VERB PHRASE ELLIPSIS: THE VIEW FROM INFORMATION STRUCTURE

LAURA KERTZ

Brown University

An analysis of antecedent mismatch effects under ellipsis is proposed to explain why some cases of verb phrase ellipsis exhibit a sizeable penalty when the elided target is not structurally matched to its antecedent, while other cases show little or no penalty at all. The proposal attributes the penalty in the former case to an information-structural constraint governing contrastive topics, and it is argued that previous accounts have misattributed that penalty to a licensing constraint on ellipsis. Results from four experiments (three off-line acceptability, one on-line self-paced reading) confirm that the relative size of the mismatch penalty can be reliably predicted based on the information structure of the clause containing the ellipsis and that acceptability differences associated with information structure are observable even in the absence of ellipsis.*

Keywords: ellipsis, parallelism, discourse, information structure, anaphora

1. INTRODUCTION. This article addresses the question of antecedent mismatch under ellipsis, a long-standing problem in the study of verbal anaphora. The 'mismatch problem', as I refer to it, has two parts. The first part of the problem is well known to linguists: VERB PHRASE ELLIPSIS (VPE) appears to require a structurally matched antecedent in order to be licensed—except in cases where it does not. The puzzle is demonstrated by data like 1–3.

(1) The driver reported the incident,
and the pedestrian did too. [report the incident]

(2) #The incident was reported by the driver,
and the pedestrian did too. [report the incident]

(3) The incident was reported by the driver,
although he didn't really need to. [report the incident]

The minimal pair in 1 and 2 differ only in the voice of the antecedent clause. In 1, the antecedent clause is active, and the VP *reported the incident* is structurally matched to the elided VP in the target (indicated in brackets). By contrast, in 2, the antecedent clause is passive, and the VP *was reported by the driver* is not structurally matched to the elided target. This mismatch between the two VPs, presumably, is the source of the violation observed for 2. The puzzle arises when we consider an example like 3, which exhibits voice mismatch comparable to that in 2, yet is not comparably degraded. The dilemma posed by data like these is a nuisance for theories of all stripes: some theories overgenerate, predicting grammaticality for 2 and 3 alike; others undergenerate, predicting ungrammaticality for both. Moreover, while a variety of proposals have been offered to account for this problem, none has in fact resolved it.

The second half of the mismatch problem has been described in the psycholinguistics literature, but has been paid scant attention in the general linguistics literature. This second problem involves the occurrence of ellipsis-like effects of mismatch in environments where there is no ellipsis. Such effects have been documented in studies that

compared the on-line processing of elliptical and nonelliptical anaphors, classified as SURFACE and DEEP, respectively, in a proposal from Hankamer and Sag (1976). In some cases, the two anaphor types show disparate behaviors, with the surface/elliptical anaphors showing a sensitivity to mismatch where the deep/nonelliptical anaphors do not. In other cases, however, the anaphor types pattern together, both showing a sensitivity to mismatch. The fact that nonelliptical anaphors have been shown to incur a mismatch penalty—one comparable to the penalty incurred under ellipsis—indicates that the mismatch problem is not unique to ellipsis. For linguistic analyses of mismatch that attribute the contrast in 1 and 2 above to a violation of the licensing conditions on ellipsis, this is problematic: either the licensing model for ellipsis is wrong, or some independent explanation for the mismatch effect outside of ellipsis is needed. For those linguistic analyses that do not invoke a matching condition on the licensing of ellipsis, the processing data are less troubling, but there is still a problem, since they fail to predict any mismatch effects at all.

In recent years there has been renewed interest in the grammar of ellipsis and the problem of antecedent mismatch. To date, however, no analysis has been proposed that would address BOTH halves of the mismatch problem. The current work offers a possible solution. I begin with a reanalysis of a generalization from Kehler (2000) that distinguishes between acceptable and unacceptable cases of antecedent mismatch. Where the relevant distinction was there cast in terms of discourse coherence, here the difference is argued to follow from information structure. Specifically, I argue that those cases of ellipsis that show a strong sensitivity to mismatch are identified by the occurrence of a contrastive topic focus in the target clause. In addition to this descriptive analysis, I posit a formal constraint that enforces topic/comment parallelism for discourse segments involved in a contrastive topic relation, as the source of the observed effects. That constraint, crucially, is not specific to ellipsis structures but instead holds for all contrastive topic relations. The analysis, as such, proposes a unified treatment of mismatch violations in both ellipsis and nonellipsis contexts.

This proposal marks a departure from previous accounts that have attributed mismatch effects to a licensing constraint on ellipsis. Where those accounts have typically invoked an identity constraint applying to VP constituents at the level of LOGICAL FORM (LF), the current proposal instead posits an information-structural constraint imposing topic/comment parallelism at the clause level. From this characterization of the issue, three distinctions emerge. First, the information-structural proposal is capable of distinguishing between cases of 'good' and 'bad' mismatch under ellipsis where an LF-identity account is not; second, the distinction proposed here is based on characteristics of overt structure (i.e. without reference to the structure of the elided VP); and finally, the proposal predicts mismatch effects for nonelliptical anaphors.

The empirical results reported here indicate that the worst cases of antecedent mismatch show effects of an information-structural violation that is not unique to ellipsis. Once the effects of that independent constraint are factored out, the penalty attributable to ellipsis turns out to be much smaller than previously assumed, raising the question of whether a categorical licensing constraint offers the best characterization of the phenomenon. The current work thus bears on a fundamental question in the study of elliptical structures: whether they should be modeled as the output of some process/ transformation or as a base-generated structure. This work also raises the question of whether the assumption of hidden structure is a useful one in the characterization of different classes of anaphora. Finally, while the proposal here is a linguistic analysis and not a processing model, it nonetheless generates novel predictions about the conditions

under which antecedent mismatch disrupts processing, as well as the time course of those effects.

2. THE MISMATCH PROBLEM. The literature on ellipsis and, specifically, the mismatch problem is largely informed by a proposal from Hankamer and Sag (1976), which argued for a distinction between two types of anaphora. I begin in this section with a review of that proposal before turning to various accounts that have succeeded it, focusing in particular on those that have addressed the mismatch problem.

2.1. THE SURFACE/DEEP HYPOTHESIS. Hankamer and Sag (1976) proposed a formal distinction between 'surface' anaphors, including verb phrase ellipsis, sluicing, gapping, and stripping, on the one hand, and 'deep' anaphors, including pronouns, sentential *it*, and null complements, on the other. The two classes are distinguished according to three criteria: (i) surface anaphors require a linguistic antecedent, while deep anaphors can be 'pragmatically controlled'; (ii) surface anaphors impose a parallelism requirement on their linguistic antecedents, while deep anaphors do not; and (iii) surface anaphors are susceptible to the 'missing antecedent' phenomenon (Grinder & Postal 1971), while deep anaphors are not.

These formal differences, it was proposed, correspond to a difference in the level of representation accessed during processing: surface anaphors require access to the surface (syntactic) representation of a contextual antecedent; deep anaphors require access either to deep structure or to some nonlinguistic element in the context. This classification was later revised (Sag & Hankamer 1984) to hold that surface anaphors require access to the logical form of an antecedent, while deep anaphors require access to a discourse model. Revised terminology was also proposed, substituting 'elliptical' for 'surface' and 'model-interpretive' for 'deep', the new terminology better reflecting the evolving theoretical commitments underlying the proposal. (As the older terminology remains entrenched, however, I adopt it here.) But the basic proposal remained unchanged and comprised two claims: one regarding the formal classification of linguistic structures, and another regarding the cognitive representations involved in interpreting those structures.

Linguistic analyses today may or may not accord with the original categorizations proposed by Hankamer and Sag. Indeed, theories today might draw no distinction, arguing that anaphora are in effect all surface or all deep. Hankamer and Sag's treatment of ellipsis as the output of a deletion operation has, however, remained influential and is the dominant theoretical approach today. By contrast, the processing claim underlying the surface/deep account (which proposed distinct mechanisms for the resolution of anaphors of different types) has not gained broad acceptance. For a time in the 1980s and 1990s a number of researchers tested the proposal experimentally, but the results from those studies failed to confirm that surface and deep anaphors are processed in a fundamentally different manner.

2.2. SURFACE EFFECTS FOR DEEP ANAPHORS. The bulk of the psycholinguistic work on the surface/deep hypothesis pursued a common strategy, comparing the effects of surface-level manipulations of antecedent form on the two types of anaphors. The results were varied and sometimes contradictory, but in considering them as a whole, two generalizations emerge. First, for on-line processing, both surface and deep anaphors are affected by surface manipulations of the antecedent, and the effect size is comparable. This was demonstrated in an early study from Murphy (1985) that tested manipulations of antecedent length, distance, and parallelism, the last of these based on voice mismatch. Second, where differences between the anaphor types were found, they re-

flected off-line judgments. For example, Tanenhaus and Carlson (1990) reported an interaction between antecedent parallelism and anaphor type by analyzing judgments in a 'makes sense' task; however, analysis of the latencies for trials where the anaphors were judged to make sense (arguably a better proxy for on-line processing) mirrored the results obtained by Murphy, with parallelism affecting both anaphor types in a comparable fashion.

The Tanenhaus and Carlson (1990) study was notable in two other respects. First, that study included two parallelism manipulations, specifically voice mismatch and syntactic category (nominal antecedent) mismatch. These two manipulations showed the same pattern of effects. Next, Tanenhaus and Carlson noted that while their off-line data show some support for the surface/deep hypothesis, their on-line findings are problematic. In explanation, they suggested that on-line effects of parallelism follow not from the syntactic effect of passivizing the antecedent, but instead from discourse effects associated with passivization. In brief, they argued that actives and passives 'map onto different discourse structures' with 'subtly contrasting logical forms' (1990:272). That view, with some modifications, is taken up and argued here. Finally, Mauner and colleagues (1995) turned up one last wrinkle in the story regarding surface effects for deep anaphors. They showed that the parallelism effect for deep anaphors occurred only in 'long passives', that is, structures that include an oblique agent in the antecedent clause. This led the authors to reject the discourse-based explanation advanced by Tanenhaus and Carlson and to propose instead that the observed parallelism effects for deep anaphors follow from awkwardness associated with expressing the oblique agent in the antecedent.

2.3. A processing explanation for ellipsis mismatch? Recent work on the processing of verb phrase ellipsis has been concerned with a very different question from that considered in the studies just described. Setting aside the comparisons between deep and surface anaphors, the goal of this recent work has been to reconcile the categorical predictions generated by linguistic analyses of ellipsis—mismatch either is grammatical or it is not—with the gradient acceptability patterns that are in fact observed. There are two ways of going about this. Under one scenario, mismatched ellipses are forbidden by the grammar but can be rendered acceptable by the processor; under the alternative, mismatched ellipses are permitted by the grammar but are disfavored by the processor.

Arregui and colleagues (2006) pursued the former approach with their recycling proposal, which holds that ellipsis interpretation involves copying an antecedent into the ellipsis site (see Frazier & Clifton 2001) and, crucially, that imperfect antecedents undergo repair. This repair introduces added processing costs, and those costs are indexed by the number of derivations necessary for converting a copy of the antecedent into the intended target structure. Additional proposals advanced in the same article argue for processing-based amelioration linked to memory, markedness, and preferences guiding antecedent search. Kim and colleagues (2011) took the alternate course, arguing that the grammar does not rule out mismatched ellipses per se, but that parser heuristics disfavor them. Those heuristics involve a general preference for structure with in-situ argument realization and an ellipsis-specific constraint preferring deletion of 'maximal' structures. The interaction of these two constraints, it is argued, accounts for the variable acceptability observed for mismatched ellipses.

The promise from these two models is that the mismatch problem—part one, at any rate—might be explained away as a side effect of processing constraints that render some ellipses more acceptable than others. The models each take different tacks, but

they share a common approach in that they predict acceptability based on the syntactic properties of the ellipsis target and, in some cases, of the antecedent. That approach, however, is not adequate to address the kinds of data that make the VPE mismatch problem so vexing. These include cases where the antecedent and the target are held constant, but acceptability varies depending on what comes in between. A minimal pair like 2 and 3, introduced above and repeated here as 4 and 5, offers a case in point.

(4) #The incident was reported by the driver,
      and the pedestrian did too. [report the incident]
(5) The incident was reported by the driver,
      although he didn't really need to. [report the incident]

In each of these examples the elided VP is a bare VP complement to an auxiliary; the antecedent clause in each case is a passive VP containing an oblique agent. As such, repair costs and parsing pressures for the two examples should be identical: neither of the gradient-acceptability models predicts a difference. Thus, neither is sufficient to address part one of the mismatch problem. Equally as important, however, because these models are specific to elliptical structure, they have nothing to say on part two. None of the factors argued to influence acceptability for mismatched ellipsis would, under these accounts, transfer to a model of interpretation for 'deep' anaphors.

2.4. LF IDENTITY AND THE MISMATCH PROBLEM. The theoretical treatment of surface anaphors has shifted over the years, adapting to conform to revised theoretical commitments (see Schwabe & Winkler 2003 for a review). Despite these adjustments, the basic claim that ellipsis is governed by a structural identity constraint has remained central to theories that model ellipsis as the output of a syntactic operation/derivation. Sag (1976), for example, recast VPE as LF deletion, as opposed to a transformation from deep to surface structure. Subsequent variations have proposed LF copying, as opposed to LF deletion, retaining the core claim that ellipsis licensing is conditioned on LF identity (Chao 1987, Lobeck 1995; see also Wasow 1972, Williams 1977, inter alia). A more recent update to LF deletion treats it as an interface phenomenon characterized as the suppression of phonological form (Chomsky & Lasnik 1993, Merchant 2001, among many others). Under this view, there is structure present in the ellipsis site (as logical form); it is just not pronounced. The key insight that unifies these proposals is that elliptical anaphors involve missing or hidden structure, and as such exhibit properties, like the mismatch effect, that distinguish them from pro-form (deep) anaphors, which involve no hidden structure.

Within the LF-identity approach, two proposals of note have worked out possible solutions to the mismatch problem. Under Fiengo and May's (1994) RECONSTRUCTION proposal, the analysis of NP traces generated by passivization was such that active and passive verb phrases, despite their apparent differences, are nonetheless LF-identical. As such, voice-mismatched ellipses do not violate the identity condition licensing ellipsis, and so are permitted by the grammar. Merchant (2008, 2013) likewise refigured the notion of LF identity such that voice-mismatched VPEs do not violate the identity condition licensing deletion. Under that proposal, the voice feature for a verb phrase is encoded on an abstract node (voice-P) dominating the verb phrase. Whether the identity condition on ellipsis is met depends on whether the constituent targeted for deletion dominates or is dominated by this abstract node: for VPE, the deleted VP is dominated by the voice-P node and so mismatch is permitted; for other types of ellipses, like sluicing and pseudo-gapping, the deleted constituent dominates/contains the voice-P node and mismatch is not possible. The proposal thus has implications for a typology of el-

liptical anaphors, predicting that 'little' ellipses, like VPE, will permit voice mismatch, while 'big' ellipses will not.

Each account offers a way around the identity constraint on ellipsis via reformulation of what 'counts' as identical at LF. Merchant's proposal makes an added prediction that does not follow from the reconstruction model, but the two approaches share a common weakness in that neither one actually resolves the mismatch problem. Rather than distinguish between good and bad mismatches, the two approaches instead shift categorical predictions such that voice-mismatched VPEs are permitted, and not prohibited, by the grammar.

2.5. BASE-GENERATION OF ELLIPSIS. An alternative approach for modeling ellipsis makes no appeal to hidden structure in the ellipsis site. This approach avoids a variety of complications associated with the LF-identity constraint, among these the proper treatment of deictics, negative polarity items, and the interpretation of bound pronominals. (See e.g. Dalrymple 1991, 2005 for a review of the arguments against hidden structure and Kennedy 2003 for a rejoinder.) Abandoning hidden structure does not, however, resolve the mismatch problem. Instead, like the variations on the LF-identity approach just described, it shifts categorical predictions.

When considering analyses that do not invoke hidden structure, it is important to point out that these do not comprise a single theory, but rather a class of theories, and like the class of LF-identity theories, these proposals are heterogeneous, invoking a variety of theoretical assumptions. As a group, these are often referred to as 'semantic licensing' proposals, to contrast with the LF-identity/'syntactic' licensing accounts. But this label suggests the application of a semantic identity constraint that licenses ellipsis, a position that is not shared across all theories. In fact, one of the advantages of the so-called 'semantic' approach is that it can admit ellipses with no antecedent: the pragmatic control cases forbidden under the surface/deep proposal, as well as category mismatches (where the antecedent to the ellipsis is a nominal constituent, as in examples 10–11, taken up below). A more apt term for grouping together the various nonsyntactic proposals is 'base-generated', and I adopt it here. This label offers a better characterization of the nonsyntactic accounts while also highlighting the more relevant contrast between the two approaches: one assumes there is LF structure that gets deleted/suppressed to form an ellipsis; the other holds that ellipsis is base-generated as either a pro-form (e.g. Hardt 1993) or a null complement to an auxiliary.

As mentioned above, assuming base-generation avoids some problems introduced by positing hidden structure, but it does not escape the mismatch problem. Nonetheless, the literature supporting a base-generation account has been the source of many of the problematic data cited in arguments against invoking hidden structure. Those arguments have typically included examples of acceptable mismatch like those considered here (Dalrymple et al. 1991, Hardt 1993), but also cases of acceptable category mismatches (Kehler 2002) and pragmatically controlled/contextually evoked ellipses (Schachter 1977, Hankamer 1978, Miller & Pullum 2012), as well as cases of split antecedents (Webber 1978, Elbourne 2008, among others). Each of these cases violates the identity condition assumed under standard LF accounts, highlighting the undergeneration problem for that approach. Many of these problematic data have been culled from corpora or naturally occurring speech, prompting some opponents of the base-generation approach to suggest that while such structures are, on occasion, produced, they are nonetheless ungrammatical (cf. Kennedy 2003) and perhaps best analyzed as speech errors (Arregui et al. 2006, Frazier & Clifton 2006). The fact that examples of this sort turn up in edited texts, in-

cluding scholarly works and news copy, however, undercuts those suggestions (see Merchant 2010 for a sizeable collection of naturally occurring data).

The great success of the base-generated approach is in admitting all of these cases that would be barred under an LF-identity approach, but this is also its failure. The base-generated approach admits good mismatches, while failing to reject the bad. What is needed is some kind of filtering mechanism that can predict when mismatch will matter. This problem for the theory, though, might also be one of its virtues, at least when part two of the mismatch problem is brought into consideration. Presumably, whatever the missing constraint is, it could explain both the conflicting data patterns observed for verb phrase ellipsis AND the occurrence of mismatch effects for deep anaphors.

2.6. DISCOURSE COHERENCE. The discourse-coherence analysis of VPE mismatch (Kehler 2000) offers a means for constraining the overgeneration of the base-generated approach. That analysis starts with the novel generalization that bad mismatches occur in discourse segments that are linked to their antecedents via RESEMBLANCE relations, while good mismatches occur in segments that are linked to their antecedents by any of a variety of other types of relation (though supporting examples overwhelmingly are drawn from the class of CAUSE-EFFECT relations). This contrast is demonstrated by the pair below, where the unacceptable mismatch in 6 occurs in a PARALLEL relation, a member of the larger class of resemblance relations, while the acceptable mismatch in 7 occurs in a VIOLATED EXPECTATION relation, a member of the cause-effect class (Kehler 2000:551, ex. 34).

(6) #This problem was looked into by John, and Bob did too.
(7) This problem was to have been looked into, but obviously nobody did.

This generalization also holds for the minimal pair introduced here, under an analysis where the coherence relations in 8–9 are resemblance and violated expectation, respectively.

(8) #The incident was reported by the driver, and the pedestrian did too.
(9) The incident was reported by the driver, although he didn't really need to.

Kehler's generalization thus serves to formally distinguish between good and bad mismatches based on coherence, attributing the difference to an interaction between ellipsis and the inferencing processes supporting the establishment of coherence. For cause-effect relations, semantic parallelism between antecedent and target is argued to be sufficient to support both the interpretation of the ellipsis and the establishment of coherence. For parallel relations, however, it is argued that semantic parallelism supports interpretation of the ellipsis, but syntactic parallelism is required to establish the intended parallel coherence relation. The claim that syntactic parallelism is necessary for computing parallel coherence receives independent support from patterns of interpretation seen for pronouns (Kehler 2000:536, Kehler et al. 2008).

The coherence analysis offers a solution for both the voice-mismatch cases considered here and also cases of category mismatch. Kehler cites, for example, the pair in 10–11 (Kehler 2000:552, ex. 40).

(10) #This letter provoked a response from Bush, and Clinton did too.
(11) This letter deserves a response, but before you do …

The coherence analysis extends further, generating predictions for the patterning of 'strict' versus 'sloppy' interpretations for pronouns in the ellipsis site. It is not clear, however, whether the proposal as framed extends to address nonellipsis data, our 'part two' problem here: elsewhere Kehler elaborates on the manner in which syntactic parallelism supports parallel coherence, without reference to ellipsis (Kehler 2002), but the

specific analysis of the ellipsis mismatch problem as described in both Kehler 2000 and 2002 invokes an 'interaction' between ellipsis resolution and the establishment of coherence.

One limitation of the coherence approach is addressed in some detail in Kehler 2002, where it is explained that there exists no set of definitive diagnostics available for determining coherence relations. Moreover, it may be the case that for some discourses, multiple relations are operative. This point is picked up in Hendriks 2004, where it is argued that the parallel relations identified by Kehler are better characterized as instances of contrastive topics, and in Kertz 2008, where it is argued that the relevant contrast involves not coherence relations but the distribution of focus in the target clause. Under the latter analysis, bad mismatches focus the subject of the target clause, while good mismatches focus the auxiliary verb, or some other nonsubject constituent in the target. Each of these thus offers a reanalysis of Kehler's basic generalization, but neither in fact explains the source of the effect. The analysis proposed by Hendriks did not address ellipsis mismatch but was instead concerned with predicting the range of interpretations available under gapping. The Kertz proposal specifically addressed mismatch, but to explain the data required the stipulation of a constraint enforcing syntactic parallelism for contrastive arguments, which, as was demonstrated in that article, is too strong.

A second critique of the coherence analysis rests on empirical grounds. While the analysis in Kehler 2002 is supported by corpus data and constructed minimal pairs, Frazier and Clifton (2006) argued that the approach is not sufficiently rigorous. Frazier and Clifton tested the predictions of the coherence account against the LF-identity approach in a series of off-line experiments, finding no evidence to demonstrate that the mismatch effect is modulated by coherence. (The experiments tested predictions about both mismatch and strict/sloppy interpretations.) It seems, however, that the types of stimuli used in those experiments were examples of the problematic cases described by Kehler, which instantiate multiple coherence relations. For example, in a stimulus set like 12–13 (Frazier & Clifton 2006:323, ex. 8), the choice of connective was intended to signal coherence relation (*just like* signals parallel; *even though* signals cause-effect), yet it seems a parallel relation might still be operative (I return to this point in §3.3).

(12) Kim looked into the problem even though/just like Lee did.

(13) The problem was looked into by Kim even though/just like Lee did.

Subsequent work by Kim and Runner (2009) has reported positive findings for an interaction between discourse coherence and voice mismatch, where cause-effect coherence minimizes, but does not eliminate, the mismatch effect. Stimuli used there also appeared to exhibit a dual coherence relation, but the choice of connective differed (*so* as opposed to *even though*).

3. An information-structural analysis. Among the proposals advanced so far to address the ellipsis mismatch problem, only the coherence analysis (Kehler 2000, 2002) can distinguish between minimal pairs that hold an ellipsis target and its antecedent constant. As described above, however, that proposal faces some challenges. The current proposal takes as a starting point Kehler's observation that there are systematic differences between relatively acceptable and relatively unacceptable cases of ellipsis mismatch and that those differences implicate discourse-level representations. I adopt here Hendriks's (2004) alternative characterization of parallel coherence relations as contrastive topics and motivate an analysis to explain why contrastive topic ellipses show the mismatch effect while other discourse relations do not.

The proposed analysis posits a clause-level constraint enforcing topic/comment parallelism on contrastive topic structures. The constraint, in brief, holds that the contrast-

ing arguments that 'anchor' a contrastive topic relation must be realized as sentence-level topics. The constraint distinguishes between the acceptable matched ellipsis in 14 and the unacceptable mismatched ellipsis in 15 based on the topic structure of the antecedent clause.

(14) The driver reported the incident, and [the pedestrian]$_F$ did too.

(15) #The incident was reported by the driver, and [the pedestrian]$_F$ did too.

In both cases, the subject/topic of the target clause (*the pedestrian*) is in focus and is interpreted contrastively with an argument in the antecedent clause, specifically 'the driver'. When *the driver* is the subject/topic in its own clause, the proposed constraint is satisfied, and the contrastive topic structure is well formed. This is the case in 14. When *the driver* appears in a nontopical oblique position, as in 15 (where *the incident* is instead the topic of the antecedent clause), the constraint is violated, and the structure is ill-formed.

The contrast in 14–15, which is predicted under the current proposal, is likewise predicted by syntactic-identity accounts that invoke hidden structure in the ellipsis site. The current proposal, however, addresses only OVERT structure, and, crucially, can distinguish between cases of good and bad mismatch. This follows because the proposed information-structural constraint is vacuously satisfied by structures that do NOT form a contrastive topic, as in 16, where the subject of the target clause is not contrastive, but instead links back to the oblique agent via coreference.

(16) The incident was reported by the driver, although he didn't [need]$_F$ to.

In this case, it is the auxiliary verb *need* that is in focus and evokes the contrast with the antecedent clause. (These focus analyses are motivated below.) Because the proposed constraint on contrastive topics is vacuously satisfied in 16, no mismatch penalty is predicted. This contrasts with the LF-identity approach, which takes into account the syntactic structure of antecedent and target VPs, as opposed to the information structure of the larger clauses containing them. LF identity thus incorrectly predicts a violation for 16 on a par with 15.

The proposal just sketched is developed in greater detail through the remainder of this section in the following steps: §3.1 reviews evidence that ellipses that exhibit the mismatch effect are contrastive topics; §3.2 introduces the proposed information-structural constraint and demonstrates that argument structure alternations induce violations of that constraint. In §3.3 I consider predictions and implications of the analysis, and then I turn briefly to a note about previous focus-based analyses of ellipsis in §3.4.

3.1. CONTRASTIVE TOPIC ELLIPSES ARE SENSITIVE TO MISMATCH. Recall that the coherence analysis (Kehler 2000, 2002) distinguished between minimal pairs based on the following generalization: unacceptable mismatches occur in parallel coherence relations, while acceptable mismatches occur in some other type of relation, typically one of a variety of cause-effect relations. That contrast is reprised in 17–18.

(17) #The incident was reported by the driver, and the pedestrian did too.

(18) The incident was reported by the driver, though he didn't really need to.

This difference in coherence is argued to underlie the difference in acceptability. But as Hendriks (2004) argued, parallel and cause-effect coherence relations also differ in how they structure the discourse: in parallel relations, each discourse segment relates independently to the prior discourse; in cause-effect relations, discourse segments 'build on' one another (Hendriks 2004:141). Hendriks applied diagnostics from Krifka 1999 to demonstrate that the parallel relations identified by Kehler form CONTRASTIVE TOPICS. Krifka identified contrastive topics as partial answers to a common discourse question

and argued that the additive particle *too* is licensed by a contrastive topic (see Krifka 1991 and citations therein). As described above, Hendriks was concerned with the range of possible interpretations for gapping structures, not the mismatch problem under ellipsis. Applying the criteria she adopts, however, to the data of interest here, we can confirm that the two conjuncts in 17 above supply partial answers to a common question, namely *who reported the incident?*. This, combined with the presence of the additive particle *too*, supports analysis of the parallel relation in 17 as a contrastive topic structure.

Further evidence supporting an analysis of Kehler's parallel relations as contrastive topics comes from arguments by analogy with a class of bona fide contrastive topics that have been much discussed in the literature. Those canonical cases of contrastive topic involve dual-focus structures, as in 19 (cf. Roberts 1996, Kadmon 2001, Büring 2003).

(19) Bill kissed Sue, and [Larry]$_{TF}$ kissed [Nina]$_F$.

The primary interest within that literature has been to distinguish between the two instances of focus. Bolinger (1961) noted that when two foci cooccur in a clause, as in 19, they are marked by distinct 'accents' or pitch contours. Subsequent work has attributed this difference to whether the focus intersects with the topic or comment portion of the utterance. When focus marks the topic of an utterance, it is referred to as a 'topic focus' (TF above); when focus marks some constituent in the comment of an utterance, it is referred to simply as 'focus' (F). The two types of focus play distinct roles in structuring the discourse, and this difference has been characterized in different ways: for example, the comment focus is 'dependent' on the topic focus (Jackendoff 1972), or the topic focus is 'fixed first' in the discourse (Roberts 1996, Büring 2003, inter alia).

The ellipsis data we are interested in here are comparable to these dual-focus contrastive topic structures, except that there is no secondary or 'dependent' focus.[1] The parallel becomes clear when we compare a dual-focus contrastive topic with a single-focus contrastive topic followed by a 'deaccented' verb phrase and a single-focus contrastive topic followed by an ellipsis (20–22, respectively).

(20) Peter ate pasta, and [Pia]$_{TF}$ ate [muesli]$_F$.
(21) Peter ate pasta, and [Pia]$_{TF}$ ate pasta too.
(22) Peter ate pasta, and [Pia]$_{TF}$ did too.

Each of the examples above shares in common the occurrence of a topic focus (TF) on *Pia*, the subject of the second conjunct. In 20, a secondary focus (F) occurs on the direct object *muesli*, forming a dual-focus contrastive topic comparable to 19 above. In 21–22 no secondary contrast occurs, and the additive particle *too* is licensed (cf. Krifka 1999: 122, ex. 48).

One important contrast between dual-focus and single-focus contrastive topics is that the former are structurally ambiguous; that is, without any additional context, it is not always possible to determine which of two foci is the topic focus. This effect is best demonstrated by considering the types of discourse questions that a contrastive topic structure might serve to answer (Roberts 1996, Büring 2003). Given the context in 23, for example, the subject NP *Pia* is the topic. Given the context in 24, the object NP *muesli* is the topic.

---

[1] Dual-focus ellipses are of course possible, and they typically involve a topic focus on the subject and a comment focus on the auxiliary verb, as in *Few states were willing to enforce the measure, but NEW YORK clearly WAS*. For the purposes of the current account, these qualify as contrastive topic structures and are predicted to pattern as such.

(23) I know what Peter ate, but what about Pia? What did Pia eat?
      [Pia]$_{TF}$ ate [muesli]$_F$.
(24) I know who ate pasta, but what about muesli? Who ate muesli?
      [Pia]$_F$ ate [muesli]$_{TF}$.

This follows because while English shows a strong correlation between topic-hood and subject-hood, that correlation is not perfect: definite NPs in subject position tend to make good sentence topics, and in out-of-the-blue contexts, the subject seems to serve as a default topic. Depending on the context, however, a nonsubject may be the topic, as in 24. For the deaccented and elided examples, this ambiguity does not exist. Indeed, we can demonstrate that focus on the subject NP is topic focus in each case by applying a test from Reinhart (1982) that explicitly marks the sentence topic using *as for*.

(25) Peter ate pasta, and as for Pia/#pasta,
      Pia ate pasta too./Pia did too.

As shown in 25, the new referent *Pia* can be introduced using *as for*. Marking the object NP *pasta* in the same way is infelicitous.

Given the foregoing, I adopt a reanalysis of Kehler's mismatch generalization in information-structural terms: ellipses that induce the greatest penalty for mismatched antecedents are those that occur in contrastive topic relations. Before turning to why this should be the case, it seems worthwhile to clarify terminology and to offer a more precise description of some of the concepts that have so far been invoked informally.

CONTRASTIVE TOPIC DEFINED. For the remainder of the discussion I draw a distinction between a 'contrastive topic focus', which describes the information-structural category of arguments that are both topic and in focus (constituents marked with the subscript 'TF'), and a 'contrastive topic relation', which describes the discourse relation holding between two clauses, one of which contains a contrastive topic focus. The former has been much discussed in the literature (see references above). To define the latter, I begin with Rooth's (1992) constraint on felicitous contrast, which he frames in terms of focus (cf. Schwarzschild 1999). That constraint holds that two utterances are felicitously contrasted when the first utterance falls within the focus semantic value of the second.

(26) CONSTRAINT ON FELICITOUS CONTRAST: A proposition B is felicitously contrasted with a preceding proposition A if the ordinary semantic value of A falls within the focus semantic value of B (cf. Rooth 1992).

This constraint captures the relationship between any two contrasting utterances. All of the data examined here—matched, mismatched, acceptable, unacceptable, elided, deaccented—have formed contrastive utterances.

To pick out the subset of contrasting utterance pairs that form contrastive topic relations, I define a contrastive topic relation as the relationship holding between two contrastive utterances, where the second utterance contains a topic focus. The definition is formulated to apply to discourse segments (construed here as a clause), not propositions, on the view that discourse segments encode a topic/comment partition, but propositions do not.

(27) CONTRASTIVE TOPIC RELATION: A discourse segment B forms a contrastive topic relation with a preceding segment A if B contains a topic focus and is felicitously contrasted with A (per 26 above).

This formulation picks out all three of the contrastive topic structures considered above—dual focus, single focus plus deaccenting, single focus plus ellipsis—as in-

stances of a contrastive topic relation. It does not, however, pick out all instances of structures containing a topic focus. For example, a sentence like 28 below would not form a contrastive topic relation, since there is no preceding discourse segment with which it contrasts.

(28) Cheri is certainly a nice enough girl, but [I]$_{TF}$ wouldn't marry her.

The topic focus in this case does not link back to a contrasting utterance, but instead plays a special role in structuring the discourse by implicating that Cheri has at least one suitor out there (see discussion in Roberts 1996). The mere presence of a topic focus, as such, does not constitute a contrastive topic relation.

All of the ellipsis data considered here contain focus in the target clause and form felicitous contrasts per the constraint in 26. But not all form contrastive topic relations. This is demonstrated by a minimal pair like 29 and 30, where each contains a focus in the target clause, but the two differ in the type of focus present. In 29, the subject NP *the pedestrian* bears a topic focus. (We can confirm that the subject is the topic using the *as for* test.) In 30, by contrast, the focused constituent is an auxiliary verb, which is not the sentence topic.[2]

(29) The driver reported the incident, and [the pedestrian]$_{TF}$ did too.

(30) The driver reported the incident, although he didn't really [need]$_F$ to.

Crucially, it is the ellipsis in the contrastive topic relation that shows degradation under mismatch, as demonstrated in 31–32.

(31) #The incident was reported by the driver, and [the pedestrian]$_{TF}$ did too.

(32) The incident was reported by the driver, although he didn't really [need]$_F$ to.

Moreover, this pattern generalizes to characterize the broad classes of data that have been used to argue for and against a syntactic identity constraint on ellipsis. Data that support a syntactic identity constraint (i.e. exhibit a sizeable mismatch penalty) are those that occur in contrastive topic relations, focusing the subject of the target clause.

(33) #The cause of the accident was investigated by the police because [the insurance company]$_{TF}$ did. (Frazier & Clifton 2006:339, ex. 9)

(34) #The student was praised by the old schoolmaster, and [the advisor]$_{TF}$ did too. (Arregui et al. 2006:241, ex. 17)

Data that go against such a constraint are those that focus the auxiliary verb in the target.

(35) This problem was to have been looked into, but obviously nobody [did]$_F$. (Kehler 2000:548, ex. 22)

(36) This information could have been released by Gorbachev, but he chose [not]$_F$ to. (Hardt 1993:37, ex. 131)

This generalization captures the basic insight from Kehler (2000, 2002) that good and bad mismatched ellipses differ in a systematic way. It recasts those differences, however, in information-structural terms. The question we are left with is why contrastive topics should behave in this way.

3.2. A CONSTRAINT ON CONTRASTIVE TOPIC RELATIONS. The data reviewed in the previous sections suggest a constraint on contrastive topic relations that induces parallelism. In this section I argue that the parallelism constraint applies not to the syntactic structure of VPs, but to information structure at the clause level (specifically the

---

[2] Presumably the topic in this sentence is the pronominal subject *he*, though a potential alternative analysis would treat the entire backgrounded proposition as topic; distinguishing between the two might depend on prior context. For now I leave open the question of how to analyze the target clause of a noncontrastive topic ellipsis, noting that whatever the sentence topic might be in 29, it is not the modal verb *need*.

topic/comment partition). I begin with the observation from Krifka (2008) that in a contrastive topic structure, topic and focus each serve their usual functions: topic identifies what the utterance is about while focus introduces alternatives—in this case, alternative topics. This set of alternative topics, or TOPIC SET, as I refer to it, has a special discourse status, which has been modeled in different ways.

THE TOPIC SET. Büring (2003), for example, models the topic set as a set of possible answers to a question in a hierarchically structured discourse: that question is presupposed by the topic focus and crucially dominates/precedes the question presupposed by the (nontopic) focus in the discourse stack.[3] The relevant restrictor defining the topic set may be explicitly evoked in the discourse, or it may remain implicit. In 37, for example, the speaker explicitly introduces the topic set, restricting it to her parents (cf. Erteschik-Shir 2007, ex. 5b).

(37) As for my parents, my mother is a teacher and my father works in an office.

Alternatively, there may be no explicit introduction of the topic set. In a case like 38, focus on *the chair* presupposes a topic set that includes 'the chair', but the speaker leaves it to the hearer to infer that *the committee* is also a member of that set.

(38) The committee apparently liked the candidate, but [the chair]$_{TF}$ was unimpressed.

The hearer may or may not infer some restrictor to that set like 'persons in a position to judge the candidate'. For our purposes it is not necessary to assume that listeners infer this restrictor. Rather, it is sufficient that they infer that there is a set and that the contrasting arguments in the two clauses are among its members.

Within the functional literature, the topic set (referred to there as a 'hypertheme') plays an additional role in structuring the discourse. Seminal work by Daneš (1974) identified a small number of well-formed 'thematic transitions' that constrain how the topic (or theme) of one sentence relates to the next. Possible transitions include a shared topic across two segments (topic chain), taking up a new topic (topic change), and introducing a new topic that is linked to the previous topic via membership in a topic set (derived theme). The derived theme as described by Daneš (1974) and later Erteschik-Shir (2007) is functionally equivalent to the contrastive topic relation described formally in the previous section, with one added piece: the link to the previous topic.

The constraint that I propose formalizes this link and states simply that for a contrastive topic relation to be well formed, both members of the topic set must be sentence-level topics.

(39) CONSTRAINT ON CONTRASTIVE TOPIC RELATIONS: A contrastive topic relation is well formed if members of the topic set are sentence topics.

This constraint captures the descriptive fact that the matched ellipsis in 40 is well formed while the mismatch in 41 is not, and it attributes this difference to the fact that *the driver* is the sentence topic in 40 but not in 41.

(40) [The driver]$_{Top}$ reported the incident, and [the pedestrian]$_{TF}$ did too.

(41) #[The incident]$_{Top}$ was reported by the driver, and [the pedestrian]$_{TF}$ did too.

This is a very different view of the problem from that of traditional LF-identity approaches, in that the current constraint invokes only overt structure in explaining the mismatch effect.

---

[3] For alternative formalizations of contrastive topic, see Ladd 1980, Vallduví 1992, Williams 1997, Steedman 2000, Wagner 2013, inter alia.

Recapping the proposal so far, I began in the previous section with Rooth's constraint on felicitous contrast (26). All of the ellipsis data examined here conform to that constraint. I then proposed a definition for a contrastive topic relation (27) and showed that only a subset of the ellipsis data forms contrastive topics. Finally, I proposed a constraint on contrastive topic relations (39) that enforces topic/comment parallelism, such that the sentence topic in each discourse segment of a contrastive topic relation is a member of the topic set. This last constraint picks out only the class of mismatched contrastive topics—our bad cases of mismatch. The analysis crucially depends on the assumption that the subject of the antecedent clause in 40 and 41 is the sentence topic. The remaining piece for the proposal is to motivate that assumption.

ARG-S ALTERNATIONS ALTER TOPIC STRUCTURE. As a default assumption, treating the subject of each clause in 40 and 41 as a topic is a sound one: in each case, the subject is referential, definite, and appears in a prominent syntactic position. As described above, however, topic-hood and subject-hood are not perfectly aligned in English. And while some diagnostics are available for specific contexts (e.g. the *as for* test), there is no sure-fire test for taking your average utterance out of the blue and determining the sentence topic (for more, see McNally 1998). Fortunately, this is not so much of a hindrance for the current proposal. This follows because we are working in a much more defined space: the information-structural effects of argument alternations are already well studied.

When considering argument structure alternations, it is important to keep in mind that there are multiple vantage points from which to describe them. One is purely syntactic. Taking the voice alternation in English as an example, we see that active- and passive-voice clauses differ in the number and type of grammatical relations supported by the verb and the structure of the VP itself. Taking the view of lexical semantics, we instead consider the mapping of thematic roles onto argument positions and how voice affects that mapping. Finally, considering the view from information structure, we know that passivization affects the relative prominence of arguments, bringing about two effects: promoting a logical object (patient/theme) to a more prominent position, and demoting the logical subject (agent) by either relegating it to a low-prominence oblique position or omitting it entirely.

Given these effects, passivization has been analyzed crosslinguistically as a form of topicalization (Siewierska 1984). The specific effects of this have been observed in English where passivization dramatically alters preferences for pronoun interpretation, a form of inference known to be sensitive to topic structure (see Garvey et al. 1975). Passivization, in this sense, is comparable to left dislocation, which displaces an argument to a clause-initial position, marking it as topic (though the syntactic means for bringing this about are distinct). Passivization has the added effect of clearly marking the demoted agent as nontopical. The *tough*-alternation has been analyzed in a comparable way. In this case, however, only one argument is involved: a logical object can appear in situ (as an object) or in a 'raised' position as subject. Comrie and Matthews (1990) proposed that in the in-situ alternate is a topic-less or 'thetic' structure. Promotion of the object marks it as topic, introducing a topic/comment partition where there was none before.

With this understanding of the information-structural effects of argument alternations, passivization in particular, we see that syntax is not the only level of representation that is affected. Argument structure, understood as the mapping of thematic roles onto argument positions, and topic structure are both implicated. When considering our question of mismatch and ellipsis, these latter two effects are crucial. It has already been shown that appealing to syntax alone is insufficient to distinguish between good

and bad cases of mismatch. However, if we instead consider how the voice alternations affect the topic structure of the antecedent clause and combine this with the proposed constraint on contrastive topic relations, an explanation of the mismatch effect emerges that picks out exactly those cases where mismatch will matter most. Argument structure mismatch leads to a violation in cases where the target clause contains a contrastive topic focus, but the contrastive topic relation between the two clauses is not supported by parallel topic/comment structure.

3.3. A RESOLUTION TO THE MISMATCH PROBLEM? The foregoing analysis offers an explanation for why a parallelism constraint on ellipsis seems sometimes to be in effect and sometimes to be flouted, by distinguishing different classes of ellipsis based on information structure. The proposal makes the same prediction, in many cases, as the coherence analysis on which it builds, but it offers a different explanation for the mismatch effect. Where Kehler appealed to an interaction between the resolution of ellipsis (an anaphoric process) and the establishment of coherence (in some cases dependent on syntax), the current proposal instead posits an information-structural constraint that imposes topic/comment parallelism for discourse segments in a contrastive topic relation.

This proposal improves on the coherence analysis by offering a simple diagnostic (in the form of the distribution of focus), which can predict whether an ellipsis will be sensitive to mismatch. Crucially, the proposal makes clear predictions for data that were problematic under the coherence approach. It predicts, for example, the null result reported by Frazier and Clifton (2006) where judgments for data like 42–43 were compared, but no effect of coherence was found.

(42) #The problem was looked into by Kim even though Lee did.
(43) #The problem was looked into by Kim just like Lee did.

In that manipulation, the choice of connective (*even though*/*just like*) was intended to signal a difference in coherence. In each case, however, the target subject was focused, evoking a contrast with the oblique argument in the antecedent clause (*Lee* contrasts with *Kim*). The current proposal predicts reduced acceptability in both cases, consistent with the observed null result. The proposal is also consistent with results from Kim and Runner (2009), which showed that while coherence, as signaled by a connective, can modulate the mismatch effect, ellipses that evoke an argument contrast consistently show reduced acceptability with mismatch. This indicates that while coherence and information structure have been confounded in past experimental work, the two may in fact make independent contributions to acceptability. (This point is taken up directly in experiment 3 here.)

In addition to the ellipsis mismatch data, the current analysis also offers insight into the more puzzling half of the mismatch problem, namely surface effects for deep anaphors. The proposal advanced here posits a general constraint applying to mismatched contrastive topics and thus predicts a penalty for contrastive topic structures with or without ellipsis. This prediction is consistent with the findings reviewed in §2 that showed effects of parallelism for processing of deep anaphors on a par with the effects observed for ellipsis. (Why ellipses and deep anaphors show different behavior off-line remains an open question.) But the prediction made here is even broader. The contrastive topic proposal predicts mismatch effects even in contexts with no anaphora. That prediction is consistent with a long-standing finding showing a facilitative effect for parallel structure in conjoined clauses. Frazier and colleagues (1984) compared on-line reading times for coordinate structures like 44 in parallel and nonparallel condi-

tions, finding increased reading times (mean RTs for the whole clause) when the second conjunct was not matched in voice to the first.[4]

(44) The adults cooked the main course, and the kids prepared the dessert.

The stimuli they tested form contrastive topic relations—indeed are examples of the 'canonical' dual-focus contrastive topic structure. The finding that such structures are sensitive to parallelism follows directly from the current account.

Thus the linguistic analysis developed here offers a potential resolution for part one of the mismatch problem, making sense of apparently contradictory patterns in the ellipsis data. It also provides a potential explanation for part two, generating predictions for apparently disparate findings in the psycholinguistics literature that show preferences for parallel structure with both nonelliptical anaphors and coordinate clauses with no anaphora (but see n. 4 for qualifications). Before turning to the empirical predictions of this proposal and the evidence supporting them, I address briefly in the following section the differences between the current account and previous analyses of ellipsis that have also invoked information structure.

3.4. An aside: focus, ellipsis, and deaccenting. The idea that ellipsis licensing is linked to focus in contrastive utterances is not new. Influential work from Rooth (1993) and Tancredi (1992) explored the links between ellipsis and deaccenting, two focus-related phenomena, setting the stage for the PF-deletion approach to ellipsis advanced by Chomsky and Lasnik (1993). There are several important differences, however, that distinguish the current account from those proposals. Chief among these is the fact that neither Rooth (1993) nor Tancredi (1992) addressed the mismatch problem. Nor, as I show below, does either of the proposals extend straightforwardly to offer a solution. Rather, the question pursued by both Rooth and Tancredi was how to explain the differences between deaccenting and ellipsis. Both appear to be licensed by focus, but ellipsis is more constrained than deaccenting, both in the range of contexts within which it can occur and the range of possible interpretations that are permitted.

The intuition that informs both Rooth's and Tancredi's proposals is that ellipsis is governed by a recoverability condition that does not apply to deaccenting. The two authors then pursued different paths in formalizing that recoverability condition. Rooth proposed that ellipsis and deaccenting are both governed by a focus-based redundancy relation, and that ellipsis is governed by an additional syntactic redundancy relation, which he implemented as an early version of the LF-identity constraint proposed in Fiengo and May's (1994) reconstruction proposal. Tancredi, by contrast, argued that both ellipsis and deaccenting are licensed by the availability of a discourse-level 'focus-based topic'. For deaccenting, this discourse object may be an entailment, but for ellipsis, it must be 'instantiated', that is, part of the asserted content.

What each proposal sought to do was to formalize the notion that a speaker may implicate a relationship between two backgrounded propositions via deaccenting, a pragmatic move that is not possible using ellipsis. Many readers will be familiar with this

---

[4] Frazier and colleagues (1984) in fact compared four alternates of these stimuli: active/active, passive/passive, and active/passive, passive/active. The slow-down induced by nonparallelism was greater in the passive/active versus the active/passive case. They suggest that this pattern might follow from discourse factors, specifically the effect of topicalizing the patient in the first conjunct. They also considered other types of nonparallelism and suggested that effects may stem from multiple sources. Subsequent work in this area has focused not only on clause-level parallelism, but also on parallelism in coordinate noun phrases (Frazier et al. 2000; see also Sturt et al. 2010). The latter, notably, is not addressed by the current work.

effect from Lakoff's (1971) famous example in 45, which implicates that calling some-one a Republican/lexicalist is a form of insult.

(45) John called Mary a Republican/a lexicalist, and then SHE insulted HIM.

Implicatures of this sort cannot be generated using ellipsis, which is permitted only if the backgrounded propositions are in fact identical. The sentence in 46, for example, can only mean that Bill called Mary a lexicalist, not that Bill insulted Mary.

(46) John called Mary a lexicalist, and then BILL did.

Rooth's syntactic identity constraint and Tancredi's instantiation requirement are both designed to capture this fact, and they do so by enforcing an added constraint on ellipsis.

It should become clear at this point why neither of these approaches is equipped to deal with the mismatch problem. Like the various formal models reviewed in §2, both Rooth's and Tancredi's proposals make categorical predictions about ellipsis mismatch. Rooth's are perhaps the most clear-cut. Given the formulation of the identity constraint he proposed, ellipsis mismatch would be categorically ungrammatical, though syntactic mismatch with deaccenting would be permitted. That account could presumably be up-dated to make use of Fiengo and May's (1994) loophole that admits voice mismatch under ellipsis. As already described, however, this does not solve the mismatch prob-lem; it merely shifts categorical predictions. Tancredi's proposal likewise makes cate-gorical predictions, though it is somewhat less clear which way those predictions run. The ambiguity arises from the treatment of the 'focus-related topic', a formal object posited by Tancredi. It appears that Tancredi intends to model it as an LF (structured syntactic/semantic) object, but the constraints that apply to it are constraints on entail-ment relations and denotational equivalence, suggesting it is a purely semantic notion. Assuming the latter, there is a categorical prediction that mismatch is possible with both ellipsis and deaccenting. Assuming Tancredi's notion of instantiation somehow imposes a syntactic identity constraint, then the prediction is that mismatch is possible with deaccenting, but not ellipsis.[5]

In summary, neither Rooth 1993 nor Tancredi 1992 was concerned with the mis-match problem, and neither analysis resolved it. Depending on implementation, both proposals could be construed to deny mismatch under ellipsis, but not deaccenting, or to admit mismatch in both ellipsis and deaccenting. Neither one, however, is able to split the difference, predicting good versus bad mismatch in ellipsis and good versus bad without. The reason neither of these proposals is capable of accounting for the mis-match facts is that neither one invokes the novel observation introduced here: ellipses may differ in where the focus in the target clause falls, and that difference is crucial for predicting acceptability under mismatch.

4. EXPERIMENTS. From the analysis developed in the previous section, two empirical predictions follow: that information structure is a reliable predictor of acceptability for mismatched ellipses, and that mismatch in a contrastive topic structure induces a penalty that is independent of ellipsis. The first prediction was tested in experiment 1, which compared off-line acceptability for matched and mismatched ellipses in con-trastive topic and simple focus structures. The second prediction was tested in experi-ment 2 by comparing the mismatched ellipsis conditions from experiment 1 to control conditions that repeated, as opposed to eliding, the target VP in the second conjunct.

---

[5] Similar ambivalence arises in Merchant 2001, where the denotational equivalence notion of Tancredi is adopted and recast as 'e-givenness'. The proposal draws the intended line between ellipsis and deaccenting, but leaves the mismatch problem for ellipsis unresolved. Merchant suggests, nonetheless, that VPs that are not matched for voice may not meet the e-givenness criterion (2001:35), with the implication that his seman-tic account would still bar mismatched ellipsis.

Experiment 3 replicated the designs from experiments 1 and 2 using a different syntactic alternation to generate matched/mismatched pairs. Experiment 3 also controlled for coherence relation, ruling out a potential alternative explanation for the results in experiments 1 and 2 based on discourse coherence. Finally, experiment 4 used on-line reading times to test whether the interaction between mismatch and information structure observed for off-line acceptability is also observed in on-line processing.

4.1. Experiment 1: focus as a predictor of mismatch acceptability. Experiment 1 tested the primary claim of the current proposal: that information structure is a reliable predictor of acceptability for mismatched ellipses. Off-line acceptability ratings were compared for matched and mismatched ellipses occurring in contrastive topic and simple focus structures. Mismatched contrastive topic ellipses were predicted to be less acceptable than all other structures.

Materials. To avoid a potential confound involving the presence/absence of a passive agent in the antecedent clause, experiment 1 made use of *tough*-alternations, which, like passivization, are characterized by promotion of a logical object/patient to subject position but do not involve demotion of a logical subject/agent. An example stimulus set is provided in Table 1. Sets were created by pairing an antecedent clause containing a raised-object (a, d) or object-in-situ (b, c) alternate with a target clause containing an ellipsis. Pairing a raised-object antecedent with a raised-object target yielded a matched antecedent condition (a), as did pairing an object-in-situ antecedent with an in-situ target (c). Alternate pairings formed the mismatch conditions (b, d).[6]

a.	Venomous snakes are easy to identify, and poisonous plants are as well.	CONTRASTIVE TOPIC	MATCH
b.	It's easy to identify venomous snakes, and poisonous plants are as well.	CONTRASTIVE TOPIC	MISMATCH
c.	It's easy to identify venomous snakes, and most experienced hikers can.	SIMPLE FOCUS	MATCH
d.	Venomous snakes are easy to identify, and most experienced hikers can.	SIMPLE FOCUS	MISMATCH

Table 1. Experiment 1 materials.

In contrastive topic conditions a contrastive argument (e.g. *poisonous plants*) was introduced in subject position in the target clause; in simple focus conditions the target subject introduced a new but noncontrastive argument (e.g. *most experienced hikers*) and was followed by an auxiliary (*can*, *do*, or *will*) encoding a focal contrast. Twenty-four stimulus sets were constructed in this way. A norming trial ($N = 30$) tested acceptability via magnitude estimation (described below) for the matched antecedent conditions to ensure that they formed a comparable baseline. (No mismatch conditions were tested in the norming trial.) The twelve stimulus sets showing the greatest variation between the two matched alternates were discarded, leaving twelve experimental stimulus sets.

Methods. Twenty-four undergraduates, all monolingual English speakers, received course credit for participation. In a magnitude-estimation paradigm, participants were asked to provide acceptability ratings for stimulus sentences as compared to a fixed

---

[6] *Language* editor Greg Carlson raised the question of whether the presence/absence of a comma before *as well* (or *too* in other stimulus sets) might affect the current result. No comma was used in the stimuli, and I cannot comment on how this omission affected the results. In consulting style guides I have found no consistent guidelines and thus cannot comment on whether readers might expect a comma. APA style disprefers the punctuation, unless needed; the Chicago Manual of Style suggests the comma is optional, unless there is an abrupt change of thought.

'modulus'. The modulus was a grammatical sentence with conjoined clauses and no ellipsis (*Ana solved the puzzle yesterday, and Carol solved it too*) with a rating fixed at 100. Statistical analysis was performed on the log-transformed ratio of the stimulus rating as compared to the modulus; means for these transformed scores are reported with results. A within-participants design was used with stimuli balanced across lists in a Latin square. For each list, experimental stimuli were presented in pseudo-random order and were interleaved with other experimental stimuli (of varying degrees of predicted acceptability) and with grammatical fillers. Methods were identical for experiments 1–3, except as noted.

STATISTICAL ANALYSIS. Multiple linear mixed-effect (LME) models were fitted to the data using the R software platform. Models with differing random-effects structures were compared via forward model selection, starting with the simplest model, which contained only random intercepts for participants. In subsequent models, effects were added in the following order: random slopes for participants, random interactions for participants, random intercepts for items, random slopes for items, and random interactions for items. An analysis of variance was conducted to compare models for goodness of fit. The model with the simplest random-effects structure providing significant improvement in fit over other models was selected, and statistical results for fixed effects are reported for that model. Fixed effects with absolute t-values greater than 2 are reported as significant. Planned comparisons were carried out via subset analysis (fitting a simple LME model with random intercepts for participants to data for one level of a condition), and p-values for those comparisons were obtained via Markov chain Monte Carlo (MCMC) sampling. This method of analysis was used for experiments 1–3.

RESULTS AND DISCUSSION. Condition means and standard error for acceptability scores, calculated as described above, are reported in Table 2. Ratings for the matched conditions were comparable, while means for the mismatched conditions diverged, with the mismatched contrastive topic condition receiving the lowest ratings. The filler mean fell between condition means and is included for comparison.

	MATCH	MISMATCH
SIMPLE FOCUS	.02 (.05)	−.13 (.04)
CONTRASTIVE TOPIC	−.02 (.02)	−.34 (.08)
FILLERS	.03 (.03)	

TABLE 2. Experiment 1 acceptability ratings with standard error.

Statistical analysis showed reliable main effects of mismatch ($|t| = 4.56$) and of focus ($t = 3.23$), as well as the predicted interaction between mismatch and focus ($t = 2.20$). Parameter estimates and t-values are reported in Table 3. The random-effects structure for the statistical model, selected via the procedures outlined above, included random intercepts, slopes, and interactions for participants. Subset analysis showed that the mismatch effect was reliable ($p < 0.001$) within both levels of focus, while the focus effect was reliable ($p < 0.001$) only in the mismatch condition.

	ESTIMATE	t
MISMATCH	−0.24 (.05)	−4.56*
FOCUS	0.12 (.04)	3.23*
MISMATCH : FOCUS	0.17 (.08)	2.20*

TABLE 3. Experiment 1 model estimates with standard error and t-values.

The absence of an effect of focus in the matched conditions confirms the success of the norming phase that was carried out to ensure that the matched antecedent conditions formed similar baselines for comparison. More importantly, the positive finding of an interaction between mismatch and focus supports the claim made here that focus is a reliable predictor of acceptability for mismatch. The results also show, however, that the simple focus condition is not immune to mismatch, as ellipses in the simple focus condition showed a numerically smaller, but still statistically reliable, effect.

4.2. Experiment 2: mismatch acceptability with repeated VPs. The linguistic analysis proposed here attributes the mismatch effect observed under ellipsis to a general (not ellipsis-specific) discourse constraint on contrastive topic relations, with the prediction that differences in mismatch acceptability will be observed in both ellipsis and nonellipsis structures. Experiment 2 tested that prediction by comparing acceptability for mismatched structures with contrastive topic versus simple focus in both ellipsis and repeated VP conditions.

Materials. Mismatched ellipsis conditions from experiment 1 were paired with control conditions that repeated, as opposed to eliding, the target VP. The resulting design crossed two levels of information structure (contrastive topic vs. simple focus) and two levels of ellipsis (ellipsis, no ellipsis). In all conditions, the target VP was not matched to antecedent (i.e. all mismatch conditions). An example stimulus set is provided in Table 4.

a.	It's easy to identify venomous snakes, and poisonous plants are as well.	CONTRASTIVE TOPIC	ELLIPSIS
b.	It's easy to identify venomous snakes, and poisonous plants are easy to identify as well.	CONTRASTIVE TOPIC	NO ELLIPSIS
c.	Venomous snakes are easy to identify, and most experienced hikers can.	SIMPLE FOCUS	ELLIPSIS
d.	Venomous snakes are easy to identify, and most experienced hikers can identify them.	SIMPLE FOCUS	NO ELLIPSIS

Table 4. Experiment 2 materials.

Methods and analysis. Methods and statistical analysis were identical to those in experiment 1. (Participant $N = 38$; item $N = 10$; modulus = *Ana solved the puzzle yesterday, and Carol eventually solved it too.*)

Results and discussion. Condition means and standard error for acceptability scores are reported in Table 5. Contrastive topic conditions were rated lower than simple focus conditions in both the ellipsis and no-ellipsis conditions. The difference was larger, however, in ellipsis conditions. Ellipsis condition means were lower than no-ellipsis means. The filler mean exceeded all condition means and is included for comparison.

	ELLIPSIS	NO ELLIPSIS
SIMPLE FOCUS	−.13 (.03)	.00 (.03)
CONTRASTIVE TOPIC	−.27 (.05)	−.05 (.02)
FILLERS	.07 (.03)	

Table 5. Experiment 2 acceptability ratings with standard error.

Statistical analysis showed reliable main effects of focus ($|t| = 3.87$) and of ellipsis ($|t| = 7.04$). While the data suggest a possible interaction, it was statistically marginal ($t = 1.91$). Parameter estimates are reported in Table 6; the model included random intercepts for participants. Inclusion of random slopes for participants improved model fit

but also led to overparameterization. Subset analysis showed that the focus effect was reliable ($p < 0.01$) within ellipsis conditions and marginal ($p = 0.06$) within no-ellipsis conditions.

	ESTIMATE	t
FOCUS	−0.09 (.02)	−3.87*
ELLIPSIS	−0.17 (.02)	−7.04*
FOCUS : ELLIPSIS	−0.09 (.05)	−1.91

TABLE 6. Experiment 2 model estimates with standard error and *t*-values.

These results demonstrate that the difference in mismatch acceptability observed for contrastive topic and simple focus ellipses persists even when the target VP is repeated, as opposed to elided. This finding confirms that the information-structural differences identified here are not confined to ellipsis contexts. However, for both types of structure (that is, both contrastive topic and simple focus), the ellipses were rated less acceptable than the repeated-VP structures. As such, the reduced acceptability observed in the lowest-rated condition—that is, the contrastive topic ellipsis condition—follows from the cumulative effects of a penalty associated with a mismatched contrastive topic and a penalty associated with ellipsis.

While this pattern is consistent with the current proposal, it raises the question of what is causing the added ellipsis penalty. One possibility is that it reflects violation of an LF-identity constraint. This would explain why the penalty arises in the ellipsis conditions, but it raises some practical concerns about the nature of the violation. Traditionally, the LF-identity constraint has been posited as a licensing constraint: identity is required to trigger deletion, and without it, the structure is ungrammatical. Here we see that the ellipsis penalty—whatever its source—is a small one, and it must be combined with an additional penalty, that is, the parallelism constraint on contrastive topics, in order to bring about the pronounced loss of acceptability that has been cited as evidence in support of an identity constraint on ellipsis.

4.3. EXPERIMENT 3: CONTROLLING FOR COHERENCE. Experiments 1 and 2 confirmed the basic claims of the current proposal, that information structure is a reliable predictor of acceptability for mismatched ellipses and that information-structural differences in mismatch acceptability are detectable whether the target VP is elided or repeated. Experiment 3 replicated the manipulations from experiments 1 and 2 with new materials. These new materials were designed to achieve three goals: to control for discourse coherence, to demonstrate that the effects generalize to different types of argument alternations, and to control for other potential confounds in the previous stimuli, as described below.

MATERIALS. Stimulus sets were formed from biclausal structures where an ellipsis target was embedded within an equative clause with an adverbial head (e.g. *as quickly as* ...). An example stimulus set is provided in Table 7. The matrix clause served as the antecedent for the ellipsis and appeared in either active or passive voice. Pairing an active-voice antecedent with an active-voice target formed the matched antecedent conditions, as did pairing a passive-voice antecedent with a passive target. In the contrastive topic conditions a contrastive argument was introduced in the subject position of the target clause (e.g. *the engineers*); in the simple focus conditions the target subject was coreferent with the logical object/patient of the antecedent clause (i.e. *it* refers to *the line*). In ellipsis conditions, the target VP was elided. In no-ellipsis conditions, the lexical head of

the target VP was repeated followed by an object pronoun. Conditions a–d replicate the design of experiment 1; conditions c–f replicate the design of experiment 2.

a.	The technicians didn't install the line as quickly as the engineers did.	CONTRASTIVE TOPIC, MATCH	ELLIPSIS
b.	The line wasn't installed by the technicians as quickly as it could have been.	SIMPLE FOCUS, MATCH	ELLIPSIS
c.	The line wasn't installed by the technicians as quickly as the engineers did.	CONTRASTIVE TOPIC, MISMATCH	ELLIPSIS
d.	The technicians didn't install the line as quickly as it could have been.	SIMPLE FOCUS, MISMATCH	ELLIPSIS
e.	The line wasn't installed by the technicians as quickly as the engineers installed it.	CONTRASTIVE TOPIC, MISMATCH	NO ELLIPSIS
f.	The technicians didn't install the line as quickly as it could have been installed	SIMPLE FOCUS, MISMATCH	NO ELLIPSIS

TABLE 7. Experiment 3 materials.

In addition, these materials have the following important features. First, the use of an equative structure explicitly marks parallel coherence for all conditions. This resolves a confound from experiments 1 and 2, where coherence and information structure covaried (all contrastive topic conditions exhibited parallel coherence). Next, in conditions where the antecedent clause was passive, an oblique agent was always realized (i.e. only long passives were used). Finally, the size of the ellipsis target was held constant across all conditions: whereas in experiments 1 and 2 the elided constituent was either VP or AP, in experiment 3 the elided constituent was always VP.

There is one further feature of the materials used here that warrants some discussion, as it pertains to a proposal from Arregui et al. 2006 regarding gradient acceptability patterns for mismatched ellipsis: the 'systematic paraphrase hypothesis' predicts variable mismatch effects depending on the relative markedness of antecedent and target (where markedness is characterized derivationally). That proposal offers a potential alternative explanation for the pattern of results observed in experiment 1, where simple focus conditions always featured an 'unmarked' object-in-situ structure in the target clause. In experiment 3, that pattern is reversed, however, with the target in the simple focus condition appearing as a 'marked' passive. The paraphrase hypothesis thus predicts lowest acceptability for the mismatched simple focus ellipses (condition d) for experiment 3, while the current proposal predicts the lowest acceptability for the mismatched contrastive topic ellipses (condition c). (For experimental results directly testing the role of markedness, and for related discussion, see Kertz 2010:42–45.)

RESULTS AND DISCUSSION. Results from experiment 3 replicated the overall pattern of findings from experiments 1 and 2. This is demonstrated in Figures 1 and 2, which depict condition means for acceptability scores across all three experiments. In experiments 1 and 3, ratings for the contrastive topic and simple focus conditions were comparable in the matched ellipsis condition, but diverged in the mismatch conditions. In experiments 2 and 3, nonellipsis control conditions showed a disparity between mismatched contrastive topics and mismatched simple focus structures. Likewise, across both experiments, the difference between the two information structure conditions was greater under ellipsis. Replication of these findings confirms that the effects of experiments 1–2 are due to information structure, and not to discourse coherence. The results here also rule out a potential alternative explanation for experiment 1 based on relative markedness.
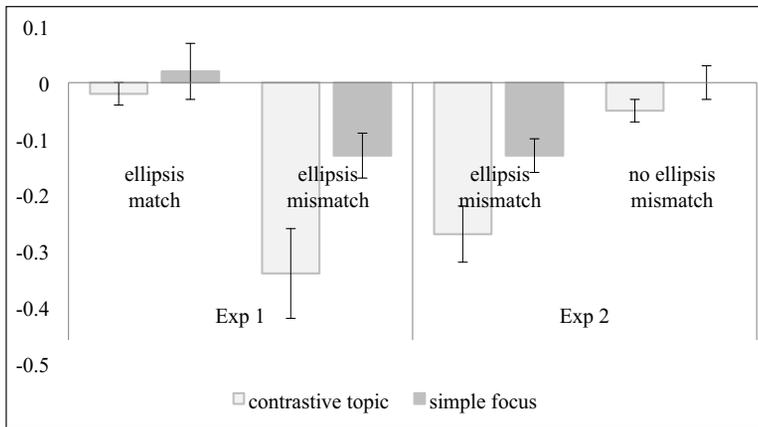
FIGURE 1. Acceptability ratings for experiments 1 and 2 (*tough*-alternates).
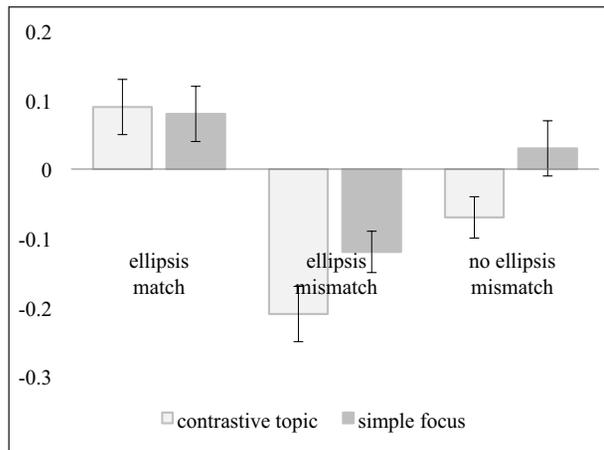

FIGURE 2. Acceptability ratings for experiment 3 (voice alternates).

Condition means and standard error for experiment 3 are reported in Table 8. Statistical analysis for experiment 3 showed reliable main effects of mismatch ($|t| = 5.19$) and of ellipsis ($t = 4.70$), as well as a reliable interaction between mismatch and focus ($t = 2.25$). There was no main effect of focus, and no interaction between focus and ellipsis. Results from the statistical model, which included random intercepts and slopes for participants, are reported in Table 9.

	ELLIPSIS MATCH	ELLIPSIS MISMATCH	NO ELLIPSIS MISMATCH
SIMPLE FOCUS	.08 (.04)	−.12 (.03)	.03 (.04)
CONTRASTIVE TOPIC	.09 (.04)	−.21 (.04)	−.07 (.03)
FILLERS	.02 (.03)		

TABLE 8. Experiment 3 acceptability ratings with standard error.

While results from experiment 3 were generally consistent with the results for experiments 1 and 2, some small differences were observed. First, overall ratings were higher

	ESTIMATE	t
MISMATCH	−0.25 (.05)	−5.19*
FOCUS	0.04 (.03)	1.52
ELLIPSIS	0.14 (.03)	4.70*
MISMATCH : FOCUS	0.10 (.04)	2.25*
FOCUS : ELLIPSIS	0.01 (.04)	0.20

TABLE 9. Experiment 3 model estimates with standard error and t-values.

in experiment 3 as compared to experiments 1 and 2, indicating that the type of alternation used to generate matched/mismatched pairs can influence acceptability. Next, where experiment 1 showed a main effect of focus, a subset analysis of the ellipsis conditions from experiment 3 did not show that effect, though the predicted interaction between focus and mismatch was replicated. Finally, while the interaction between ellipsis and focus was not statistically reliable in either experiment 2 or 3, the absolute t-value for experiment 2 (1.91) was considerably larger than the value obtained in experiment 3 (0.20). This last point is informative when we consider the relationship between ellipsis and the information structure manipulation, since it indicates that in experiment 3, ellipsis contributed a merely additive penalty, while in experiment 2, there is the possibility that ellipsis may have exacerbated an already existing penalty.

From these three experiments we can conclude that at least part of the penalty previously attributed to violations of an LF-identity constraint on ellipsis in fact follows from an information-structural constraint that is orthogonal to ellipsis. When we factor out this independent penalty, the differences left to be accounted for are much smaller than would be suggested by the types of minimal pairs cited in the literature as evidence of LF-identity violations. A similar finding was reported by Duffield and Matsuo (2003), who showed that ellipses with finite targets were routinely judged less acceptable than those with infinitival targets, whether matched or not. Findings such as these raise the possibility that penalties observed for 'bad' ellipses can be parceled out across a variety of different factors that do not directly implicate ellipsis licensing. Duffield and Matsuo interpreted this as a challenge to the categorical nature of ellipsis licensing, a view that is consistent with the larger stance taken here. Returning to the current results, however, while we see clear evidence of a general discourse penalty, it is also clear that the ellipsis conditions are incurring some added penalty. I postpone a fuller discussion of this point until the general discussion.

4.4. EXPERIMENT 4: MISMATCH EFFECTS IN ON-LINE PROCESSING. Having demonstrated in experiments 1–3 that the relative acceptability of mismatched structures under ellipsis—and even in the absence of ellipsis—is conditioned on information structure, experiment 4 turns to a consideration of on-line processing. The first question addressed here is a simple one: whether off-line acceptability patterns correlate with ease of processing. Under the assumption that they do, the next question is one of timing: at what point in processing does the mismatch effect emerge? A natural place to look for such an effect is following the ellipsis. A prediction that follows directly from the off-line data presented already is that mismatch will lead to increased reading times post-ellipsis, and that mismatch will disproportionately affect contrastive topic relations.

Based on the linguistic analysis offered here to explain the source of that interaction, however, a second, perhaps less expected, prediction emerges. Recall that the current proposal attributes the mismatch effect in contrastive topic structures to a defective topic transition; that is, the violation observed in those cases is due not to the mismatch in syntactic structure between the elided VP and its antecedent, but rather to the non-

topical status of one member of the topic set in an intended contrastive topic relation. If it is the case that the violation stems from a defective topic transition, one would expect evidence of that violation to be detectable quite early—as early as the subject of the target clause. Such a finding would show that mismatch can disrupt processing even before the reader encounters the ellipsis, strong evidence that the constraint enforcing parallelism is independent of ellipsis.

MATERIALS. Experiment 4 followed the design of experiment 1, in which two levels of information structure (contrastive topic/simple focus) were crossed with two levels of antecedent form (match/mismatch). Stimuli from experiment 1 were adapted to include a five-word spillover region following the ellipsis, as shown in Table 10.

a.	Venomous snakes are fairly easy / to identify, / and / poisonous plants / are also, / as far / as I know.	CONTRASTIVE TOPIC	MATCH
b.	It's fairly easy to identify / venomous snakes, / and / poisonous plants / are also, / as far / as I know.	CONTRASTIVE TOPIC	MISMATCH
c.	It's fairly easy to identify / venomous snakes, / and / most experienced hikers / can, / as far /as I know.	SIMPLE FOCUS	MATCH
d.	Venomous snakes are fairly easy / to identify, / and / most experienced hikers / can, /as far / as I know.	SIMPLE FOCUS	MISMATCH

TABLE 10. Experiment 4 materials.

Regions for analysis are indicated with slashes and were defined as follows: the final phrase of the antecedent clause contained a two-word infinitival phrase in raised-object conditions (e.g. *to identify*) and a two-word noun phrase in object-in-situ conditions (e.g. *venomous snakes*); the conjunction (*and*) constituted a single-word region introducing the target clause; the target subject region contained an adjective and noun (e.g. *poisonous plants*) in the contrastive topic condition and a quantifier, adjective, and noun in the simple focus condition (e.g. *most experienced hikers*); the VP region consisted of an auxiliary verb plus an additive particle in the contrastive topic condition (e.g. *are also*) and an auxiliary verb alone in the simple focus condition (e.g. *can*); the post-ellipsis region contained the first two words following the ellipsis site (e.g. *as far*).

METHODS. In a self-paced reading task with a moving window display, participants pressed the space bar to advance through stimuli one word at a time. The display showed dashes in place of characters with whitespace between words. Participants were thus able to gauge how far along in the sentence they were and the length of upcoming words, but were not able to view upcoming words or words already read. A yes/no question testing comprehension was asked after each stimulus, and participants received immediate feedback indicating whether they had answered the question correctly. Participants were instructed to read as naturally as possible, making sure that they understood what was read. Participants were instructed to pay attention to feedback on answers to comprehension questions and to treat negative feedback as a cue to read more carefully. Practice items and questions were presented prior to the experiment, which took roughly thirty minutes to complete. (Participant $N = 48$; item $N = 12$.)

STATISTICAL ANALYSIS. Statistical analyses were computed for comprehension scores and for reading times by fitting mixed-effects linear models in the manner described for experiments 1–3. All trials were included in the comprehension analysis; only correct comprehension trials were included in the reading-time analysis. For the reading-time analysis, measurements above 2,000 ms were discarded. Means and standard deviations

were computed for each combination of region and condition, and measurements falling more than three standard deviations from the mean were removed. This procedure resulted in a loss of 1.8% of the analyzable data.

To adjust for differences in word length across conditions, residual reading times were used as the dependent variable in statistical analysis. Residual reading times were obtained by computing a regression equation predicting reading time based on word length for each participant and then subtracting the reading time predicted by the participant's regression equation from the recorded reading time. Residual reading times were averaged across words in critical regions, and model fitting and comparison were carried out for each region of interest. For regions in the target clause, statistical models crossed two fixed effects: antecedent form (match/mismatch) and information structure (contrastive topic/simple focus). For the final phrase in the antecedent clause, statistical models included the form of the antecedent (in-situ/raised object) as the sole fixed effect. For the conjunction introducing the target clause both of these models were tested, as described below. Details of random-effects structures for each analysis are reported with model results.

Results: comprehension. Overall means for comprehension scores of experimental items was 81%. Scores were generally high across items and across conditions, with the exception of one item set for which scores were extremely low (8%). That item set was retained in the analysis of comprehension scores (exclusion raises the comprehension score mean to 88%) but was excluded from the reading-time analysis. Means and standard error for comprehension scores are reported in Table 11. The statistical model used to analyze comprehension scores included random intercepts for participants. Parameter estimates with standard error and t-values for fixed effects are reported in Table 12.

	MATCH	MISMATCH
CONTRASTIVE TOPIC	.91 (.02)	.94 (.02)
SIMPLE FOCUS	.83 (.03)	.86 (.03)

Table 11. Experiment 4 comprehension scores with standard error.

	ESTIMATE	t
MATCH	3.47 (2.58)	1.35
FOCUS	−6.94 (2.58)	−2.69*
MATCH : FOCUS	1.39 (5.16)	0.27

Table 12. Experiment 4 model estimates for analysis of comprehension scores with standard error and t-values.

Comprehension scores were higher ($|t| = 2.69$) in contrastive topic conditions as compared to simple focus. There was no effect of mismatch and no interaction between mismatch and focus. The fact that there was no main effect of mismatch confirms that the syntactic manipulation did not affect participants' ability to understand the stimuli. The main effect of focus suggests that participants may have had an easier time interpreting the contrastive topic structures. In interpreting this result, however, we must keep in mind that the focus manipulation involved changing the lexical items used to make up the target clause, and as a result it was necessary to construct two sets of comprehension questions, one for each focus condition. As such, the observed focus effect may indicate a difference in difficulty of the questions used to test for accuracy, as opposed to a true effect of focus on ease of comprehension. (This was unavoidable, since using a single set would have precluded questioning the content of the target clause, which might, in

turn, have led some participants to fail to attend to the target.) Crucially, there was no interaction, and accuracy was generally high.

RESULTS: READING TIME. Raw reading times are plotted in Figure 3, which shows word-by-word reading-time measures beginning at the final phrase of the antecedent up until two words following the ellipsis in the target clause. Residual reading times, adjusted for word length and participant reading rate, and averaged across regions as described above, are reported in Table 13 below.
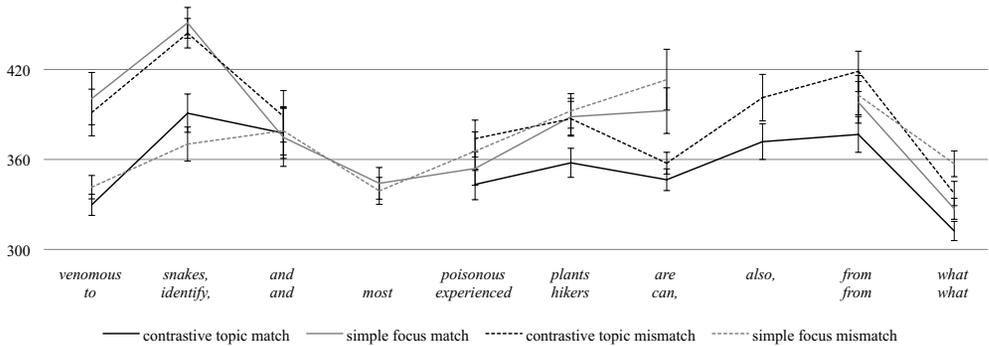

FIGURE 3. Raw word-by-word reading times (ms) for experiment 4.

As seen in Fig. 3, reading times at the end of the antecedent clause were influenced by the form of the antecedent. Reading times for the noun phrase *poisonous plants* (which appeared clause-finally in the contrastive topic mismatch and simple focus match conditions) were longer than reading times for the infinitival verb phrase *to identify* (which appeared clause-finally in the contrastive topic match and simple focus mismatch conditions). That difference is visible for the raw reading times charted in the figure, and analysis of residual reading times confirms a reliable main effect of antecedent type (raised object vs. in-situ) when differences in word length and participant reading rate are taken into account ($|t| = 1.38$). Random-effects structure for the statistical model was fully specified (random intercepts, slopes, and interactions for participants and items).

	FINAL PHRASE OF ANTECEDENT	CONJUNCTION	TARGET SUBJECT	TARGET VP	POST-ELLIPSIS
CONTRASTIVE TOPIC, MATCH	−35 (8)	−2 (11)	−61 (7)	−22 (7)	−47 (8)
CONTRASTIVE TOPIC, MISMATCH	2 (10)	14 (9)	−25 (7)	4 (8)	−8 (8)
SIMPLE FOCUS, MATCH	16 (13)	4 (9)	−42 (7)	19 (13)	−20 (7)
SIMPLE FOCUS, MISMATCH	−38 (8)	−2 (9)	−50 (8)	33 (19)	−14 (7)

TABLE 13. Experiment 4 mean residual reading times with standard error by region.

While reading times in the antecedent clause diverged based on the form of the antecedent, they converged at the conjunction introducing the target clause, which is perhaps not surprising, given that the conjunction is a high-frequency function word. Two statistical analyses were carried out for residual reading times at the conjunction. The first tested for a simple effect of antecedent form (raised object vs. in-situ), comparable to the model tested for the antecedent clause. That analysis (which included random intercepts for participants) showed no effect of antecedent form ($|t| = 1.13$). The second model tested for effects of the experimental manipulations used in the target clause, that

is, focus crossed with match. Any effects of those manipulations at the conjunction would be spurious (because those manipulations involved material coming later in the clause), and the analysis was conducted to rule out the possibility of spurious effects influencing later regions of interest. Notably, the difference in residual reading times at the conjunction for the contrastive topic condition was larger than the difference for the simple focus conditions (16 ms vs. 6 ms; see Table 13). Variability in this region was also high, however, with standard error for residual reading times ranging from 9 to 11 ms across conditions. LME analysis showed no reliable effects of focus or syntax, and no reliable interaction ($|t| = 1.09$). Results for the optimal model, which included random intercepts for participants, are reported in Table 14. The highest t-value observed for the interaction, across all models tested, was 1.12.

	CONJUNCTION		SUBJECT		VP		POST-ELLIPSIS	
	ESTIMATE	t	ESTIMATE	t	ESTIMATE	t	ESTIMATE	t
MATCH	5.18 (9.77)	0.53	2.60 (8.41)	0.31	20.88 (15.08)	1.38	20.34 (9.36)	2.17*
FOCUS	–5.45 (9.77)	–0.56	–2.03 (6.45)	–0.32	37.98 (12.89)	2.95	3.65 (7.05)	0.52
MATCH : FOCUS	–21.26 (19.55)	–1.09	–36.46 (12.60)	–2.89*	–8.37 (31.17)	–0.27	–36.29 (18.53)	–1.96

TABLE 14. Experiment 4 model estimates with standard error and t-values for analysis of residual reading times in target-clause regions.

The first reliable effects of the target-clause manipulations emerged at the subject noun phrase of the target, and those effects followed the predicted interactive pattern, with contrastive topic structures showing a larger effect of mismatch than simple focus structures. As seen in Fig. 3, reading times for *most experienced hikers*, the subject in the simple focus conditions, are comparable for matched and mismatched conditions (4–10 ms difference in raw reading times across words). Meanwhile, reading times for *poisonous plants* were faster (by 30 ms at each word) in matched antecedent conditions, where the noun phrase *venomous snakes* appeared as a subject in the antecedent clause, as compared to mismatched conditions, where *venomous snakes* appeared as a direct object of an embedded verb in the antecedent. Statistical analysis of mean residual reading times for the subject noun phrase region showed a reliable interaction between match and focus ($|t| = 2.89$), with no reliable main effects. Results from the statistical model, which included random intercepts and slopes for participants, are reported in Table 14.

Turning next to the verb phrase region, visual inspection of raw reading times in Fig. 3 suggests the first evidence of a mismatch effect for simple focus structures. Analysis of residual reading times in this region, however, consistently showed reliable effects of focus (longer reading times for simple focus conditions), but no reliable effect of mismatch. Three analyses were conducted for this region. The first averaged over words in the VP region as described above. That model showed a reliable effect of the focus manipulation ($t = 2.95$) but no effect of match ($t = 1.38$). The model included random intercepts, slopes, and interactions for participants and is reported in Table 14. The absence of a main effect of mismatch was somewhat surprising for this region, and two follow-up analyses were conducted. The first compared residual reading times for the auxiliary verbs alone (no averaging in the contrastive topic condition). That analysis likewise showed a reliable effect of focus ($t = 3.58$), but not match ($t = 1.11$). A final analysis, this one comparing residual reading times at the word immediately preceding the ellipsis site (the auxiliary verb, e.g. *can*, in simple focus conditions, and the additive particle, e.g. *also*, in contrastive topic conditions), also showed a reliable effect of focus

($t = 2.19$), but not mismatch ($t = 1.63$). Models for each of these follow-up analyses included random intercepts, slopes, and interactions for participants.

In the post-ellipsis region, the predicted interaction between focus and mismatch was once more observed. As seen in Fig. 3, contrastive topic conditions showed a sizeable difference between matched and mismatched conditions for the two words immediately following the ellipsis; simple focus conditions showed a small difference at the first word post-ellipsis, growing larger by the second word. Statistical analysis of residual reading times averaged across both words showed a reliable effect of mismatch ($t = 2.68$) and a reliable interaction between focus and mismatch ($t = 2.33$). That model included random intercepts for participants and no other random effects. Model comparison showed no improvement in fit with inclusion of additional random effects, though t-values were diminished. The most conservative estimate for t-values comes from a fully specified model (containing random intercepts, slopes, and interactions by participants and by items). In that model, also reported in Table 14, the t-value for mismatch remained reliable at 2.17, while the t-value for the interaction fell below significance to 1.96. Statistical analysis of just the first word post-ellipsis, using a fully specified model, showed a marginal effect of mismatch ($t = 1.81$) and a reliable interaction ($|t| = 2.37$). Statistical analysis of the second word post-ellipsis, using a fully specified model, showed reliable effects of mismatch ($t = 3.27$) and of focus ($t = 2.60$), but no interaction. These effects can be seen in Fig. 3, where the contrastive topic condition shows a mismatch penalty immediately following the ellipsis site, but the simple focus condition shows a penalty one word later.

DISCUSSION. The on-line results from experiment 4 confirm that the patterns of relative acceptability observed off-line in experiments 1 and 3 correlate with on-line ease of processing. Specifically, the expected interaction between mismatch and focus was observed immediately following the ellipsis, in confirmation of the current account. Further evidence consistent with the current proposal was found in the pre-ellipsis region. Analysis of that region showed that for contrastive topic ellipses, the time spent reading the target subject, for example, *poisonous plants*, depended on the structure of the antecedent. When the contrasting argument (*venomous snakes*) appeared as a subject in the preceding clause, reading times were faster; when it appeared as an object, reading times were slower. For simple focus ellipses, reading times at the target subject were not affected in this way. Both of these findings support the analysis proposed here, which holds that the ellipsis mismatch effect implicates a defective topic transition.

There are additional points worthy of mention, however, with respect to the pre-ellipsis result. These involve issues related to both finding and interpreting that result. First, detecting the early (pre-ellipsis) mismatch effect was supported by the specific information-structural properties of the *tough*-construction alternation. Since the in-situ variant of a *tough*-alternation is a thetic or topic-less structure, it cannot serve as the first segment in a contrastive topic transition. Immediately upon encountering the intended contrastive topic in the target cause, a violation should be detected. In a voice-mismatched structure, by contrast, the passive alternate does not lack a topic; rather, the intended topic transition is not supported by that topic structure (the 'wrong' argument is in subject/topic position). In such a case, the defective transition may not induce a violation until later in the clause, when the intended contrast becomes apparent.

Next, it is possible that the differences in the contrastive topic condition follow not from a penalty associated with mismatch, but instead from an advantage for the matched condition. Indeed, when we consider reading patterns in the target clause together with those from the antecedent, a possible interpretation of the result is that the

matched contrastive topic condition retains the advantage of a facilitative effect for raised-argument structures (as compared to argument-in-situ) in the antecedent clause. The question then arises of why the simple focus structures do not behave in the same way; that is, why would simple focus structures fail to carry over the advantage from the antecedent into the subject of the target clause? Some variation on the proposal made here, which implicates topic transitions between the antecedent and the target, might offer an explanation, but a better understanding of the nature of processing across clause boundaries is necessary to fully develop that account.

Finally, there is an additional possible interpretation of the pre-ellipsis result, one having nothing to do with topic structure, that must be considered. That explanation attributes the increased reading times in the mismatched contrastive topic to a garden-path effect whereby readers misanalyze clausal coordination as NP conjunction. The potential ambiguity crops up only in the mismatched contrastive topic conditions, where two semantically similar NPs are separated by a conjunction, leaving open the possibility of misanalysis as in 47.

(47) It's fairly easy to identify [venomous snakes, and poisonous plants]$_{NP}$ …

If readers are inclined to posit NP coordination in this way, and they do so in only this condition, this could offer an alternative explanation for the observed data pattern.[7]

The data in this case, however, are not consistent with a traditional garden-path analysis for three reasons. First, for all stimuli, a comma appeared at the close of the antecedent clause, unambiguously signaling that readers had reached the end of the clause. We can confirm that readers in fact were attentive to this cue, as reading times across all conditions were elevated at the final word, indicating clause-final 'wrap-up' (see Fig. 3). Next, if it were the case that readers did not attend to the presence of the comma, misanalyzing the target subject as a conjoined object, a reading-time increase indicating reanalysis would not be expected until later, when evidence of misanalysis becomes available. That evidence becomes available to the reader only at the auxiliary verb. The fact that increased reading times are observed at the target subject, prior to encountering evidence of misanalysis, argues against this alternative explanation. But there remains the possibility that the pattern seen here follows from a so-called 'hallucinated garden path'. Under this scenario, readers would correctly attend to the comma at the end of the antecedent clause, but upon encountering a semantically similar noun phrase in an immediately following subject position, a reanalysis of that earlier parse would be entertained. This type of disruption has been modeled as 'uncertainty over prior input' (Levy 2011) and could potentially explain the pattern of data seen here. Whether this effect reflects difficulty in resolving the intended discourse transition or uncertainty about prior input, or some combination of the two, cannot be determined at this point.

5. General discussion. Findings from the experiments reported in the previous section showed clear evidence of the influence of information structure on ellipsis acceptability, consistent with the claim made here that contrastive topics are subject to a parallelism constraint that does not apply to simple focus structures. It was shown that the information structure of the target clause in an ellipsis (whether focus fell on a subject/topic as opposed to an auxiliary verb) reliably distinguished between relatively acceptable and relatively unacceptable instances of mismatched ellipsis (experiments 1 and 3). The differences in acceptability observed for mismatched contrastive topic and

---

[7] Thanks to an anonymous referee for raising this point.

simple focus structures were furthermore shown to arise whether the target clause contained an ellipsis or not (experiments 2 and 3), though notably, the difference was smaller in cases with no ellipsis. Finally, an interaction between antecedent match and information structure consistent with the pattern observed in experiments 1 and 3 was observed in on-line measures, affecting reading times both before and after the ellipsis site in the target clause. These various findings offer positive evidence that the proposed constraint on contrastive topics plays a role in ellipsis acceptability and that the constraint applies independently of ellipsis.

An obvious question raised by these results is the role of the constraint in the ellipsis-licensing process, more specifically, whether an understanding of the constraint and its effects might serve to adjudicate between competing models of ellipsis licensing. The proposal developed here is in fact agnostic on whether ellipsis might occur as the result of deletion or instead be base-generated as a null constituent. In several respects, however, the current proposal is more consistent with a base-generated account. Indeed, the constraint on contrastive topics can be seen to supply the filter required by a base-generated model to curb overgeneration of unacceptable structures. Moreover, from the empirical data presented here it becomes clear that the types of examples typically used to argue in support of an LF-identity constraint—that is, those data that show a dramatic contrast between matched and unmatched ellipsis pairs—in fact suffer from the cumulative effects of perhaps multiple penalties. On the one hand, when the effect of information structure is isolated, the remaining penalty reflects a loss of acceptability that is much smaller than would be expected from the violation of a grammatical constraint licensing phrase structure. On the other hand, the source of that residual penalty remains unexplained, and it was shown clearly in experiments 1 and 3 that even for simple focus ellipses, matching argument structure across antecedent and target is preferred. That finding suggests that structure matching may yet be implicated in ellipsis licensing, though it might operate in a more limited fashion or apply in a different manner than traditional deletion-under-identity models have assumed.

5.1. OTHER 'DEFECTIVE' ANTECEDENTS. When assessing the relative contributions of the contrastive topic and LF-identity constraints in modeling the grammar of ellipsis, it is helpful to consider additional types of apparent licensing violations beyond the argument structure mismatches that have constituted the primary data set under consideration here. There are, of course, other types of 'defective' antecedents that have been described in the literature. These include, for example, syntactic category mismatches (nominal antecedents), split antecedent ellipses (interpreted via reference to more than one source clause), and so-called 'exophoric' or 'pragmatically controlled' ellipses (with no linguistic antecedent), all of which are predicted to be ungrammatical under traditional deletion-under-identity accounts. Structures of this sort nonetheless occur, and are judged by many speakers to be acceptable. Examples previously cited in the literature are offered in 48–50.

(48) This letter deserves a response, but before you do, …   (Kehler 2000:549, ex. 26)
(49) Wendy is eager to sail around the world and Bruce is eager to climb Kilimanjaro, but neither of them can because money is too tight.
(Webber 1978, ex. 63)
(50) Once in my room, I took the pills out. 'Should I?' I asked myself.
(Miller & Pullum 2012:10, ex. 25)

The question to ask is which of the constraints inducing parallelism—the constraint on contrastive topics or the constraint enforcing LF identity—is better equipped to address

data of these sorts. It turns out that data like these are particularly problematic for a dele-tion-based approach to ellipsis precisely because they fail to satisfy the LF-identity con-straint licensing deletion, at least in its traditional formulation. Proposals from Merchant (2008, 2013) and Fiengo and May (1994), as already described, have carved out an ex-ception for voice mismatches (and potentially other types of argument structure mis-match). But even modifications like these are not sufficient to admit the full range of exceptions to identity that seem to arise. By contrast, the contrastive topic constraint ex-cludes none of the cases in 48–50, since none of them instantiates a contrastive topic.

From the perspective of the current work, there are two possible routes in dealing with this broader class of mismatched ellipses. One possibility is to abandon a deletion-based approach to ellipsis, shifting entirely to a base-generation account while relying on the contrastive topic constraint to filter out bad mismatches. That approach follows the basic model set up in Kehler 2000, 2002, where it was argued that VPE is strictly (deep) anaphoric and that mismatch effects emerge as the result of reconstruction trig-gered by the need to establish parallel coherence. The innovation under the current ac-count is to replace the constraint on parallel coherence with a constraint on contrastive topics. The improvement on the coherence approach is the extension to nonelliptical targets and the treatment of structures with multiple coherence relations. The limitation to this approach, however, is that there remains a considerable risk of overgeneration, and without an account of the types of discourse constraints governing the various ad-ditional contexts—beyond contrastive topics—where ellipsis can occur, the account remains incomplete. Note, however, that various proposals have been described else-where in the literature, which suggest ways in which these excluded data sets might be accommodated by a more elaborated theory (Webber 1978, Elbourne 2008, Miller & Pullum 2012, inter alia).

As an alternative to a shift to a base-generated approach, a grammar model might in-stead retain the basic conception of VPE as the result of a deletion process, but admit a variety of constraints, particularized to different contexts, that trigger deletion. This would stand in contrast to the traditional approach that has invoked LF identity as the sole 'triggering' mechanism. For contrastive topic ellipses, which appear to conform to LF identity, it could be the case that a contrastive topic configuration, combined with some constraint on recoverability, serves to license deletion. For other types of dis-course configurations, however, it would be necessary to describe alternative con-straints licensing deletion. Here, crucially, the theoretical burden of explaining ellipsis licensing in all those remaining contexts is identical to the burden faced by the base-generated account. With respect to the current proposal then, the relevant point is not that the contrastive topic analysis precludes the existence of hidden structure in the el-lipsis or even a view of ellipsis as the output of deletion. Rather, an appreciation of an expanded set of data makes clear that the LF-identity constraint argued to trigger dele-tion is too strong, and the constraint on contrastive topics offers an alternative means of ruling out apparently ungrammatical ellipses.

5.2. Other types of anaphora. Another way of assessing the relative contributions of the contrastive topic constraint versus LF identity to the grammar of ellipsis is to con-sider other types of anaphora beyond just VPE. Anaphors are a mixed bag, and they can be subdivided in a variety of ways, based, for example, on the type of constituent in-volved (nominal, verbal, clausal), its syntactic form (clitic, pro-form, lexically headed phrase), and its lexical status (reduced form, repeated form, related form). It was Han-kamer and Sag (1976, and later Sag & Hankamer 1984) who suggested that pro-form ver-

sus ellipsis was a particularly salient distinction, a point that was argued in the context of a hypothesis about the role that LF representations play in not only licensing, but also interpreting, ellipsis. Adherence to or rejection of this distinction has served, over the years, as a bright divide between two approaches to anaphora: those that distinguish between 'surface' and 'deep' and those that are 'all deep'. But the surface/deep divide is not the only means of categorizing verbal anaphors, and the contrastive topic analysis predicts a natural class of anaphors that show features previously attributed to a constraint on LF representations (i.e. an intolerance for argument structure mismatch). The proposal does so, however, by constraining overt structure, without making reference to LF.

The natural class of anaphors that is predicted by the contrastive topic analysis of mismatch includes gapping, pseudo-gapping, and stripping (all of which occur in contrastive topic relations), together with the subset of VPE cases that form contrastive topics. Each of these anaphor types is expected to be sensitive to argument structure mismatches across antecedent and target clauses. This prediction overlaps in many respects with Hankamer and Sag's original (1976) categorization of surface anaphora (though they did not address pseudo-gapping) and to a certain degree with Merchant's (2008, 2013) proposal (which identifies pseudo-gapping as crucially distinct from VPE). But the current proposal differs from each of these in offering an explanation for why VPE seems to pattern in some cases like a surface anaphor, and in some cases like a pro-form. It also differs in admitting nonelliptical *do so* and *do it*, just in case these expressions are realized in a structure that does not form a contrastive topic. As such, the contrastive topic analysis offers a potential increase in empirical coverage (by including these nonelliptical cases), but a question once more arises regarding the role of the contrastive topic constraint: does it apply in addition to the constraint on LF identity or instead of it?

One way of answering this question is to consider whether LF identity does any extra work beyond what could be achieved with the contrastive topic constraint alone. At first blush, a consideration of sluicing structures suggests that it does. Sluicing structures, as demonstrated in 51–52, are characterized by a question fragment (either matrix or embedded) that contains a wh-word.

(51) The pudding was eaten by someone, but we don't know who/by who.
(52) Someone ate the pudding, but we don't know who/*by who.

Under an ellipsis-based approach to sluicing, such structures differ minimally from VPE in that where a VPE deletes a VP constituent, a sluice deletes an IP. On the assumption that deletion is triggered by LF identity, sluices are predicted to show matching effects. Given that the entire IP target is deleted, under this approach, antecedent/target matching might appear to be a moot point, but evidence of a matching constraint can be seen in certain contexts, as, for example, in 52 above, where case marking the wh-word in the remnant (*who*) that corresponds to a subject argument in the antecedent (*someone*) results in a violation. That result is predicted if sluicing is licensed by LF identity. However, insofar as sluices do not form contrastive topics (and it appears they do not), the result is not predicted under the contrastive topic analysis of mismatch. Under the current account, the penalty associated with mismatched sluices must be attributed to an alternative source.

A closer look at sluicing, however, reveals substantial differences between the two structures, differences that might suggest that they warrant independent treatments. For example, where the target clause in an ellipsis forges a link back to its antecedent via a focal contrast, in a sluice, a focused wh-phrase supplies a link back to an underspeci-

fied (or even implicit) argument in a prior utterance. Moreover, as alluded to already, while mismatch effects are evidenced in VPE by voice marking on the auxiliary verb governing the ellipsis site, they are evidenced in sluicing by case matching between arguments in the antecedent and the 'remnant' of a sluice. Case-matching effects of this sort are known to arise outside of sluicing contexts, however—for example, in question/answer pairs. While it is possible to extend an ellipsis analysis to question/answer contexts (see e.g. Merchant 2004), alternative models of question/answer matching have been proposed (Groenendijk & Stokhof 1984, Ginzburg & Sag 2000), and it is possible that those accounts might instead be extended to address sluicing. Moreover, case matching for question/answer pairs has also been described in the psycholinguistics literature, where a very different characterization of the phenomenon is on offer. Levelt and Kelter (1982), for example, demonstrated modulation of the effect based on the amount of material intervening between question and answer, in support of a memory-based model.

Before leaving this discussion of the typology of anaphors, I note that the views adopted here overlap in many ways with proposals from Winkler (1997, 2005), who likewise marks an important role for information structure in characterizing verbal anaphora. Under Winkler's view, verbal anaphors are divided into sentence-bound and discourse-bound (cf. Williams 1977), and the role focus plays in each case is different. Given Winkler's categorization, gapping and ellipsis do not form a natural class, since gapping is sentence-bound, and VPE is discourse-bound (2006). Winkler does not address the mismatch question directly, noting instead that mismatch is possible under VPE (given its status as discourse-bound). Nonetheless, one way of reconciling the conclusions reached in this work with Winkler's would be to argue that VPE can be either sentence-bound (in just those cases where they form contrastive topic relations) or discourse-bound (in the remaining cases).

5.3. The residual mismatch effect. Summarizing to this point, I have argued that the contrastive topic analysis of ellipsis mismatch is itself agnostic on the question of whether there is hidden structure in an ellipsis site. Indeed, the issues raised by the mismatch problem do not follow directly from the assumption of hidden structure, but instead arise from the adoption of an identity constraint that applies to that hidden structure. As a constraint on ellipsis licensing, LF identity is simply too strong, and the contrastive topic analysis offers a potential alternative characterization of the violation. The obvious next question is whether the contrastive topic constraint is too weak. The empirical results reported here suggest yes, in two respects. First, when comparing mismatched elliptical and nonelliptical structures, the ellipses consistently showed an added penalty above and beyond the penalty induced by a mismatched contrastive topic (experiments 2 and 3). Next, even in cases of simple focus (noncontrastive topic) structures, matching ellipses were preferred over mismatches (experiments 1 and 2). This latter result is consistent, moreover, with the Mauner et al. 1995 finding that a matching preference for ellipsis persists even when a short passive is used (in contrast with *do it*, which requires a long passive to induce a mismatch penalty).

Because effects like these arise specifically in the context of ellipsis, it might seem reasonable to fall back on an ellipsis-specific constraint, like deletion under identity, to explain them. I have identified, however, a variety of problems that follow from invoking the LF-identity constraint on ellipsis, at least as it has been traditionally formulated: first, the constraint does not apply to all ellipses; next, in those cases where it does apply, effects are gradient (arguing against a categorical constraint); and finally, the ef-

fect emerges only weakly in the current data; meanwhile, in a host of examples culled from corpora there is apparently no violation at all. Each of these observations argues against invoking LF identity to shore up the empirical predictions of the contrastive topic analysis.

From the perspective of the current work, rather than pursing an LF-based explanation for residual mismatch effects, a more promising alternative might be instead to pursue a more elaborated model of discourse factors that constrain the use of ellipsis. The goal of the current work has been to consider the value of a single constraint—the constraint on contrastive topics—for assessing a difference between two broad classes of data. It remains possible that there are additional constraints that have not been identified here, but that nonetheless are implicated in licensing ellipsis, and that might underlie distinctions of an even finer grade. Any post hoc proposal suggested to account for the current data would of course be speculative, but past work on the role that discourse context plays in licensing ellipsis offers some examples of the types of constraints that might be at play. Miller (2011), for example, counters the default assumption that the different classes of verbal anaphors are interchangeable (tacit in the many works comparing VPE and *do it*) by demonstrating with corpus data subtle differences in contextual support for anaphors of different types. Similarly, Malt (1985) showed that VPE, specifically, is processed more easily when the ellipsis is contained within an utterance that supplies an answer to a question, as opposed to following up on a prior statement. In a related vein, Grant and colleagues (2012) have argued that the salience of an implicit question in a discourse context can ease processing for ellipsis, in some cases weakening the mismatch penalty. Each of these findings suggests added constraints on ellipsis licensing that can modulate acceptability and ease of processing, and that are crucially discourse-based. (See also Frazier & Clifton 2005 and Garnham & Oakhill 1987 for effects of discourse on interpretation of ambiguous ellipsis.)

In addition to such discourse-based considerations, recent work has highlighted a role for extragrammatical processing constraints in explaining gradient patterns of acceptability for ellipsis. This includes, notably, Arregui et al. 2006 and Kim et al. 2011, both of which specifically address the mismatch problem. While I have argued that those proposals are not equipped to account for the information-structural differences at issue in the current work, the memory and parsing-based effects that are predicted under those accounts would be expected to induce additional penalties, offering a potential explanation for the residual effects seen here. A further consideration raised directly by the current work involves the ways in which information packaging influences ellipsis processing in real time. The analysis developed here has been framed in terms of a static constraint on well-formed structures, with some tentative discussion of how that constraint might be applied during on-line processing. For example, reading-time data from experiment 4 were argued to be consistent with the proposal that a well-formed topic transition is necessary for successful interpretation of an ellipsis further downstream, on the view that the topic structure is an important cue to the identification of contrasting arguments. Presumably, however, the role that information structure can play in this respect might vary depending on context.

A contrast noted by an anonymous referee illustrates this point. Consider the pair below, in which 53 seems to be less acceptable than 54. The former is precisely the type of mismatched structure that is filtered out by the contrastive topic constraint (though here we see active-passive, not passive-active order), and its reduced acceptability is predicted. The sentence in 54 does not exhibit a voice mismatch but does share with 53 the property that *Larry* and *Bill* are intended to be contrastive, despite appearing in par-

allel argument positions. The crucial difference between these cases is that *Larry* is realized as a subject in 53 but as a displaced object in 54.

(53) #Pia likes Bill, and Larry is too. [liked by Pia]
(54) Pia likes Bill, and Larry, she does too. [like *t*]

This leads to an important difference not only in the missing material to be inferred at the ellipsis site (a VP containing a passive verb plus oblique agent in 53 vs. an active verb plus direct object trace in 54), but also the number of arguments encountered before the ellipsis. In the displaced-object case, prior to encountering the ellipsis, the reader has encountered two sets of referents: *Pia* and *Bill*, and *Larry* and *Pia* (evoked by the pronoun *she*), and presumably the work of aligning these arguments—that is, recognizing that *Bill* and *Larry* form a set—is done. In 53, by contrast, at the ellipsis site, the reader has encountered a pair of arguments (*Pia* and *Bill*) plus one more (*Larry*). It is only at the ellipsis site that it becomes apparent that *Larry* forges a link back to *Bill*, and the structure is judged to be ill-formed. Of course, alternative explanations for the contrast in 53 and 54 could be offered, and the proposal just sketched, once tested, could be found wanting. The intended point of this digression is to suggest that the role that information structure plays in parceling out information during real-time production and comprehension might serve as yet another source of variability in the acceptability and ease of processing for elliptical structures.

6. CONCLUSION. The research presented here proposed a novel account of the ellipsis mismatch problem, demonstrating that relative acceptability and ease of processing for mismatched ellipses can be reliably predicted based on information structure. The linguistic analysis posited a well-formedness constraint that predicts penalties associated with mismatched contrastive topic relations both in elliptical and nonelliptical structures. Empirical evidence from three off-line studies and one on-line study supporting this proposal were presented. Those empirical findings confirmed that the worst cases of antecedent mismatch showed effects of an information-structural violation that is not unique to ellipsis. Once the information-structural effect is accounted for, the remaining penalty turns out to be much smaller than previously assumed. These results raise the question of whether a categorical licensing constraint offers the best characterization of ellipsis in a grammar model, while highlighting the importance of discourse factors for both off-line acceptability and on-line processing of elliptical structures.

REFERENCES

ARREGUI, ANA; CHARLES CLIFTON, JR.; LYN FRAZIER; and KEIR MOULTON. 2006. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language* 55.232–46.
BOLINGER, DWIGHT L. 1961. Contrastive accent and contrastive stress. *Language* 37.83–96.
BÜRING, DANIEL. 2003. On D-trees, beans, and B-accents. *Linguistics and Philosophy* 26. 511–45.
CHAO, WYNN. 1987. *On ellipsis*. Amherst: University of Massachusetts, Amherst dissertation.
CHOMSKY, NOAM, and HOWARD LASNIK. 1993. The theory of principles and parameters. *Syntax: An international handbook of contemporary research*, ed. by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 506–69. Berlin: Mouton de Gruyter.
COMRIE, BERNARD, and STEPHEN MATTHEWS. 1990. Prolegomena to a typology of tough movement. *Studies in typology and diachrony: For Joseph H. Greenberg*, ed. by William Croft, Keith Denning, and Suzanne Kemmer, 43–58. Amsterdam: John Benjamins.
DALRYMPLE, MARY. 1991. Against reconstruction in ellipsis. Technical report SSL-91-114. Palo Alto: Xerox PARC.

DALRYMPLE, MARY. 2005. Against reconstruction in ellipsis. *Ellipsis and nonsentential speech* (Studies in linguistics and philosophy 81), ed. by Reinaldo Elugardo and Robert J. Stainton, 31–55. Berlin: Springer.

DALRYMPLE, MARY; STUART M. SHIEBER; and FERNANDO C. N. PEREIRA. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy* 14.399–452.

DANEŠ, FRANTISEK. 1974. Functional sentence perspective and the organization of the text. *Papers on functional sentence perspective*, ed. by Frantisek Daneš, 106–28. The Hague: Mouton.

DUFFIELD, NIGEL, and AYUMI MATSUO. 2003. Factoring out the parallelism effect in VP-ellipsis. *Chicago Linguistic Society* 39.591–603.

ELBOURNE, PAUL. 2008. Ellipsis sites as definite descriptions. *Linguistic Inquiry* 39.191–220.

ERTESCHIK-SHIR, NOMI. 2007. *Information structure: The syntax-discourse interface*. Oxford: Oxford University Press.

FIENGO, ROBERT, and ROBERT MAY. 1994. *Indices and identity*. Cambridge, MA: MIT Press.

FRAZIER, LYN, and CHARLES CLIFTON, JR. 2001. Parsing coordinates and ellipsis: Copy alpha. *Syntax* 4.1–22.

FRAZIER, LYN, and CHARLES CLIFTON, JR. 2005. The syntax-discourse divide: Processing ellipsis. *Syntax* 8.121–74.

FRAZIER, LYN, and CHARLES CLIFTON, JR. 2006. Ellipsis and discourse coherence. *Linguistics and Philosophy* 29.315–46.

FRAZIER, LYN; ALAN MUNN; and CHARLES CLIFTON, JR. 2000. Processing coordinate structures. *Journal of Psycholinguistic Research* 29.343–70.

FRAZIER, LYN; LORI TAFT; TOM ROEPER; CHARLES CLIFTON, JR.; and KATE EHRLICH. 1984. Parallel structure: A source of facilitation in sentence comprehension. *Memory and Cognition* 12.421–30.

GARNHAM, ALAN, and JANE OAKHILL. 1987. Interpreting elliptical verb phrases. *Quarterly Journal of Experimental Psychology* 39A.611–27.

GARVEY, CATHERINE; ALFONSO CARAMAZZA; and JACK YATES. 1975. Factors influencing assignment of pronoun antecedents. *Cognition* 3.227–43.

GINZBURG, JONATHAN, and IVAN A. SAG. 2000. *Interrogative investigations*. Stanford, CA: CSLI Publications.

GRANT, MARGARET; CHARLES CLIFTON, JR.; and LYN FRAZIER. 2012. The role of non-actuality implicatures in processing elided constituents. *Journal of Memory and Language* 66.326–43.

GRINDER, JOHN, and PAUL M. POSTAL. 1971. Missing antecedents. *Linguistic Inquiry* 2.269–312.

GROENENDIJK, JEROEN ANTONIUS GERARDUS, and MARTIN JOHAN BASTIAAN STOKHOF. 1984. *Studies on the semantics of questions and the pragmatics of answers*. Amsterdam: University of Amsterdam dissertation.

HANKAMER, JORGE. 1978. On the nontransformational derivation of some null VP anaphors. *Linguistic Inquiry* 9.66–74.

HANKAMER, JORGE, and IVAN SAG. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7.391–426.

HARDT, DANIEL. 1993. *Verb phrase ellipsis: Form, meaning, and processing*. Philadelphia: University of Pennsylvania dissertation.

HENDRIKS, PETRA. 2004. Coherence relations, ellipsis and contrastive topics. *Journal of Semantics* 21.133–53.

JACKENDOFF, RAY. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

KADMON, NIRIT. 2001. *Formal pragmatics*. Oxford: Blackwell.

KEHLER, ANDREW. 2000. Coherence and the resolution of ellipsis. *Linguistics and Philosophy* 23.533–75.

KEHLER, ANDREW. 2002. *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.

KEHLER, ANDREW; LAURA KERTZ; HANNAH ROHDE; and JEFFREY L. ELMAN. 2008. Coherence and coreference revisited. *Journal of Semantics* 25.1–44.

KENNEDY, CHRIS. 2003. Ellipsis and syntactic representation. In Schwabe & Winkler, 29–53.

KERTZ, LAURA. 2008. Focus structure and acceptability in verb phrase ellipsis. *West Coast Conference on Formal Linguistics (WCCFL)* 27.283–91.

KERTZ, LAURA. 2010. *Ellipsis reconsidered*. San Diego: University of California, San Diego dissertation.

KIM, CHRISTINA; GREGORY M. KOBELE; JEFFREY T. RUNNER; and JOHN T. HALE. 2011. The acceptability cline in VP ellipsis. *Syntax* 4.318–54.

KIM, CHRISTINA, and JEFFREY T. RUNNER. 2009. Discourse structure and parallelism in VP ellipsis. Poster presented at Semantics and Linguistic Theory (SALT) 19, The Ohio State University, Columbus.

KRIFKA, MANFRED. 1991. A compositional semantics for multiple focus constructions. *Proceedings of Semantics and Linguistic Theory (SALT) (Cornell Working Papers in Linguistics 11)* 1.127–58. Online: http://elanguage.net/journals/salt/article/view/1.127.

KRIFKA, MANFRED. 1999. Additive particles under stress. *Proceedings of Semantics and Linguistic Theory (SALT)* 8.111–28. Online: http://elanguage.net/journals/salt/article/view/8.111.

KRIFKA, MANFRED. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55.3–4.243–76.

LADD, D. ROBERT. 1980. *The structure of intonational meaning: Evidence from English*. Bloomington: Indiana University Press.

LAKOFF, GEOFFREY. 1971. Presupposition and relative well-formedness. *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology*, ed. by Danny D. Steinberg and Leon A. Jakobovits, 329–40. Cambridge: Cambridge University Press.

LEVELT, WILLEM J. M., and STEPHANIE KELTER. 1982. Surface form and memory in question answering. *Cognitive Psychology* 14.78–106.

LEVY, ROGER. 2011. Integrating surprisal and uncertain-input models in online sentence comprehension: Formal techniques and empirical results. *Proceedings of the 49th annual meeting of the Association for Computational Linguistics*, 1055–65.

LOBECK, ANNE. 1995. *Ellipsis: Functional heads, licensing, and identification*. Oxford: Oxford University Press.

MALT, BARBARA C. 1985. The role of discourse structure in understanding anaphora. *Journal of Memory and Language* 24.271–89.

MAUNER, GAIL; MICHAEL K. TANENHAUS; and GREG N. CARLSON. 1995. A note on parallelism effects in processing deep and surface verb-phrase anaphora. *Language and Cognitive Processes* 10.1–12.

MCNALLY, LOUISE. 1998. On recent formal analyses of topic. *The Tbilisi Symposium on Language, Logic, and Computation: Selected papers*, ed. by Jonathan Ginzburg, Zurab Khasidashvili, Carl Vogel, Jean-Jacques Lévy, and Enric Vallduví, 147–60. Stanford, CA: CSLI Publications.

MERCHANT, JASON. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

MERCHANT, JASON. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27.661–738.

MERCHANT, JASON. 2008. An asymmetry in voice mismatches in VP-ellipsis and pseudo-gapping. *Linguistic Inquiry* 39.169–79.

MERCHANT, JASON. 2010. Voice mismatches and the dark side of ellipsis. Chicago: University of Chicago, MS.

MERCHANT, JASON. 2013. Voice and ellipsis. *Linguistic Inquiry* 44.77–108.

MILLER, PHILIP. 2011. The choice between verbal anaphors in discourse. *Anaphora processing and applications: 8th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 2011)* (Lecture notes in artificial intelligence 7099), ed. by Iris Hendrickx, Sobha Lalitha Devi, António Branco, and Ruslan Mitkov, 82–95. Berlin: Springer.

MILLER, PHILIP, and GEOFFREY K. PULLUM. 2012. Exophoric VP post-auxiliary ellipsis. Paris: Université Paris Diderot, MS. Online: http://www.lel.ed.ac.uk/~gpullum/Exophoric.pdf.

MURPHY, GREGORY L. 1985. Processes of understanding anaphora. *Journal of Memory and Language* 24.290–303.

REINHART, TANYA. 1982. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27.53–94.

ROBERTS, CRAIGE. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. *OSU Working Papers in Linguistics (Papers in semantics)* 49.91–136.

ROOTH, MATS. 1992. A theory of focus interpretation. *Natural Language Semantics* 1.75–116.

ROOTH, MATS. 1993. Ellipsis redundancy and reduction redundancy. *Proceedings of the Stuttgart Ellipsis Workshop* (Arbeitspapiere des Sonderforschungsbereichs 340.29), ed. by S. Berman and Arild Hestvik. Stuttgart and Tübingen: Universities of Stuttgart and Tübingen.

SAG, IVAN. 1976. *Deletion and logical form*. Cambridge, MA: MIT dissertation.

SAG, IVAN, and JORGE HANKAMER. 1984. Toward a theory of anaphoric processing. *Linguistics and Philosophy* 7.325–45.

SCHACHTER, PAUL. 1977. Does she or doesn't she? *Linguistic Inquiry* 8.763–67.

SCHWABE, KERSTIN, and SUSANNE WINKLER. 2003. Exploring the interfaces from the perspective of omitted structures. *The interfaces: Deriving and interpreting omitted structures*, ed. by Kerstin Schwabe and Susanne Winkler, 1–28. Amsterdam: John Benjamins.

SCHWARZSCHILD, ROGER. 1999. GIVENness, AvoidF, and other constraints on the placement of accent. *Natural Language Semantics* 7.141–77.

SIEWIERSKA, ANNA. 1984. *The passive: A comparative linguistic analysis*. London: Croom Helm.

STEEDMAN, MARK. 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry* 31.649–89.

STURT, PATRICK; FRANK KELLER; and AMIT DUBEY. 2010. Syntactic priming in comprehension: Parallelism effects with and without coordination. *Journal of Memory and Language* 62.333–51.

TANCREDI, CHRISTOPHER D. 1992. *Deletion, deaccenting, and presupposition*. Cambridge, MA: MIT dissertation.

TANENHAUS, MICHAEL K., and GREG N. CARLSON. 1990. Comprehension of deep and surface verb phrase anaphors. *Language and Cognitive Processes* 5.257–80.

VALLDUVÍ, ENRIC. 1992. *The information component*. New York: Garland.

WAGNER, MICHAEL. 2013. A compositional analysis of contrastive topics. Montreal: McGill University, MS. Online: http://semanticsarchive.net/Archive/mIwZTIwM/wagner11.pdf.

WASOW, TOM. 1972. *Anaphoric relations in English*. Cambridge, MA: MIT dissertation.

WEBBER, BONNIE LYNN. 1978. *A formal approach to discourse anaphora*. Cambridge, MA: Harvard University dissertation.

WILLIAMS, EDWIN S. 1977. Discourse and logical form. *Linguistic Inquiry* 8.101–39.

WILLIAMS, EDWIN S. 1997. Blocking and anaphora. *Linguistic Inquiry* 28.577–628.

WINKLER, SUSANNE. 1997. *Ellipsis and information structure in English and German: The phonological reduction hypothesis*. (Arbeitspapiere des Sonderforschungsbereichs 340.121.) Stuttgart and Tübingen: Universities of Stuttgart and Tübingen.

WINKLER, SUSANNE. 2005. *Ellipsis and focus in generative grammar*. Berlin: Mouton de Gruyter.

WINKLER, SUSANNE. 2006. Ellipsis. *The encyclopedia of language and linguistics*, 2nd edn., ed. by Keith Brown, 109–13. Oxford: Elsevier.

Department of Cognitive, Linguistic,   [Received 24 August 2010;
  and Psychological Sciences   revision invited 17 May 2011;
Brown University   revision received 16 April 2012;
Providence, RI 02912-1821   accepted 25 January 2013]
[laura_kertz@brown.edu]