It is precisely in these cases where extraction from control complements of VP-ellipsis is substantially degraded, as shown in (55).

(55) *Mary tried, over the summer, to read MOBY DICK, but I don't know what BILL did ~~try, over the summer, to read t~~.

The above data suggest a strong correlation between the correlate's ability to QR and the ability to extract from VP-ellipsis, one that follows straightforwardly from parallelism with no appeal to MaxElide.

The parallelism-based analysis extends to the other cases in the previous section where asymmetric extraction from VP-ellipsis was impossible irrespective of intervening focus. Recall that extraction out of a DP complement in VP-ellipsis is ungrammatical. The examples are repeated below.

(56) a. *ABBY heard a lecture about a Balkan language, but I don't know what kind of language BEN did. Lasnik and Park (2013, 240)

b. *ABBY should give a lecture on ART, but I don't know what BEN should.

Once more the answer here lies in considering the scopal properties of the correlate: in order to create parallel binding relations in the AC, it must QR out of the DP, but this is not possible because DP is a scope island (May 1985, Larson 1985, Charlow 2010; cf. Sauerland 2005) just like finite CPs are. Since the correlate cannot QR out of the DP, it is impossible to create the necessary binding relations in the AC, and so parallelism fails. A slightly different analysis holds for the case where the *wh*-phrase is a degree phrase like *how upset*. Recall that these could not be extracted from VP-ellipsis even if the extraction was local:

(57) *John became very upset, but I don't know how upset BILL did.

[12]The correlation between the exceptional wide scope of universal QPs in embedded contexs and the availability of long-distance extraction from VP-ellipsis is not perfect. For instance, Farkas and Giannakidou (1996) note that universal QP subjects can take "extrawide scope" out of finite complements of predicates like *make sure*, yet the same complements do not allow for long-distance extraction from VP-ellipsis.

(i) A student made sure that every invited speaker had a ride. $\forall \gg \exists$

(ii) *Sally made sure JOHN got a job, but I don't know who SARAH did ~~make sure t got a job~~.

It is perhaps relevant that Farkas and Giannakidou's account of (i) is not stated in terms of covert movement but rather in terms of how the lexical semantics of the embedding predicate ensures that the embedded subject and matrix subject behave as if they were co-arguments. Thus, it could be the case that these finite clauses are barriers for QR, allowing us to retain the parallelism-based account of (ii), while some other mechanism ensures inverse scope of the two quasi-co-arguments. Clearly more work is needed to make this argument work though.

We propose that parallelism fails here because the correlate here is a predicate, and predicates are non-quantificational and hence unable to undergo QR. As before, this means that the extraction in EC is not matched by covert A′-extraction in AC, and so parallelism is violated. No amount of intervening focus can salvage the VP-ellipsis option here, although it is correctly predicted that it will be grammatical if there is parallel extraction in AC (as in (35)).[13]

Before moving on, there is one more case of unsalvageable extraction from VP-ellipsis which remains to be accounted for: matrix *wh*-object extractions like Hartman's (15a), repeated here.

(15a) Mary will kiss Bill. Who will JOHN *(kiss)?

Going by the discussion at the beginning of section 3, we may assume that (15) follows straightforwardly from parallelism, since there is no parallel head movement in AC and EC. This would indeed follow if we took the correct LF for EC in (15a) to be as in (58): extraction of the object makes the VP a rebinding configuration, and since the binder of the variable in the object position is in Spec,CP then it is this whole domain which needs to be taken into account in calculating parallelism. And since there is no parallel head movement in AC, parallelism will not be satisfied, even if the object QRs in parallel to Spec,CP. Note that traces of A-movement need not be taken into account here.

(58) $[_{CP}$ who $\underline{\lambda x\ [_{C'}\ \text{will}\ \lambda y\ [_{TP}\ \text{John}\ [_{T'}\ y\ [_{VP}\ \text{kiss}\ x\ ]]]]]}$

However this analysis begins to come apart at the seams once we assume, as we have done above, that *wh*-movement is successive-cyclic, stopping off at intermediate *v*P and CP projections on the way to the final scope position. If we add this to our schematic, and assume that each intermediate step of movement creates a separate variable binding configuration (as does Hartman 2011), then the projection "closed off" by the binder left in the intermediate landing site, $\lambda x'$ in (59), ought to create a PD in which ellipsis would apply to derive VP-ellipsis. This is not what we want, since VP-ellipsis is never possible in these configurations.

(59) $[_{CP}$ who $\lambda x\ [_{C'}\ \text{will}\ \lambda y\ [_{TP}\ \text{John}\ [_{T'}\ y\ [_{vP}\ x\ \underline{\lambda x'\ [_{VP}\ \text{kiss}\ x'\ ]]]]]]$

Note that reintroducing A-traces into the LFs is not the way to go: while this would give us an account of matrix *wh*-objects, recall that it would lead us to expect the same behaviour from matrix *wh*-adverbials, in that the latter would also be incorrectly predicted to be unsalvageable. Finally, we may recall that it does indeed seem to be T-to-C movement that is implicated in making (15a) unsalvageable, since we saw that VP-ellipsis can be salvaged with embedded *wh*-objects, and indeed with parallel T-to-C extraction alongside *wh*-extraction (cf. (17)). Thus, what we need is an analysis in which any extraction from VP "catches" the binding path left by T-to-C movement, resulting in an unsalvageable parallelism violation with non-parallel extractions, while extraction of TP-adjuncts in matrix *wh*-adverbial questions does not.

We propose that the relevant configuration arises if we assume that all auxiliary verbs that occur in T move there from a lower *v*P projection. For (15a), this would give us the LF in (60).

(60) $[_{CP}$ who $\underline{\lambda x\ [_{C'}\ \text{will}\ \lambda y\ [_{TP}\ \text{John}\ [_{T'}\ y\ \lambda y'\ [_{vP}\ x\ \lambda x'\ [_{v'}\ y'\ [_{VP}\ \text{kiss}\ x'\ ]]]]]]]$

Here we see that the trace of *v*-to-T movement ensures that the *v*P is no longer a PD, since it contains a trace which is rebound from T, and the result is that the smallest possible PD is the one created by the second step of successive-cyclic *wh*-movement. But since this PD contains within it the path of T-to-C movement, the result is that the AC must also contain T-to-C movement in order for parallelism to be satisfied. Thus, we correctly predict that matrix extraction of a *wh*-object from VP-ellipsis where the antecedent does not

[13] Supporting evidence comes from the fact that DP-predicates are also degraded as remnants of extraction from VP-ellipsis:

(i) *I'm sure *John* will become A FOOTBALLER, but I don't know what HIS BROTHER will.

This indicates that it is not necessarily the categorial status of AP-remnants which rules them out as remnants, but rather restrictions on what can QR.

also involve inversion, like in (15), will always involve a parallelism violation, since the binders created by the moved object and the moved auxiliary will always overlap and hence "extend" the smallest putative PD up to CP, where non-parallelism with respect to T-to-C movement is found. Crucially, this does not upset our analysis of matrix *wh*-adverbials or *wh*-subjects, since the paths of *wh*-movement and *v*-to-T movement will not intersect to extend the smallest possible PD all the way to the left periphery in the same way. We can see this by looking at the revised LF for the matrix *wh*-adverbial question in (14b), repeated here: the domain formed by *v*-to-T is a PD containing no rebound variables, and since this excludes the trace left by T-to-C we do not expect non-parallelism in this domain to lead to an unsalvageable parallelism violation.

(14b) If Anna isn't going to resign today, then when WILL she?

(61) $[_{CP}$ when λx $[_{C'}$ will λy $[_{TP}$ x $[_{TP}$ she $[_{T'}$ y $\underline{\lambda y' [_{vP} [_{v'} y' [_{VP} \text{ resign }]]]]]]]}$

Thus we neatly capture the distinction between VP-extractions on the one hand and IP-domain extractions on the other. The only issue is that this comes at the expense of assuming that all auxiliaries, including the modals and *do* as well as *be/have*, are heads of *v*P-projections and that the uniformly move to T rather than being base-generated there. However there are proposals in the literature which support this view: Embick and Noyer (2001), Bjorkman (2011) and Thoms (to appear b) propose analyses where *do* is a spellout of *v* when it has moved to T, while Iatridou and Zeijlstra (2013) propose that the scopal interactions of modals and negation indicate that these, too, must be base-generated in a lower position and then moved to T. We therefore take this assumption to be well-supported and submit that this analysis, insofar as it is successful, can be taken to support this view of auxiliaries in English.

4.2 Ellipsis and derivational economy

To account for the remaining cases, we follow the core intuition of Merchant's (2008) original proposal by proposing an account in terms of derivational economy. However we depart from Merchant and others since him in rejecting the ellipsis-specific economy constraint MaxElide, which states a preference for ellipsis of larger constituents over smaller ones, as we have seen already that it has a number of empirical problems which make stating its domain of application in a way that doesn't enforce larger ellipsis domains at all times very difficult. Rather, we propose that the interaction of ellipsis processes and movement is such that derivations involving certain ellipsis processes can require fewer movement steps, with the result that the shorter derivation is preferred to the longer one, making the latter degraded (Chomsky 1991, 1993, Kitahara 1997, Epstein 1992).

In the context of the present analysis, this is most straightforward in the case of local object extractions like (1), repeated below.

(1) Mary was kissing someone, but I don't know who (*she was).

Recall from section 4.1 that we assumed, following Fox and Lasnik (2003), that *wh*-movement can move in one fell swoop just in case sluicing applies; or, to put it another way, sluicing bleeds successive-cyclic movement, a fact that would follow from an approach where the need of moved phrases to pass through certain projections is phonological in nature (as in Fox and Pesetsky 2005; see also Bošković 2007). In the case of object extraction, this means that a derivation which involves sluicing will be more economical than one which involves VP-ellipsis, since the latter will require two steps of A′-movement but the former will only require one, as the schematics in (62) show.[14]

[14] A reviewer notes that this only goes through if we adopt the view of derivational economy according to which the metric employed for effort is a count of a the number of derivational steps (e.g. Chomsky 1995). As they point out, this may not be correct, as it is plausible that the correct notion of economy is one where what matters is the length of dependencies as counted in terms of nodes crossed, with direct consequences for our account, and they note furthermore that the general preference for shorter dependencies over longer ones evidenced by psycholinguistic studies (e.g. Crain and Fodor 1985, Gibson 2000, Phillips et al. 2005) could be a factor that affects offline judgments of the kind discussed above. Although we agree that this is an issue that one should consider when it comes to economy-based arguments, it is not clear to us whether the preference for shorter

(62) a. ... [$_{CP}$ who$_i$ ~~[$_{TP}$ she was [$_{VP}$ [$_{VP}$ kissing t$_i$]]]]~~ sluicing
b. ... [$_{CP}$ who$_i$ [$_{TP}$ she was [$_{vP}$ t$_i$ ~~[$_{VP}$ kissing t$'_i$]~~]]] VPE

We assume that VP-ellipsis and sluicing are competing derivations, with both having the same information-structural properties and the only difference being the size of the ellipsis site. Competition only applies when the two ellipsis options are both contained within the same PD, that is, if the target for VP-ellipsis is contained within the target for sluicing and it is fully parallel to the AC. If the two compete, it follows that the sluicing derivation will block the VP-ellipsis one here because it is more economical, requiring one less step of A′-movement. Crucially, if sluicing is ruled out by placing focus in the IP-domain, hence making the information structure properties of sluicing and VP-ellipsis distinct, then it will not compete and VP-ellipsis will be possible, so we correctly predict that intervening focus will save the VP-ellipsis option with embedded *wh*-objects and related extractions from VP.[15]

An important property of this account is that this only predicts sluicing to be more economical when the remnant is extracted from within the VP by successive-cyclic movement. If the remnant is extracted from the IP-domain, as with subject questions or *wh*-adverbial questions where the adverbial modifies the IP-domain, then the *wh*-phrase will not need to make any successive-cyclic stop-offs in either the sluicing or VP-ellipsis derivation, and so the number of steps of A′-movement with a sluicing derivation and a VP-ellipsis derivation will be the same. This predicts there to be no competition between sluicing and VP-ellipsis when the subject is extracted locally from the IP-domain, as with subjects and IP-level *wh*-adverbial modifiers. As Schuyler (2001), Merchant (2008) and Hartman (2011) note, this is correct for cases like thes:

(63) Someone left, but I don't know who (did).

(64) You say you'll pay me back but you haven't said when (you will).

Our account also improves on the MaxElide account by not predicting competition between the different VP-ellipsis in examples like (36), repeated here, since there is no A′-extraction and there seems to be no good reason to believe that one ellipsis derivation would be more economical than the other.

(36) a. John has been singing, and Mary has (been), too.
b. John shouldn't be drinking, and Mary shouldn't (be), either.

Indeed our account only predicts competition between ellipsis options when one of the ellipsis options is sluicing; that is, it is not the presence of A′-movement itself which brings about competition (as in Merchant 2008), but the one-fell-swoop derivation for sluicing. That this is the case is attested by the fact that there is no competition between big and small ellipsis options in configurations where there is A′-movement from the ellipsis targets but there is no option to apply sluicing; this is demonstrated by the fact that ellipsis of the

dependencies would necessarily lead to a preference for dependency formation derivations involving a sequence of shorter steps ahead of ones with longer steps where the two ultimately involved creating a global dependency of the same length, as evidence for the short-step preference comes primarily from experiments which show preferences for creating shorter global dependencies than the ones that are required (i.e. preferences for subject relativization over object relativization).

[15] As an anonymous reviewer notes, we must find a way for sluicing to block VP-ellipsis, but not block a sentence without any ellipsis. As (i) shows, sluicing in (ia) does not block (ib), though presumably (ib) contains more steps of movement than (ia).

(i) a. John was kissing someone but I don't know who.
b. John was kissing someone but I don't know who he was kissing.

Intuitively, we only want derivations that include the process of ellipsis to be competitors. But how do we formalize this? Here is one way. Let's assume that only derivations that have the same numerations compete. Let's also assume, following Merchant (2001), that ellipsis is licensed by an E(llipsis)-feature. Departing from Merchant slightly, let's assume that the E-feature is a morpheme that merges on to certain functional heads and triggers non-pronunciation of that head's complement. For both the sluicing and VP-ellipsis derivations, the numerations would contain the ellipsis licensing E-feature. The difference between the two being which functional head the feature merges with (C for sluicing and T for VP-ellipsis). A sentence without any ellipsis would not have the E-feature present in its numeration and thus would not be a competitor against the ellipsis derivations.

VP containing the control complement in (65) (a variant of (31) above) does not block the option to elide just the infinitival complement.[16]

(65) JOHN wants to kiss MARY, but I don't know ...
  a. ... who BILL does.
  b. ... who BILL wants to.

MaxElide would predict competition between (65a) and (65b), leading to the prediction that (65b) would be blocked by (65a). No such prediction is made by our account, since we only predict competition between the economical option of one-fell-swoop sluicing and other elliptical derivations.

Now let us turn to the case of matrix *wh*-adverbial questions. Recall that these differ from their embedded counterparts like (64) in requiring intervening focus to ensure that VP-ellipsis is possible; that is, sluicing seems to outcompete with VP-ellipsis here in the matrix examples, but not in the embedded ones.

(9) We know Anna is going to resign. The only question is: when (*is she)?

(14b) If Anna isn't going to resign today, then when WILL she?

As noted above, Hartman provides data from Indian and Irish dialects of English which indicates that the crucial difference between (9) and (14b) is T-to-C movement: VP-ellipsis is blocked when there is T-to-C movement, at least when both sluicing and VP-ellipsis are possible. Our proposal is that sluicing blocks VP-ellipsis in (9) because *sluicing bleeds T-to-C movement*, resulting in a more economical derivation. Specifically, we argue that the landing site for T-to-C movement is within the sluicing site, and deletion of the landing site bleeds T-to-C as it is driven by a PF-condition which dictates that the null C_{+wh} is affixal and must be supported by an overt head like T. This makes the sluicing derivation shorter than the VP-ellipsis, since the VP-ellipsis derivation involves an extra movement step, and so economy prefers the sluicing option all other things being equal.

(66) $[_{CP}$ when$_i$ $[_{C'}$ ~~C $[_{TP}$ t_i $[_{TP}$ subj T VP ~~$]]]]$ sluicing
$[_{CP}$ when$_i$ $[_{C'}$ C+T$_j$ $[_{TP}$ t_i $[_{TP}$ subj t_j ~~VP~~ $]]]]$ VP-ellipsis

This predicts intervening focus to have ameliorating effect with the VP-ellipsis option, and of course it also predicts that this effect will only be seen when T-to-C is involved.

To make this analysis convincing, we need to provide some support for the component parts of the analysis, namely (i) that the sluicing site contains the landing site for T-to-C movement, and (ii) that T-to-C movement is motivated by PF-conditions and so can be bled by ellipsis. The first component is motivated by Merchant's (2001) "sluicing-comp generalization," which states that no non-operator material may survive sluicing. This is motivated by the fact that overt complementizers never occur to the left of the sluicing remnant, even when they can co-occur in the non-elliptical structures in the language in question (with evidence from Slavic, Germanic and Celtic languages), and this strongly suggests that the constituent deleted in sluicing is large enough to contain the complementizers in the CP-domain. We do not dwell on the matter of how to explain the sluicing-comp generalization,[17] but we take it that any account of this restriction would

[16]Takahashi and Fox (2005) argue that the opposite effect holds in examples like (i), with parallel extraction in AC and EC:

(i) a. *I don't know which puppy you should agree to adopt, but I know which one you should NOT agree to.
  b. I don't know which puppy you should agree to adopt, but I know which one you should NOT.

We are unsure why there is variation in these effects. We note however that the contrast in (i) would not follow from any of the other accounts if they were to adopt the widely-held assumption that *wh*-movement is cyclic, since on this account the lambda operators left by cyclic movement through the embedded VP or CP would demarcate PDs in which the application of MaxElide would derive (ia) as grammatical.

[17]There are different ways to capture this. One is to follow Rizzi (1997) in assuming that the CP-domain is split into a number of different projections, and to further assume that the landing site for *wh*-movement is in a higher CP-projection than the one that is targeted by T-to-C, with a head of the higher projections licensing sluicing. Such an approach is explored in Craenenbroeck

generalize to account for missing complementizers in sluices would also account for the absence of T-to-C movement, as suggested by Merchant (2001), and we posit that the most plausible analysis is one where the complementizers are contained in the ellipsis site. The second component of our analysis, the claim that T-to-C movement is driven by PF-conditions and can be bled by ellipsis, is not new, having been advanced for English by Lasnik (1999, 2001) and van Craenenbroeck and Lipták (2008) for Hungarian. Lasnik's argument comes from the fact that matrix sluices do not retain the auxiliary, as in (67), but this is undermined by the sluicing-comp generalization, which subsumes this effect to the deletion of complementizers.

(67) A: John kissed someone.
B: *[$_{CP}$ Who [$_{C'}$ did ~~[$_{TP}$ he kiss]~~]]

Nevertheless the claim that null complementizers in English are affixes which need support from some other head in the structure has been made in different forms in Pesetsky (1991), Bošković and Lasnik (2003) and Kim (2008) (see also Bruening to appear for an alternative PF-based analysis of inversion), so we take it that the proposal that our assumptions about C in English are reasonably well-founded. As for the claim that verb movement to such a target may be bled by ellipsis, van Craenenbroeck and Lipták (2008) analyse an interesting set of facts in Hungarian that seem to provide compelling evidence for this. The evidence comes from so-called *focus sluices* like (68), where the remnant of sluicing is a focussed non-*wh* XP in a yes-no question. As they note, the head which realizes C in yes-no questions, which normally surfaces adjoined to the verb in C, is found attached to the ellipsis remnant in the focus sluices. They interpret this as indicating that the verb has failed to undergo head movement to that C position because it has been bled by ellipsis.

(68) János meghívott egy lányt, de nem tudom hogy ANNÁT*(-e).
John invited a girl but not I.know COMP Anna-Q
'John invited a girl, but I don't know if it was Anna.'

Thus it seems our basic assumptions about T-to-C movement are reasonably well-supported, although clearly more work needs to be done to unearth further evidence for this effect.

So far we have argued that ellipsis may bleed verb movement and successive-cyclic movement in certain contexts, and that this has consequences for the economy of derivation which are reflected in the data normally attributed to MaxElide. At this point, one may wonder whether there is a principled way of predicting which types of movement could be bled by ellipsis, as it would complicate our analyses substantially if it turned out that a larger class of movements were bled by ellipsis. The null hypothesis is that only movement that is motivated by PF-constraints can be bled, and we have indicated that this may hold for the movements which were bled by ellipsis in our analyses, namely intermediate steps in successive-cyclic A′-movement and T-to-C movement. What other movement rules can be analysed this way? One candidate is A-movement to Spec,TP, as there are proposals in the literature for PF-based accounts (Sauerland and Elbourne 2002, Landau 2007) and indeed it has been proposed by Merchant (2001) and van Craenenbroeck and den Dikken (2006) that A-movement is indeed bled by sluicing. This would have implications for our analysis of subject *wh*-questions like (63): if ellipsis bled A-movement, then the derivation for a subject sluice might be able to omit the step of A-movement prior to *wh*-movement, making the sluice more eco-

(2010) and Baltin (2010). This has a few problems though, such as the fact that it divorces T-to-C movement from movement to the spec of the same projection, losing the core insight of criterial accounts that hold that one movement causes the other. It also leads us to expect that in languages with highly "isolating" CP-fields with overt realizations for all the heads of CP-projections, the complementizer which heads the higher CP-projections hosting the *wh*-phrase and licensing ellipsis will be realized in sluices. This does not seem to be correct though, as Welsh obeys the sluicing-comp generalization, yet it seems to be a good candidate for a language with an isolating CP-field, as it may realize up to three distinct C-heads simultaneously (Hendrick 2000). An alternative analysis of the sluicing-comp generalization is provided by Thoms (2010). Thoms rejects the idea that ellipsis is licensed by a specific set of lexically specified heads, like T in the case of VP-ellipsis and C in the case of sluicing, and instead he proposes that ellipsis is generally licensed by overt movement, with ellipsis effectively being another way of doing copy deletion. According to this theory, the licensor of ellipsis in sluicing is the moved *wh*-phrase itself, which licenses deletion of its structural complement; this includes the complementizer, thus deriving the sluicing-comp generalization.

nomical and thus leading to the incorrect prediction that sluicing would outcompete VP-ellipsis. We do not take this particular case to be a problem here, though, since the PF-based theories of A-movement have a number of problems (see e.g. Lasnik and Park 2003 and Barros et al. 2014 on the claim that A-movement is bled by sluicing), but it is illustrative of the wider issue for our account. We must leave this as a topic for future research, although we note optimistically that it may be possible to turn things around and use "MaxElide" effects as a probe for identifying movement rules which are driven by PF-conditions.

5 Conclusion

In this reply we have argued against accounts of interactions between *wh*-movement and ellipsis in terms of MaxElide, which enforces competition between sluicing and VP-ellipsis in narrowly defined domains. We showed that extraction from VP-ellipsis is more restricted than we would expect on the basis of MaxElide alone, with many extractions remaining ungrammatical even when the competing sluicing derivations are ruled out. We argued that this large class of cases can be explained in terms of parallelism alone, which this required a reassessment of which movement types count for the calculation of parallelism, according to which only A′-traces and head traces were taken into account. This left just a small class of extractions where ruling out sluicing did affect the grammaticality of extraction from VP-ellipsis, and we argued that these can be analysed in terms of general derivational economy.

There are three important implications of our analysis which we would like to bring attention to. First, our analysis allows us to dispense with the ellipsis-specific constraint, MaxElide. This is a welcome theoretical result, since it is not clear how such a constraint could be learned, and it is also difficult to see how it could be said to derive from general functional pressures to "say less," since there are many cases where such a constraint would seem to be inappropriate. Second, the data analyzed here provides strong evidence for a structural notion of parallelism, as the semantic definition found in the previous literature on MaxElide appeared to be too weak to account for many of the contrasts presented here. Obviously more work needs to be done on this front as there are data that seem to require a less stringent definition of parallelism. We believe that this data can be accounted for with a structural approach to parallelism with the mechanism of *accommodation* (see e.g., Fox 1999b, Thoms to appear a). Third, the data discussed here indicate that A-traces do *not* count for the calculation of parallelism, contrary to Hartman (2011)'s central claim that traces are uniform with respect to how they are interpreted at the syntax-semantics interface. Although a uniform analysis of movement is a laudable aim, we believe that this separation of A-movement from the other movement types is justified, as it is well-known that A-movement often behaves with respect to reconstruction as if it does not leave a trace (Chomsky 1995, Lasnik 1998), although it is now well-established that A-movement cannot be analysed as traceless movement altogether (Fox 1999a, Lebeaux 2009, Iatridou and Sichel 2011). Figuring out how to account for this non-uniform picture is a big topic for future research, and we speculate that examining the interaction of the two empirical phenomena considered here – reconstruction and ellipsis parallelism – may be the way to go.

References

Abels, Klaus. 2011. Don't fix that island! it ain't broke, paper presented at the Islands in Contemporary Linguistic Theory conference, Vitoria-Gasteiz.

Baltin, Mark. 2010. The nonreality of doubly-filled comps. *Linguistic Inquiry* 41:331–335.

Baltin, Mark. 2011. Deletion versus pro-forms: An overly simple dichotomy? *Natural Language and Linguistic Theory* 30:381–423.

Barros, Matthew, Elliott, Patrick, and Thoms, Gary. 2014. There is no island repair, Ms., Rutgers/UCL/University of Edinburgh.

Bjorkman, Bronwyn. 2011. Be-ing default - the morphosyntax of auxiliaries. Ph.D. thesis, MIT, Cambridge, Massachusetts.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38:589–644.

Bošković, Željko and Lasnik, Howard. 2003. On the distribution of null complementizers. *Linguistic Inquiry* 34:527–546.

Bruening, Benjamin. to appear. Alignment in syntax: Quotative inversion in English. *Syntax* .

Charlow, Simon. 2010. Can DP be a scope island? In Thomas Icard and Reinhard Muskens, eds., *ESSLLI 2008 and ESSLLI 2009 Student Sessions. Selected Papers*, volume 6211, 1–12, Lecture Notes In Computer Science.

Chomsky, Noam. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2:303–350.

Chomsky, Noam. 1986. *Barriers*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Robert Freidin, ed., *Principles and Parameters in Comparative Grammar*, 417–454, Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Ken Hale and Jay Keyser, eds., *The view from Building 20*, 1–52, Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels, and Juan Uriagereka, eds., *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–156, MIT Press.

Chung, Sandra, Ladusaw, William A., and McCloskey, James. 1995. Sluicing and Logical Form. *Natural Language Semantics* 3:239–282.

Craenenbroeck, Jeroen van. 2010. *The syntax of ellipsis: Evidence from Dutch dialects*. Oxford: Oxford University Press.

Crain, Stephen and Fodor, Janet D. 1985. How can grammars help parsers? In D. Dowty, L. Kartunnen, and A. M. Zwicky eds., *Natural language parsing: Psycholinguistic, computational and theoretical perspectives* , 94–128. Cambridge, UK: Cambridge University Press.

Embick, David and Noyer, Rolf. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.

Epstein, Samuel David. 1992. Derivational constraints on A'-chain formation. *Linguistic Inquiry* 23:235–260.

Farkas, Donka F. and Giannakidou, Anastasia. 1996. How clause-bounded is the scope of universals? In Teresa Galloway and Justin Spence, eds., *Proceedings of SALT VI*, 35–52, Ithaca, NY: CLC Publications.

Fiengo, Robert and May, Robert. 1994. *Indices and identity* Cambridge, Massachusetts: MIT Press.

Fox, Danny. 1999a. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30:157–196.

Fox, Danny. 1999b. Focus, parallelism and accommodation. In Tanya Matthews and Devon Stolovitch, eds., *Proceedings of SALT IX*. 70—90. Ithaca, NY: CLC Publications.

Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, Massachusetts: MIT Press.

Fox, Danny and Lasnik, Howard. 2003. Successive-cyclic movement and island repair: the difference between Sluicing and VP-ellipsis. *Linguistic Inquiry* 34:143–154.

Fox, Danny and Pesetsky, David. 2005. Cyclic linearization of syntactic structure. *Theoretical Linguistics* 31:1–46.

Gibson, Edward. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, T. Miyashita, & W. O'Neil eds., *Image, Language, Brain: Papers from the first Mind Articulation Project symposium* , 95-126. Cambridge, MA: MIT Press.

Grant, Margaret. 2008. A psycholinguistic investigation of maxelide in variable-binding contexts, paper presented at NELS 39.

Griffiths, James and Lipták, Anikó. 2014. Contrast and island-sensitivity in clausal ellipsis. *Syntax* 17:189–234.

Hardt, Daniel. 2006. Re-binding and the derivation of Parallelism Domains. In *Proceedings of Berkeley Linguistics Society*, volume 32, Berkeley, California, paper presented at the Berkely Linguistics Society.

Hartman, Jeremy. 2011. The semantic uniformity of traces: evidence from ellipsis parallelism. *Linguistic Inquiry* 42:367–388.

Heim, Irene. 1997. Predicates or formulas? Evidence from ellipsis. In Aaron Lawson and Enn Cho, eds., *Proceedings of SALT VII*, 197–221, Cornell University: CLC Publications.

Hendrick, Randall. 2000. Celtic initials. In Andrew Carnie and Eithne Guilfoyle, eds., *The Syntax of Verb Initial Languages*, 13–38, Oxford University Press.

Iatridou, Sabine and Sichel, Ivy. 2011. Negative DPs, A-movement and scope diminishment. *Linguistic Inquiry* 42:595–629.

Iatridou, Sabine and Zeijlstra, Hedde. 2013. Negation, modality and deontic modals. *Linguistic Inquiry* 44:529—-568.

Johnson, Kyle. 2000. How far will quantifiers go? In Roger Martin, David Michaels, and Juan Uriagereka, eds., *Step by Step: Essays on minimalist syntax in honor of Howard Lasnik*, 187–210, Cambridge, Massachusetts: MIT Press.

Kayne, Richard S. 1998. Overt vs. covert movement. *Syntax* 1:128–191.

Kennedy, Chris. 1997. Antecedent contained deletion and the syntax of quantification. *Linguistic Inquiry* 28:662–688.

Kim, Kwan-sup. 2008. English C moves downward as well as upward: An extension of Boškovič and Lasnik's (2003) approach. *Linguistic Inquiry* 39:295–307.

Kitahara, Hisatsugu. 1997. *Elementary Operations and Optimal Derivations*. Cambridge, Massachusetts: MIT Press.

Krifka, Manfred. 2006. Association with focus phrases. In V. Molnár and Susanne Winkler, eds., *The architecture of focus.*, number 82 in Studies in Generative Grammar, 105–135, Berlin: Mouton de Gruyter.

Landau, Idan. 2007. EPP extensions. *Linguistic Inquiry* 38:485–523.

Larson, Richard K. 1985. Quantifying into NP, manuscript, MIT.

Lasnik, Howard. 1998. Some reconstruction riddles. In A. Dimitriadis, ed., *Proceeding of the Penn Linguistics Colloquium*, 83–98, Philadelphia, Pennsylvania: University of Pennsylvania.

Lasnik, Howard. 1999. On feature strength: three minimalist approaches to overt movement. *Linguistic Inquiry* 30:197–217.

Lasnik, Howard. 2001. When can you save a structure by destroying it? In Minjoo Kim and Uri Strauss, eds., *Proceedings of the North East Linguistic Society 31*, 301–320, Georgetown University: GLSA.

Lasnik, Howard and Park, Myung-Kwan. 2003. The EPp and the Subject Condition under Sluicing. *Linguistic Inquiry* 34:649–660.

Lasnik, Howard and Park, Myung-Kwan. 2013. Locality and MaxElide in extraction out of elided VP. In Yoichi Miyamoto, Daiko Takahashi, Hideki Maki, Masao Ochi, Koji Sugisaki, and Asako Uchiboro, eds., *Deep Insights, Broad Perspectives: Essays in Honor of Mamoru Saito*, 235–256, Tokyo: Kaitakusha.

Lebeaux, David. 2009. *Where does Binding Theory apply?*. Cambridge, Mass.: MIT Press.

May, Robert. 1985. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.

Merchant, Jason. 2001. *The syntax of silence: sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

Merchant, Jason. 2008. Variable island repair under ellipsis. In Kyle Johnson, ed., *Topics in Ellipsis*, 132–153, Cambridge, United Kingdom: Cambridge University Press.

Pesetsky, David. 1991. Zero syntax vol.2: infinitives, Ms., MIT (available at http://web.mit.edu/linguistics/people/faculty/pesetsky/infins.pdf).

Phillips, Colin, Kazania, Nina, and Abada, Shani H. 2005. ERP effects of the processing of the processing of syntactic long-distance dependencies. *Cognitive Brain Research* 22(3):407–428.

Reich, Ingo. 2004. Association with focus and choice functions: a binding approach. *Research on Language and Computation* 2:463–489.

Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20:335–397.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman, ed., *Elements of grammar*, 281–337, Dordrecht: Kluwer Academic Publishers.

Rooth, Mats. 1992. Ellipsis redundancy and reduction redundancy. In Stepehen Berman Berman and Arild Hestvik, ed., *Proceedisn of the Stuttgart Ellipsis Workshop*, Heidelberg: IBM Germany.

Ross, John Robert. 1969. Guess who? In Robert I. Binnick, Alice Davison, Georgia M. Green, and Jerry L. Morgan, eds., *Chicago Linguistics Society*, 252–286, Chicago, Illinois.

Sauerland, Uli. 2005. DP is not a scope island. *Linguistic Inquiry* 36:303–314.

Sauerland, Uli and Elbourne, Paul. 2002. Total reconstruction, PF-movement, and the derivational order. *Linguistic Inquiry* 33:283–319.

Schuyler, Tamara. 2001. Wh-movement out of the site of VP ellipsis. In Séamas Mac Bhloscaidh, ed., *Syntax at Santa Cruz*, 1–20, UC, Santa Cruz: Linguistics Department.

Takahashi, Shoichi and Fox, Danny. 2005. MaxElide and the re-binding problem. In Effi Georgala and Jonathan Howell, eds., *Proceedings of Semantics and Linguistic Theory XV* 223–240, Ithaca, NY: CLC Publications.

Takahashi, Shoichi and Hulsey, Sarah. 2009. Wholesale late merger: beyond the A/A-bar distinction. *Linguistic Inquiry* 40:387–426.

Thoms, Gary. 2010. 'Verb-floating' and VP-ellipsis: Towards a movement account of ellipsis licensing. *Linguistic Variation Yearbook* 10:252—297.

Thoms, Gary. to appear a. Syntactic identity, parallelism and accommodated antecedents. *Lingua* .

Thoms, Gary. to appear b. From Last Resort insertion to locality: English *do*-support. In Balasz Suranyi, ed., *Minimalist Approaches to Syntactic Locality*, Cambridge: Cambridge University Press.

Thoms, Gary. to appear c. Pseudogapping, parallelism and the syntax of focus. *Syntax* .

van Craenenbroeck, Jeroen and den Dikken, Marcel. 2006. Ellipsis and EPP repair. *Linguistic Inquiry* 37:653–664.

van Craenenbroeck, Jeroen and Lipták, Anikó. 2008. On the interaction between Verb Movement and Ellipsis: New Evidence from Hungarian. In Charles B. Chang and Hannah J. Haynie, eds., *Proceedings of WCCFL XXVI*, 138–146, Somerville, Massachusetts: Cascadilla Press.

van Urk, Coppe and Richards, Norvin. 2015. Two components of long-distance extraction: Successive cyclicity in Dinka. *Linguistic Inquiry* 46(1):113—155.

Wold, Dag E. 1996. Long distance selective binding: The case of focus. In Teresa Galloway and Justin Spence, eds., *Proceedings of SALT VI*, 311–328, Ithaca, NY: CLC Publications.