Verb Phrase Ellipsis Resolution as a Side Effect of Discourse Coherence

Adam C. Baker
University of Chicago
adamc@uchicago.edu

May 30, 2007

1 Ellipsis, Economy and Discourse

> "One might expect that [ellipsis] would increase the interpretation burden on the hearer, and thus explicitness would always be preferred. In reality, the felicitous use of ellipsis appears to have the opposite effect." –Kehler (2002)

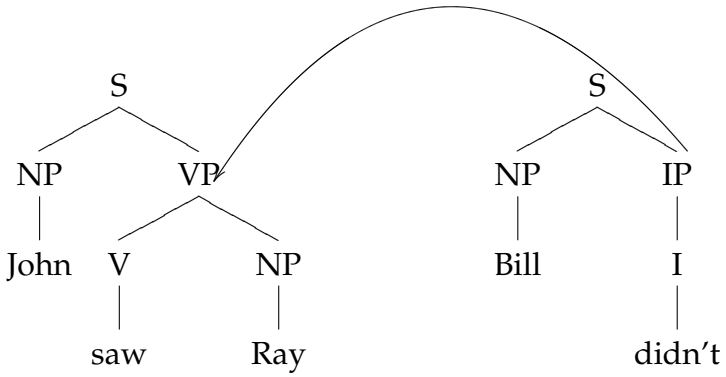
A great deal of recent research has focused on explaining patterns in natural language in terms of economy principles. Research within the Minimalist Program (Chomsky, 1995) frequently attempts to explain patterns using derivational economy. Fox (2000), for instance, uses a derivational principle of scope economy to explain certain scope possibilities of quantifiers under ellipsis. But Kehler's quote above suggests a potentially different notion of economy is also at work in ellipsis: computational economy, a sort of least-effort principle.

Ellipsis clearly requires less effort for the speaker by allowing them to say less. Laboratory evidence suggests that it also requires less effort for the hearer. Frazier and Clifton (2001) shows that processing difficulty of elided material does not depend on the size of the understood meaning of the ellipsis site. We might interpret this as showing that ellipsis also helps the hearer by allowing them to use something that they have computed from a previous utterance to fill in the elided material. One contentious debate in the ellipsis literature is whether this "something" that is reused from a previous utterance is a syntactic structure (Sag, 1976; Williams, 1977; Fiengo and May, 1994; Merchant, 2004b,

to appear, 2007) or a meaning (Dalrymple et al., 1991; Hardt, 1993, 1999; Culicover and Jackendoff, 2005). At first blush the Frazier and Clifton findings seem to support the semantic view, where only a meaning from a previous utterance is reused. If there is syntax in the the ellipsis site, it's not clear how the silent syntactic structure is generated without requiring effort that Increases with the size of the elided constituent.

Frazier and Clifton interpret their findings as showing that VPE makes use of a constant time syntactic copy operation. One way to formalize this proposal is through structure sharing: the IP node that dominates the stranded auxiliary remnant of VPE also dominates the VP of a previous utterance. This sort of analysis is roughly illustrated in (1).

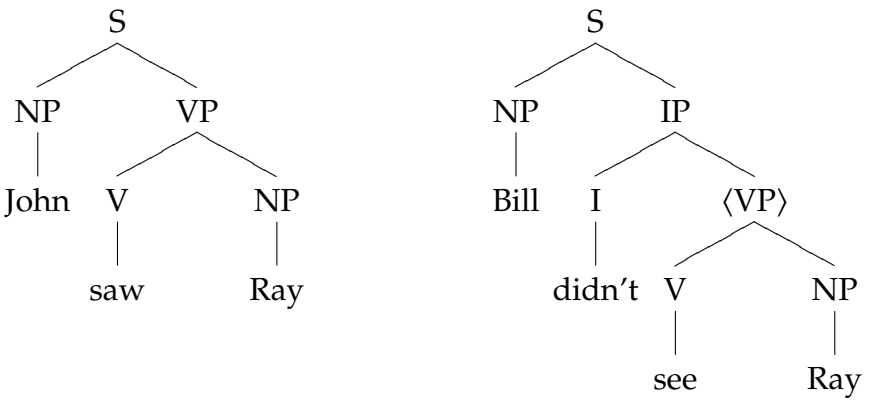
(1) a. John saw Ray.
b. Bill didn't.
c.



This "copy by reference" operation can copy an antecedent syntactic structure into the ellipsis site requiring effort that increases as the complexity of the copied structure increases.

Typically parse trees for sentences with VPE are written as in (2), where the angle brackets around the VP node indicate that everything it dominates is unpronounced. But unless this is just another notation for the multidominance graph in (1c), I know of no way the material under the VP can be generated without requiring effort proportional to the size of the elided constituent.

(2)



A second line of research has shown deep connections between ellipsis and discourse grammar. Williams (1977) argues that VPE involves a null anaphor that is filled in by the logical form of an antecedent VP by a discourse level grammar rather than a sentence level grammar. Tancredi (1992) and Rooth (1992) demonstrate a number of striking similarities between deaccented and elided VPs and work out analyses where similar discourse constraints license both phenomena. And Hardt (2003) shows that a clause with VPE must stand in some discourse relation to its antecedent clause.

Prüst, Scha, and van den Berg (1994) (henceforth PSV) advance a hypothesis connecting these two lines of research: VPE is resolved as a side effect of discourse coherence establishment. They represent an elided VP as a unification variable in the logical form. The variable gets bound by a unification grammar that is motivated independently for calculating discourse coherence. This bound variable ends up referring to the logical form from the previous utterance, resulting in the sort of structure sharing or multidominance graph represented in (1c), and not a structure like (2).

PSV's account faces a number of challenges. Their treatment of VPE only looks at the simple core cases like (1), where the elided and antecedent VPs are straightforwardly identical both structurally and semantically. Work on VPE has revealed that the antecedent possibilities for VPE are far more complicated, and it is these more complex cases that are the focus of current theoretical research on ellipsis. The goal of this paper is to revise PSV's framework to account for these cases.

One problem is the interaction of voice alternations and VPE. PSV assigns active and passive sentences different logical forms which do not unify in their grammar, predicting that VPE is impossible when an antecedent clause is active and the ellipsis clause is passive. This is contrary to fact, as we can see in (3) and (4). The importance of these voice mismatches to a theory of ellipsis is a major focus of current research (eg Merchant, 2001,

to appear, 2007; Frazier and Clifton, 2006; Kehler, 2002; Arregui et al., 2006).

(3) Active antecedent, passive ellipsis.
   a. Steve asked me to send the set by courier through my company insured, and it was. [sent by courier through my company insured](Kehler, 2002, p 53)
   b. Actually, I have implemented it [a computer system] with a manager, but it doesnt have to be. [implemented with a manager] (Kehler, 2002, 53)
   c. Speaker A: Someone mugged Tom yesterday.
   Speaker B: Oh yeah?
   Speaker C: You know, the same thing happened to Mary.
   Speaker B: Wow!
   Speaker A: You know, now that I think of it, Sandy was, too. [mugged]
   (Sag, 1976, 75, footnote 2)

(4) Passive antecedent, active ellipsis.
   a. A lot of this material can be presented in a fairly informal and accessible fashion, and often I do. [present this material in a fairly informal and accessible fashion] (Chomsky, quoted in Dalrymple et al., 1991, 440)
   b. This problem was to have been looked into, but obviously nobody did. [look into this problem](Kehler, 2002, 53)
   c. The ice cream should be taken out of the freezer, if you can. [take the ice cream out of the freezer](Hardt, 1993, 37)

Similarly, we might expect the causative/inchoative alternation to produce different logical forms for the VP (or *v*P, if that is the relevant level for ellipsis), and so on PSV's analysis, a mismatch of this type between an elided VP and its antecedent should be unacceptable as well. And in fact, most researchers who have looked at the problem come to this conclusion. For example, Merchant (2007, 24) presents the contrasts between (5) and (6) as evidence of this.

(5) Mismatch, no ellipsis
   a. This can freeze. Please freeze this.
   b. Bill melted the copper vase, and the magnesium vase melted, too.

(6) Mismatch with ellipsis
   a. * This can freeze. Please do. [freeze this]
   b. * Bill melted the copper vase, and the magnesium vase did, too. [melt]

However, altering the discourse environment can improve these sorts of examples, at least for some speakers.

(7) a. ? The fire burned my neighbor's house down just like my house had burned down.

b. ? We were able to sink the enemy fleet because many of their ships already had sunk in a previous battle.

(8) a. ? The fire burned my neighbor's house down just like my house had. [burned down]

b. ? We were able to sink the enemy fleet because many of their ships already had in a previous battle. [sunk]

I find the (8) examples as bad or perhaps only slightly worse than their unelided (7) counterparts, and I've confirmed these judgments with some of my colleagues. Other colleagues find (8) robustly unacceptable. The degraded status of these examples and the variance in acceptability among speakers must be explained, and the PSV model of VPE resolution, along with most models of VPE fail to do this.

Webber (1978) introduces another challenging set of data involving "split antecedents", where an elided VP draws it's understood meaning from more than one clause in the discourse. Two examples of this are given in (9) from Schwarz (2000). Parallel examples are discussed at length in Hardt (1999), Merchant (2004a), and Elbourne (to appear).

(9) a. When John had to cook, he didn't want to. When he had to clean, he didn't either.

b. Bob wants to sail round the world and Alice wants to climb Kilimanjaro, and they both will, if they can afford to.

(9a) is interesting because the second instance of VPE, *he didn't either*, is understood to mean that John didn't want to clean either. The understood meaning is constructed from *want to* from the first sentence and *clean* from the second sentence. Similarly, the VPE clauses in (9b) must contain both *sail round the world* and *climb Kilimanjaro*. The challenge here is not only to ensure that the elided VPs contain both *sail round the world* and *climb Kilimanjaro* and that both instances of *they* contains both *Bob* and *Alice*, but also to ensure that each elided VP is correctly matched to its corresponding subject so that we don't end up with a reading consistent with Alice doing both activities and Bob doing neither.

Another sort of shortcoming is the PSV account of discourse coherence. Their unification grammar only calculates coherence when the discourse is coherent in virtue of the

propositions expressed in the discourse being about parallel entities in parallel situations. But there are many other ways in which a discourse may be coherent. For example, a discourse may describe a sequence of events, one leading to the next, or the situation described by one proposition may have a causal relation to that of another. A more complete characterization of, for instance, cause-effect coherence may be useful for explaining the possibility of ellipsis in other cases where the logical form of the antecedent may not match that of the elided VP, such as (10) and (11), from Hardt (1993).

(10) Martha and Irv had planned to nominate each other, but Martha couldn't, because of political obligations. [nominate Irv]

(11) Harry used to be a great speaker, but he can't anymore, because he lost his voice. [speak]

An apparent problem with their discourse grammar is that it uses a context free phrase structure grammar to parse the discourse. It has been demonstrated that coherence can be established between clauses in a non-context free way (Wiebe, 1993; Forbes et al., 2003). But Webber et al. (2003) factor discourse relations into two types: (1) anaphoric relations, expressed by discourse adverbials such as *then*, *otherwise*, *nevertheless*, and (2) structural relations, expressed by discourse adjacency, coordinating conjunctions such as *and*, *or*, *but*, and *so* or subordinating conjunctions such as *instead*, *although*, *whereas*, and *when*. They show that, although anaphoric relations can connect clauses fairly freely, structural discourse relations can only connect clauses in a context-free way. Therefore, a context-free phrase structure grammar like the one used by PSV may be adequate for the structural fragment of the discourse grammar. For this reason, this paper will look only at VPE examples where the elided VP and its antecedent stand in a structural discourse relation, leaving an analysis of VPE with anaphoric discourse relations connecting the elided VP to its antecedent for future work.

This paper attempts to address some of these shortcomings by incorporating an analysis of Cause-Effect coherence establishment and the resolution of pronoun reference advanced by Hobbs (1979) and Kehler (2002, chapter 6). Like PSV, the Hobbs/Kehler analysis uses a unification grammar, but discourse connected clauses are not unified with each other. Instead, they are unified with a third formula from the participants knowledge base that states an implication connecting the two clauses. I will generalize this mechanism account for VPE, and then explore the consequences of extending this model of unification through an "implicational bridge" to account for the voice and argument structure mismatches in (3) – (8).

Given the unresolved issues with the PSV account of VPE I've just pointed out, we

may be asking why we should revise the PSV model rather than using some other model. One reason to use a refined PSV model is its conceptual tractability. PSV's context-free grammar for parsing discourses is simpler and more familiar than alternatives such as D-LTAGS of Forbes et al. (2003) or Segmented Discourse Representation Structures of Asher (1993). Using this simpler model will allow me to make precise statements that cover a smaller range of data, and will hopefully reveal insights that future work can incorporate into a more complex framework with better empirical coverage.

The most compelling reason to stick with the PSV model is that it offers a single mechanism connecting the apparently unconnected discoveries about ellipsis regarding processing economy and discourse. The following two assumptions are sufficient to establish this connection.

1. The inferences that establish discourse coherence between clauses can be modeled using a unification algebra with an additional binary operator, the Most Significant Common Denominator.

2. Unstressed pronouns and VP ellipsis are fundamentally similar. They are both represented as unification variables at logical form.

Because unification entails "copy by reference", assigning a value to a unification variable is a constant time operation with respect to the complexity or size of the value assigned, consistent with the Frazier and Clifton (2001) findings. These two assumptions also predict Hardt (2003)'s observation that a discourse relation must connect the VPE clause to its antecedent and the similarity between VPE and deaccenting discussed in Rooth (1992) and Tancredi (1992) and much subsequent literature.

Hardt (1993, 1999) and Elbourne (to appear) have both claimed that pronouns are basically the same phenomena using radically different systems. Hardt treats pronoun resolution as a semantic phenomena and claims that elided VPs are pro-verbs; Elbourne treats VPE as a syntactic phenomena, and claims that pronouns are cases of NP ellipsis where the pronoun is a stranded determiner. The important insight is that both pronoun reference and VPE should be modeled using the same mechanism, supporting the second assumption above.

The second assumption is challenged by Hankamer and Sag (1976)'s distinction between surface anophors such as VPE and deep anaphors such as pronouns. This distinction is based on the fact that, for instance, pronouns, but not elided VPs, may take antecedents from the non-linguistic environment of the utterance. This distinction can be modeled in the present system by allowing additional resolution mechanisms for resolving deep anaphors that are unavailable for surface anaphors. However, providing

such mechanisms, along with principled reasons these mechanisms should be available for pronouns but not VPE, is beyond the scope of this paper[1].

To address the challenges presented above, I propose that discourse coherence establishment may proceed through an "implicational bridge". I argue that:

1. Implicational bridge inferences are crucial to coherence establishment. Through coherence, they become crucial to ellipsis resolution.
2. The use of an implicational bridge is costly. They are used only when needed to establish coherence. An utterance that forces the use of an implicational bridge may cause degraded acceptability, varying with the salience of the bridging proposition.

Hardt (2005) makes both of these claims in some form, but this investigation will provide a new perspective on why they hold.

The second point is very similar to the recycling hypothesis of Arregui et al. (2006). They posit a recycling operation that can repair syntactic mismatches between an elided VP and its antecedent that minimally alters the antecedent structure to allow it to fit into the syntactic frame required by the ellipsis site. My proposal implements this hypothesis using proof-theoretic implicational reasoning over logical forms, making use of the whatever fragment of a participant's knowledge base is salient as postulates for deduction. Perhaps a better way to make my proposal consistent with Arregui et al. (2006)'s recycling hypothesis is to assume a lower level VPE resolution mechanism, with reanalysis occurring as a side effect of discourse coherence establishment.

Since the logical forms produced by the sentence syntax are the same structures that are used to represent the knowledge base of speakers and hearers, these logical forms might have more in common with conceptual structures in Culicover and Jackendoff (2005) than Logical Form in, eg Heim and Kratzer (1998), making this more like a semantic identity theory of VPE. But to maintain compatibility with the Arregui et al. (2006) findings, the logical forms of active voice VPs must be distinct from that of passive voice VPs. Since the distinction between passive and active voice has previously been assumed to be a purely morpho-syntactic distinction, this entails that I am arguing for a syntactic theory of VPE. My position may therefore appear contradictory. But I find it consistent to assume that the syntactic level of Logical Form might also be used to store world knowledge in the knowledge base. When talking about cases of ellipsis with no linguistic antecedent, Elbourne (2005) says:

> One sometimes gets the impression that some theorists think that verb-phrase

[1]See Merchant (2004b) for one approach to this apparent problem.

> meanings that are merely contextually salient or able to be worked out, as opposed to occurring as the value of some constituent in the linguistic environment, should not come to be represented as syntactically fully-fledged VPs. It is unclear what the grounds for this view could be, however. If it ever happens that we think of things and then put our thoughts into words, which is not implausible, we are extraordinarily adept at moving from non-linguistic to linguistic modes of representation.

I simply suggest that our world knowledge can be converted to the a linguistic mode of representation for use in ellipsis resolution, for the same reason.

Likewise, since voice affects how a conversation participant's attention is directed by an utterance, I find it reasonable to assume voice has a reflex in conceptual structure. Put another way, active and passive paraphrases seem to highlight different objects as important. If you think of the meaning of an utterance as an message from a speaker to a hearer about what is important to their mutual goals, then active and passive paraphrases must have different meanings. Accepting this, it becomes difficult to empirically distinguish between a syntactic level of Logical Form and a semantic level of Conceptual Structure. I assume that other purely morpho-syntactic alternations, such as gerunds and infinitives, are indistinguishable at this level. This is expected if the logical forms are a semantic structure, but Kennedy (2003) argues that this will also be true for a syntactic level of Logical Form as well, again blurring the empirical consequences for the distinction[2].

This investigation is laid out as follows: In section 2, I go through the necessary background from PSV and Kehler (2002), laying out a theory of coherence establishment through Parallelism and through Cause-Effect coherence. I show how this theory accounts for reciprocals and nominal antecedents. In section 3, I discuss voice and argument structure alternations. In section 4 I explain how the theory I advance can handle the "split antecedent" cases in (9). The discussion throughout will be focused on VP ellipsis and the relevant parallel facts with pronuons. Other forms of ellipsis, such as sluicing, stripping, and fragment answers will not be examined here. They show different patterns, such as not tolerating voice mismatches, and must be restricted from using the mechanisms I posit for VPE resolution.

[2]But cf. Merchant (2001, 2004b) for arguments from sluicing and stripping constructions that weigh in favor of a syntactic view of lf. He analyzes these constructions as remnant movement followed by deletion (phonological suppression) of a Tense Phrase. This predicts a number of syntactic connectivity effects that remain difficult for a semantic approach to ellipsis, even one with apparently linguistic structure such as voice added to the semantic structures.

2 A Unification Grammar for Establishing Coherence

2.1 The Most Specific Common Denominator

PSV start displaying their gramar with simple discourses like (12), and I will too.

(12) a. Mary likes John
b. Susan adores Peter

In analyzing discourses, PSV use a phrase structure grammar with rules for building complex discourse constituent units (DCUs) from simpler DCUs. Clauses are the most basic DCUs. More complicated discourses are constructed incrementally, clause by clause. I will simplify their discourse grammar by ignoring incremental parsing, and just present a discourse grammar that generates a fragment of the discourses that speakers accept as coherent.

The central notion of PSV's discourse grammar is the Most Specific Common Denominator. I will give a precise definition of the MSCD in section 2.1.1. Intuitively, the MSCD of two DCUs ϕ and ψ, written $\phi ¢ \psi$, is whatever syntactic/semantic structure they have in common.

In the theory advanced here, a DCU is a set of attribute-value pairs. Many attributes relevant to discourse grammar and coherence, such as speaker, tense, or modal indexes, will be ignored in this article. I will focus on the following three DCU attributes.

logical form (lf): The context independent syntactic/semantic structure of the DCU. This may contain unresolved discourse variables (called wildcards). PSV call this attribute the **sem** value.

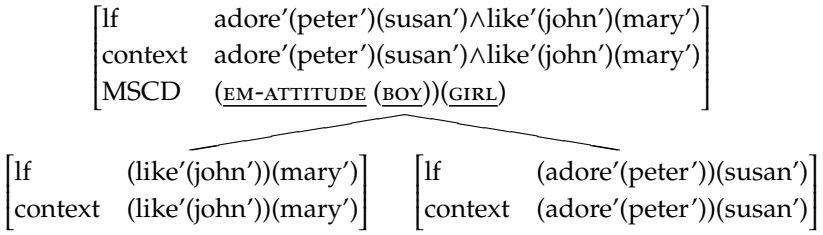
MSCD: The MSCD of the daughters of a complex DCU. PSV call this attribute the **schema** value.

context: The context dependent syntactic/semantic structure of the DCU. This may not contain unresolved wildcards, as it represents the fully resolved meaning of the DCU. PSV called this attribute the **consem** value.

This theory of discourse is illustrated for example (12) in (13). The MSCD value for the root node of the tree is that given for this example in PSV.

(12) a. Mary likes John.

b. Susan adores Peter.

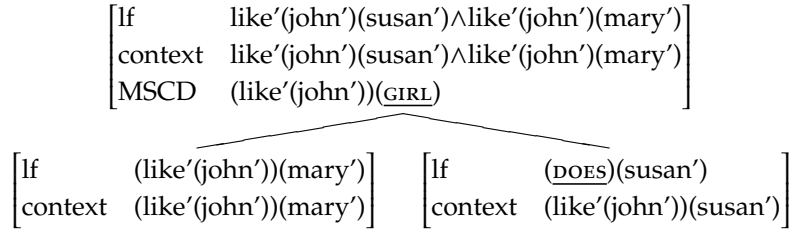
(13)


The symbols EM-ATTITUDE, BOY, and GIRL are **wildcards**. They are each a variable over a **sort**. A sort is a subset of the objects in the sort's type; GIRL is the subset of the domain of entities containing all and only girls, BOY is the subset of the entities containing all and only boys, and EM-ATTITUDE is the subset of the two place relations that are true just in case the first argument has an emotional attitude towards the second argument. Thus, the MSCD attribute of the overall discourse captures the fact that both sentences are about a girl having an emotional attitude towards a boy, and it is in virtue of these commonalities that the discourse in (12) is seen as coherent.

But (12) doesn't show the most important use of wildcards. In (14) a wildcard is used in the lf attribute as the meaning of an elided VP. This wildcard is resolved in the context value for that clause by unifying the lf value with the parents MSCD value.

(14) a. Mary likes John.

b. Susan does too.

c.


In the next section, I will describe the details of why the MSCD for this discourse is (like′(john′))(GIRL). This value is used to resolve the value of the wildcard DOES, which is reflected in the context value of the second clause and the mother node of the discourse.

2.1.1 Logic and Definitions

In this section I'll characterize the data structures that the Most Significant Common Denominator is calculated over and give a precise definition of the MSCD and the Most General Unification (MGU), which is used to resolve ellipsis in (14).

The structures that the MSCD is computed over are the well-formed formulas of a standard typed logic, augmented with a set of sorts for each type, an order relation on the sorts of a type, and wildcards for each sort. I will briefly describe the properties of sorts and wildcards for calculating the MSCD. See PSV, section 2 for the details of the syntax and semantics of this logic.

This logic is similar to the Intensional Logic used in Montague Grammar defined in Dowty, Wall, and Peters (1981). Unlike the IL used in Montague Grammar, the sorted logic used here is not a dispensable step in derivation of the model theoretic semantics. It is an indispensable data structure over which discourse constraints are specified, and it is used directly for discourse coherence calculations. This makes the sorted logic more like the predicate-argument structure from Steedman (2000) or the syntactic level of Logical Form in Heim and Kratzer (1998).

Adding sorts and wildcards to the type logic requires the following definitions, in addition to the standard syntax and semantics of a typed logic.

CS_α The set of sortal constants of type α. In particular, $CS_t = \{\text{TV}\}$. Sortal constants already mentioned include $\{\text{BOY}, \text{GIRL}\} \subset CS_e$, $\text{EM-ATTITUDE} \in CS_{\langle e,\langle e,t\rangle\rangle}$, and $\text{DOES} \in CS_{\langle e,t\rangle}$. Each sortal constant denotes a subset of the domain of α, D_α.

Order $\preceq_\alpha$ An ordering relation on the sorts of type α. If σ and τ are sorts of type α, then $\sigma \preceq_\alpha \tau$ iff the subset of D_α denoted by σ is a subset of that denoted by τ

Wildcards A wildcard of a given sort is an anonymous unification variable, and is indicated by underlining it. This means that the wildcard $\underline{\sigma}$ of sort σ,

1. ranges over objects in the denotation of σ, and
2. is interpreted independently of all other occurrences of $\underline{\sigma}$.

This lets us put sorted wildcards into our logic, and with some special bookkeeping we can force the assignment function to interpret each occurrence of a wildcard independently (see PSV for definitions that accomplish this). But we still need to define the MSCD and MGU operations that are used to calculate coherence and resolve ellipsis. The MGU is defined as a greatest lower bound (join operator) on the objects in an algebra ordered

by specificity relation. So to define the operations we need, we must define a specificity preordering relation. This is done by defining:

$s(\phi)$ A function from logical formulas to sorts. For any formula ϕ of type α, $s(\phi) = \sigma$, where σ is the smallest sort of type α according to $\preceq_\alpha$ containing ϕ.

Preorder $\sqsubseteq$ A preorder relation over formulas of the logic. $\phi \sqsubseteq \psi$ is read ϕ is at least as specific as ψ or ψ is at least as general as ϕ[3]. The relation is defined by specificity axioms. The basic axioms are:

1. For logical constants $a, b; a \sqsubseteq b$ iff $a = b$
2. For sorts σ, τ of type α; $\underline{\sigma} \sqsubset \underline{\tau}$ iff $\sigma \preceq_\alpha \tau$
3. For logical formula $\phi; \phi \sqsubseteq s(\phi)$
4. For formulas ϕ, ψ of type $\langle \alpha, \beta \rangle$ and formulas χ, ξ of type α; $\phi(\chi) \sqsubseteq \psi(\xi)$ iff $\phi \sqsubseteq \psi$ and $\chi \sqsubseteq \xi$.

Using this definition of the 'is more specific than' preorder, the MGU is defined as a join operation on the preorder.

Most General Unification The most general unification of ϕ and ψ of the same type, written $\phi \sqcap \psi$, is the object χ of that type such that

1. $\chi \sqsubseteq \psi$,
2. $\chi \sqsubseteq \phi$, and
3. $\forall \sigma$ of that type such that $\sigma \sqsubseteq \psi$ and $\sigma \sqsubseteq \phi$, $\sigma \sqsubseteq \chi$.

When no such χ exists in the model, we say that ϕ and ψ don't unify, written $\phi \sqcap \psi = \bot$. This is just notation; $\bot$ is not a member of the unification algebra.

The MSCD is given a definition similar in style to the definition of the MGU.

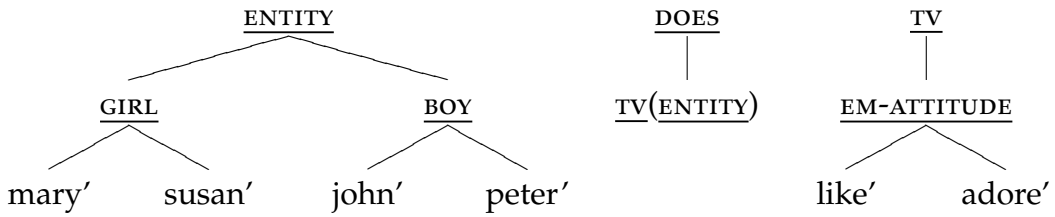
Most Specific Common Denominator The MSCD of ϕ with respect to ψ of the same type, written $\phi ¢ \psi$, is the object χ of that type such that

1. $\phi \sqsubseteq \chi$,
2. $\chi \sqcap \psi \neq \bot$, that is χ and ψ unify, and
3. $\forall \sigma$ of that type such that $\phi \sqsubseteq \sigma$ and $\sigma \sqcap \psi \neq \bot$, $\chi \sqsubseteq \sigma$

[3]This is a preorder and not an order because, for instance $\exists y[P(y)] \sqsubseteq \exists x[P(x)]$ and $\exists x[P(x)] \sqsubseteq \exists y[P(y)]$ yet $\exists y[P(y)] \neq \exists x[P(x)]$

Assuming the algebraic structure in (15) let's us compute the MSCD and MGU values that we need to model coherence and resolve ellipsis in (12) and (14).

(15)



(12) a. Mary likes John.
lf: like′(john′)(mary′)

b. Susan adores Peter.
lf: adore′(peter′)(susan′)

c. (12a) ¢ (12b) = EM-ATTITUDE(BOY)(GIRL)

(14) a. Mary likes John.
lf: like′(john′)(mary′)

b. Susan does too.
lf: DOES(susan′)

c. (14a) ¢ (14b) = like′(john′)(GIRL)

d. (14c) ⊓ (14b) = like′(john′)(susan′)

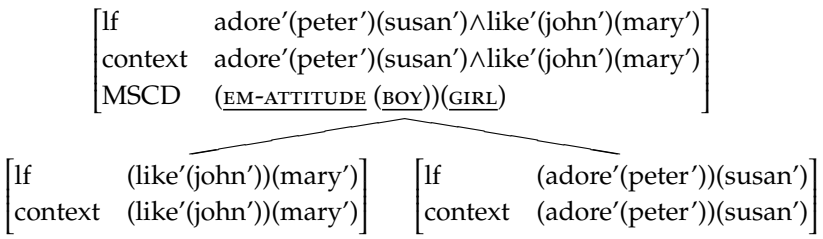
Now we need to use these operators to define rules for constructing discourses from clauses.

2.1.2 Discourse Grammar

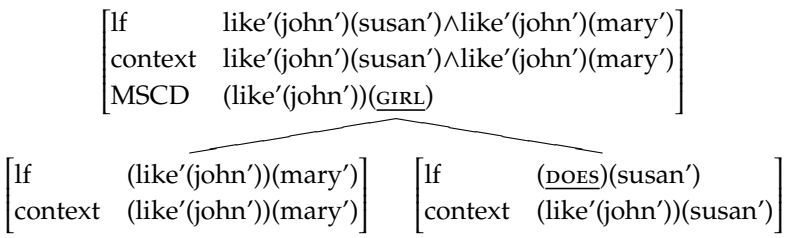
PSVs discourse grammar parses discourses into Discourse Constituent Units (DCUs), which are composed of smaller DCUs connected by a rhetorical relation. Rhetorical relations are usually expressed by coordinating or subordinating conjunctions or sentence adverbials, though they may remain unexpressed, in which case they must be inferred. Atomic DCUs are clauses.

PSV give many rules for constructing complex DCUs from simpler DCUs. I will focus on their rule called List, which I call Parallel following Kehler (2002). The Parallel rule is used to assemble the complex DCUs in (13) and (14), repeated from above.

(13)

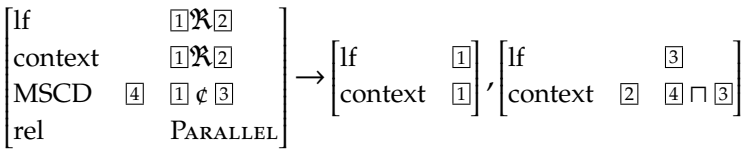


(14)



Trees (13) and (14) have quite a few properties in common that should be captured by the parallel rule. The first daughter DCU has no wildcards that need to be resolved, and the lf and context values for the first daughter DCU are identical. Generally Parallel does not allow cataphora, where pronouns or VPE in the first clause is resolved to a referent introduced later, so this should be encoded in the Parallel rule. The MSCD value of the parent is just the MSCD of its daughters. The context value of the second daughter DCU is the MGU of its lf value with the MSCD value of the parent DCU. And the lf and context value of the parent are just the conjoined values of the daughter context values. More generally, the lf and context values of the parent are the values of the daughter context values related by whatever sentence connective is used to connect the two clauses, with conjunction being the default connective when none is overly expressed. These constraints are captured in the Parallel rule given in (16).

(16) PARALLEL:



Constraints:

1 has no wildcards.

4 has no trivial wildcards.

$\mathfrak{R}$ is the propositional relation expressed by a clausal connective. The default relation is propositional conjunction when no other relation is expressed.

The boxed numbers that appear as the values of some attributes and as labels of others,

such as $\boxed{1}$, are used to indicate identity constraints[4]. For instance, the context value of the second daughter in this rule is labeled $\boxed{2}$, so any other occurrences of $\boxed{2}$ in the rule have the value of the context value of the second daughter. This value itself is the formula $\boxed{4} \sqcap \boxed{3}$. Since $\boxed{4}$ is the label given to the value of the MSCD attribute of the parent and $\boxed{3}$ is the value of the lf attribute of the second daughter, this formula says that the value of the context attribute is the Most General Unification of its lf value and the MSCD of its parent, which is what we want.

The Parallel rule requires that the MSCD value of the parent has no trivial wildcards. PSV define a trivial wildcard as a wildcard of a sort that is marked as trivial. Sorts that are not marked as trivial are meant to be psychologically natural generalizations of the objects of that sort. For the purposes of this paper, a sort is trivial if it is the top sort of its type. The trivial sorts we'll see in future examples are ENTITY, the top sort of type e, and VP, the top sort of type $\langle e, t \rangle$.

It's this triviality constraint that rules out (17) as an infelicitous discourse. But PSV note that discourse (18) is acceptable. For this discourse, they introduce another discourse relation that they call rhetorical subordination, which can apply whenever the sentence connective is a discourse subordinator such as *because*. This discourse rule is the same as Parallel, but it allows a trivial MSCD.

(17) a. John goes to the theater.
b. # Lia gives a concert.

(18) a. John goes to the theater.
b. It's because Lia gives a concert.

While relaxing the triviality requirement for rhertorical subordination is necessary, the MSCD account still fails to capture what it is about these discourses that makes them coherent. For instance, for (18) the MSCD based grammar does not account for the presuppositions that are required to see this discourse as coherent, like the fact that that Lia is giving the concert at the theater that John is going to or that John wants to see Lia's concert. To get this, we need a new way to establish a discourse as coherent that does not rely on calculating the MSCD.

[4]This is an abuse of notation. Sometimes two occurrences of a numbered tag must be reference identical (like the two VPs in (1c)) and sometimes they must be value identical (like the two VPs in (2)). This distinction is not relevant to the discussion that follows.

2.2 Coherence Relations

Kehler (2002) is a more recent view of discourse coherence establishment. He proposes three broad categories of coherence relations, illustrated by (19) – (21).

(19) **Resemblance (Parallel)**
Dick Gephardt organized rallies for Gore, and Tom Daschle distributed pamphlets for him.

(20) **Cause-Effect (Explanation)**
George is dishonest. He's a politician.

(21) **Contiguity (Occasion)**
A flashy-looking campaign bus arrived in Iowa. Soon afterward, George W. Bush gave his first speech of the primary season.

I will ignore Contiguity relations and instead focus on Resemblance and Cause-Effect relations.

Of the Resemblance relations Kehler identifies, I will only attempt to model Parallel. Parallel requires that the main predicates of two clauses share some salient property, and that the parallel arguments of those clauses also share a common property. So (19) is coherent because Dick Gephardt and Tom Daschle are both high-ranking Democrats, orgainzinig rallies and distributing pamphlets are both ways of supporting a candidate, and *Gore* and *him* both refer to the same person. The MSCD is just a way of calculating these common properties. In this case, the MSCD of the two clauses is for′(gore′)(SUPPORT)(DEMOCRAT).

The other Resemblance relations Kehler identifies are shown in the table below, taken from Kehler (2002, ch. 2).

Resemblance Relations

Relation	Constraints	Conjunctions
Parallel	$p(p_1) \wedge p(p_2); q_i(a_i) \wedge q_i(b_i)$	and
Contrast	$p(p_1) \wedge \neg p(p_2); q_i(a_i) \wedge q_i(b_i)$ $p(p_1) \wedge p(p_2); q_i(a_i) \wedge \neg q_i(b_i)$	but
Exemplification	$p(p_1) \wedge p(p_2); q_i(a_i) \subset q_i(b_i)$	for example
Generalization	$p(p_1) \wedge p(p_2); q_i(b_i) \subset q_i(a_i)$	in general
Exception	$p(p_1) \wedge \neg p(p_2); q_i(a_i) \subset q_i(b_i)$ $p(p_1) \wedge \neg p(p_2); q_i(b_i) \subset q_i(a_i)$	however nonetheless
Elaboration	$p_1 = p_2; a_i = b_i$	that is

Cause-Effect Relations presupposes that the truth of the proposition expressed by one of the clauses normally gives one good reason to expect that the proposition expressed by the second clause, or its denial, is true. So (20) is coherent because people expect politicians to be dishonest, or to put it another way, being a politician explains why George is dishonest. Notice that (22) is also coherent.

(22) George is honest, even though he's a politician.

It is coherent because the first clause is a denial of the result one would expect given the second clause.

The four Cause-Effect coherence relations Kehler identifies are shown in the table below, also taken from Kehler (2002, ch 2). The squiggly arrow, $\rightsquigarrow$, is an operator which means that the first operand explains or makes it reasonable to expect the second operand.

Cause-Effect Relations

Relation	Presupposes	Conjunction
Result	$P \rightsquigarrow Q$	and (as a result) therefore
Explanation	$Q \rightsquigarrow P$	because
Violated Expectation	$P \rightsquigarrow \neg Q$	but
Denial of Preventer	$Q \rightsquigarrow \neg P$	even though despite

I will propose a model for these inferences similar to the MSCD model of Parallel. This model starts with Kehler's coherence-driven model for pronoun reference resolution, based on the model proposed by Hobbs (1979).

2.2.1 Coherence and Pronouns

In Chapter 6 of his book, Kehler advances a theory of pronouns that breaks pronominal reference possibilities into two cases that are relevant for this paper.

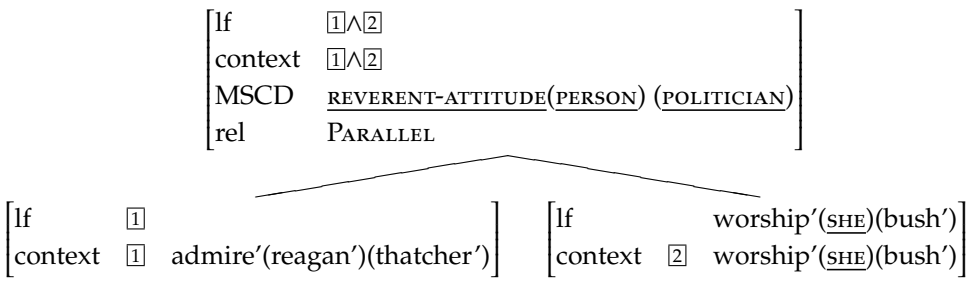
1. When Resemblance holds, the pronoun may only refer to it's parallel argument.
2. When Cause-Effect holds, pronoun reference is resolved by unification with the relevant presupposition.

As I said earlier, the MSCD account models the inferences of Kehler's Parallel relation. Since establishing Parallel matches parallel arguments and resolves wildcards by

unification, our present model agrees with the first case as far as it makes any predictions at all. (52) is an especially striking confirmation of this theoretical prediction because even a ϕ-feature mismatch and a very plausible non-parallel antecedent does not allow a pronoun to take a non-parallel antecedent. (24) shows the results of applying Parallel to these two clauses. It fails to resolve the pronoun because SHE and Ronald Reagan have a wildcard as their MSCD, which unifies with another wildcard to produce a wildcard instead of a grounded term needed for pronoun resolution.

(23) # Margaret Thatcher admires Ronald Reagan, and George W. Bush absolutely worships her. [her ≠ Margaret Thatcher]

(24)



Kehler's account of pronoun resolution when a Cause-Effect relation holds is illustrated by (25). The pronoun *they* is resolved by unification with a presupposed bit of world knowledge. Kehler represents the relevant knowledge as (26).

(25) City council denied the protesters a permit because...
  a. ...they feared violence.
  b. ...they advocated violence.

(26) [fear'(v)(x)∧advocate'(v)(y)∧ allow-to-cause'(v)(y)(z)] ⇝deny'(w)(y)(x)

The representation in (26) contains variables that can be bound by unification, but unlike PSV's wildcards, the variables in (26) are **named**. Multiple occurrences of a named variable must have the same reference. There is no reason why we should not introduce named variables over sorts to the logic as well. Doing so allows us to make world knowledge like (26) more specific without introducing extra conjunctions with predicates restricting the variables. I represent a named variable of sort σ as $\underline{\sigma}_i$, where i is an integer index that is distinct for distinct named variables. Anonymous variables (wildcards) of sort σ are still written as $\underline{\sigma}$, without the index. This means that the knowledge represented as (26) is now represented as the more specific formula (27).

(27) [fear′(RESULT$_1$)(GROUP$_3$) ∧advocate′(RESULT$_1$)(GROUP$_2$) ∧
allow-to-cause′(RESULT$_1$) (GROUP$_2$)(ENTITY$_4$)]⇝deny′ (ENTITY$_4$)(GROUP$_2$) (GROUP$_3$)

Kehler derives the attested readings of (25a) and (25b) as follows. First, the logical form derived by the sentence grammar is (28a). *Because* signals that Explanation holds between the two clauses, allowing one to match the first clause in (25) with the consequent in (27), producing (28b) as the presupposition that establishes the Explanation.

(28) a. deny′(permit′)(protesters′)(city-council′)
b. [fear′(RESULT$_1$)(city-council′)∧advocate′ (RESULT$_1$)(protesters′)∧
allow-to-cause′ (RESULT$_1$)(protesters′)(permit′)]
⇝deny′(permit′)(protesters′)(city-council′)

What happens next depends on whether (25a) or (25b) is uttered. If (25a) follows (25), then the logical form of the next clause is (29). This only unifies with the *fear* term of the presupposition, resolving the presupposition to (29a) and the contextual meaning of the second clause to (29b). This forces *they* to be interpreted as the city council.

(29) fear′(violence′)(THEY)
a. [fear′(violence′)(city-council′)∧advocate′ (violence′)(protesters′)∧
allow-to-cause′ (violence′)(protesters′)(permit′)]
⇝deny′(permit′)(protesters′)(city-council′)
b. fear′(violence′)(city-council′)

(30) advocate′(violence′)(THEY)
a. [fear′(violence′)(city-council′)∧advocate′ (violence′)(protesters′)∧
allow-to-cause′ (violence′)(protesters′)(permit′)]
⇝deny′(permit′)(protesters′)(city-council′)
b. advocate′(violence′)(protesters′)

If (25) continued with (25b) then the logical form of the second clause is (30). It only unifies with the *advocate* term of the presupposition, resolving the presupposition to (30a) and the second clause's contextual meaning to (30b). This forces *they* to be interpreted as the protesters.

2.2.2 Generalizing from Pronoun Reference to Coherence Establishment

If we continue to pursue the idea that this coherence driven pronoun resolution is a side effect of the inferences that establish coherence, then this unification of with a bridging implication should be part of establishing any Cause-Effect coherence relation. And representing VPE as a wildcard will automatically resolve the ellipsis, again as a side effect of the coherence calculation[5].

The Cause-Effect relations are presented in the same format as Parallel. They are given as rules for combining DCUs with constraints on their application. They also introduce a new DCU attribute: presup, whose value is the presupposition that the truth of one of the propositions, possibly augmented with a few other assumptions, explains or at least makes it reasonable to expect the other proposition. That is, the "cause" leads one to normally expect the "effect". (31) shows the rule for Explanation.

(31) EXPLANATION:

$$\begin{bmatrix} \text{lf} & \boxed{1}\ \mathfrak{R}\ \boxed{2} \\ \text{context} & \boxed{1}\ \mathfrak{R}\ \boxed{2} \\ \text{presup} & ((\phi \wedge \psi) \rightsquigarrow \chi) \sqcap ((\boxed{3} \wedge \psi) \rightsquigarrow \boxed{4}) \\ \text{rel} & \textsc{Explanation} \end{bmatrix} \rightarrow \begin{bmatrix} \text{lf} & & \boxed{4} \\ \text{context} & \boxed{1} & (\chi \sqcap \boxed{4}) \end{bmatrix}', \begin{bmatrix} \text{lf} & & \boxed{3} \\ \text{context} & \boxed{2} & (\phi \sqcap \boxed{3}) \end{bmatrix}$$

$((\phi \wedge \psi) \rightsquigarrow \chi) \sqcap ((\boxed{3} \wedge \psi) \rightsquigarrow \boxed{4}) \neq \bot$

$\phi \wedge \psi \rightsquigarrow \chi$ is salient world knowledge or can be accommodated by the participants in the conversation.

$\mathfrak{R}$ is the propositional relation expressed by a clausal connective. The default relation is propositional conjunction when no other relation is expressed.

Most of this rule is straightforwardly similar to Parallel. The most interesting part is the formula $\phi \wedge \psi \rightsquigarrow \chi$. To see what this is doing, consider (25b). To establish Result for this example,

1. ϕ = advocate'($\underline{\text{RESULT}_1}$)($\underline{\text{GROUP}_2}$)
2. ψ = fear'($\underline{\text{RESULT}_1}$)($\underline{\text{GROUP}_3}$) $\wedge$ allow-to-cause'($\underline{\text{RESULT}_1}$)($\underline{\text{GROUP}_2}$) ($\underline{\text{ENTITY}_4}$)
3. χ = deny'($\underline{\text{ENTITY}_4}$)($\underline{\text{GROUP}_2}$) ($\underline{\text{GROUP}_3}$)

This straightforwardly encodes Kehler's analysis of pronoun reference in Cause-Effect relations as a rule of discourse grammar. Similar rules for Result and Violated Expectation

[5] In Chapter 3, Kehler provides a different theory of VP Ellipsis based on his coherence theory. See Frazier and Clifton (2006) for experimental evidence disconfirming this theory.

are given in (32) and (33), with the list of additional constraints suppressed, which are the same as the constraints on Explanation.

(32) RESULT:

$$\begin{bmatrix} \text{lf} & \boxed{2}\ \mathfrak{R}\ \boxed{1} \\ \text{context} & \boxed{2}\ \mathfrak{R}\ \boxed{1} \\ \text{presup} & ((\phi \wedge \psi) \rightsquigarrow \chi) \sqcap ((\boxed{3} \wedge \psi) \rightsquigarrow \boxed{4}) \\ \text{rel} & \textsc{Result} \end{bmatrix} \rightarrow \begin{bmatrix} \text{lf} & & \boxed{3} \\ \text{context} & \boxed{2} & (\phi \sqcap \boxed{3}) \end{bmatrix}, \begin{bmatrix} \text{lf} & & \boxed{4} \\ \text{context} & \boxed{1} & (\chi \sqcap \boxed{4}) \end{bmatrix}$$

(33) VIOLATED EXPECTATION:

$$\begin{bmatrix} \text{lf} & \boxed{2}\ \mathfrak{R}\ \boxed{1} \\ \text{context} & \boxed{2}\ \mathfrak{R}\ \boxed{1} \\ \text{presup} & ((\phi \wedge \psi) \rightsquigarrow \chi) \sqcap ((\boxed{3} \wedge \psi) \rightsquigarrow \boxed{4}) \\ \text{rel} & \textsc{Violated Expectation} \end{bmatrix} \rightarrow \begin{bmatrix} \text{lf} & & \boxed{3} \\ \text{context} & \boxed{2} & (\phi \sqcap \boxed{3}) \end{bmatrix}, \begin{bmatrix} \text{lf} & & \neg\boxed{4} \\ \text{context} & \boxed{1} & \neg(\chi \sqcap \boxed{4}) \end{bmatrix}$$

With these rules, the coherence rule now models the inferences a hearer makes in a normal context that are needed to establish the coherence of (18). Recall that this discourse normally lead one to conclude that John wants to see the concert Lia is giving and that concert is in the theater that John is going too.

(18) a. John goes to the theater.
b. It's because Lia gives a concert.

The relevant presuppositions can be encoded as (34).

(34) [give′($\underline{\textsc{performance}_4}$) ($\underline{\textsc{human}_1}$)∧in′($\underline{\textsc{location}_3}$) ($\underline{\textsc{performance}_4}$)∧ want-to-see′ ($\underline{\textsc{performance}_4}$)($\underline{\textsc{human}_2}$)] ⇝ go-to′($\underline{\textsc{location}_3}$)($\underline{\textsc{human}_2}$)

The logical form for *John goes to the theater*, go-to′(theater′)(john′), will unify with the consequent binding $\underline{\textsc{human}_2}$ to john′ and $\underline{\textsc{location}_3}$ to theater′. The logical form of *Lia gives a concert*, give′(concert′)(lia′) will unify with the *give* term in the antecedent, binding $\underline{\textsc{performance}_4}$ to concert′ and $\underline{\textsc{human}_1}$ to lia′. This then fills in the presuppositions that the concert is in the theater John's going to and that John wants to see the concert in the course of establishing an Explanation between the two clauses. This parse of the discourse is shown in (35), which satisfies all of the constraints of the Explanation relation and delivers the correct presuppositions for the discourse.

(35)
$$\begin{bmatrix} \text{lf} & \boxed{1}\ \textsc{because}\ \boxed{2} \\ \text{context} & \boxed{1}\ \textsc{because}\ \boxed{2} \\ \text{presup} & [\text{give}'(\text{concert}')(\text{lia}') \wedge \text{in}'(\text{theater}')(\text{concert}') \wedge \text{want-to-see}'(\text{concert}')(\text{john}')] \\ & \rightsquigarrow \text{go-to}'(\text{theater}')(\text{john}') \\ \text{rel} & \textsc{Explanation} \end{bmatrix}$$

$$\begin{bmatrix} \text{lf} & \boxed{1} \\ \text{context} & \boxed{1}\ \text{go-to}'(\text{theater}')(\text{john}') \end{bmatrix} \quad \begin{bmatrix} \text{lf} & \boxed{2} \\ \text{context} & \boxed{2}\ \text{give}'(\text{a-concert}')(\text{lia}') \end{bmatrix}$$

This analysis of cause effect coherence coupled with the previous assumption that elided VPs are represented as unification variables allows VPE resolution to occur as a side effect of Cause-Effect coherence establishment as well as Parallel establishment. This gives us a new perspective on reciprocal antecedents such as (10). The relevant presupposition is (37a). The logical forms of the two clauses in (10) are (37b) and (37c). The first clause unifies with the antecedent in (37a) and the part of the second clause unifies with the consequent, resolving the VP refernece in the course of establishing a Violated Expectation. This understanding of the the discourse is represented in (38).

(36) Martha and Irv had planned to nominate each other, but Martha can't. [nominate Irv]

(37) a. $\text{plan}'(\text{nominate}'(\text{each-other}'))\ (\underline{\textsc{human}}_1 + \underline{\textsc{human}}_2)$
$\rightsquigarrow \text{can}'(\text{nominate}'(\underline{\textsc{human}}_2))(\underline{\textsc{human}}_1)$

b. $\text{plan}'(\text{nominate}'(\text{each-other}'))(\text{martha}' + \text{irv}')$

c. $\neg\text{can}'(\underline{\text{VP}})(\text{martha}')$

d. $\neg\text{can}'(\text{nominate}'(\text{irv}'))(\text{martha}')$

(38)
$$\begin{bmatrix} \text{lf} & \boxed{1} \wedge \boxed{2} \\ \text{context} & \boxed{1} \wedge \boxed{2} \\ \text{presup} & \text{plan}'(\text{nominate}'(\text{each-other}'))(\text{martha}'+\text{irv}') \rightsquigarrow \text{can}'(\text{nominate}'(\text{irv}'))(\text{martha}') \\ \text{rel} & \textsc{Violated Expectation} \end{bmatrix}$$

$$\begin{bmatrix} \text{lf} & \boxed{1}\ \text{plan}'(\text{nominate}'(\text{each-other}'))(\text{m}'+\text{i}') \\ \text{context} & \boxed{1} \end{bmatrix} \quad \begin{bmatrix} \text{lf} & \neg\text{can}'(\underline{\text{VP}})(\text{martha}') \\ \text{context} & \boxed{2}\ \neg\text{can}'(\text{nominate}'(\text{i}'))(\text{m}') \end{bmatrix}$$

The case of nominal antecedents are analyzed in the same way. For (39), we can make use of the expectation in (40) to resolved the ellipsis and establish coherence for the discourse according to (41).

(39) Harry used to be a great speaker, but he can't anymore. [speak]

(40) great′(speaker′)($\underline{\text{HUMAN}_1}$) $\rightsquigarrow$ can′(speak′)($\underline{\text{HUMAN}_1}$)

(41)

$$\begin{bmatrix} \text{lf} & \boxed{1} \wedge \boxed{2} \\ \text{context} & \boxed{1} \wedge \boxed{2} \\ \text{presup} & \text{great′(speaker′)(harry′)} \rightsquigarrow \text{can′(speak′)(harry′)} \\ \text{rel} & \textsc{Violated Expectation} \end{bmatrix}$$

$$\begin{bmatrix} \text{lf} & \boxed{1} & \text{great′(speaker′)(harry′)} \\ \text{context} & & \boxed{1} \end{bmatrix} \qquad \begin{bmatrix} \text{lf} & & \neg\text{can′}(\underline{\text{VP}})(\underline{\text{HUMAN}}) \\ \text{context} & \boxed{2} & \neg\text{can′(speak′)(harry′)} \end{bmatrix}$$

So far I have avoided the issue of how the right bridging expectation is selected over other possible expectations. But a predictive model of expectation generation is needed to make this analysis of VPE non-trivial. Without a predictive model, the analyst is free to select whatever bridging expectation derives the observed results. While a predictive model for generating appropriate expectations for a discourse is beyond the scope of this paper, I'll do some preliminary speculation about some plausible factors affecting bridge selection.

First, only presuppositions with an antecedent conjunct that unifies with the first clause encountered or a consequent that unifies with the first clause need to be considered, eliminating any possible bridges that fail to accomplish this. Second, the conjunct encountered further reduces the possibilities by signaling which Cause-Effect is most likely to apply according to the Cause-Effect relation table in section 2.2. As the second clause is encountered, other possible bridges that do not have an appropriate term that unifies with the second clause can be eliminated. Of the remaining possible bridges, the most trivial remaining bridge should be chosen, since it represents the most conservative assumptions about what the speaker meant. But this conservative desire for the most trivial bridge is counterbalanced by the fact that the bridge must be a reasonable or plausible expectation to have, evaluated with respect to known general or specific facts. But how we should measure the trivialness of a bridging expectation and how it interacts with plausibility is still mysterious.

Even without a predictive model of expectation generation for Cause-Effect coherence establishment, this model of VPE makes a prediction: reciprocal reference like that in (36) and nominal antecedents like (39) should not be possible in Parallel discourses. And in fact it is difficult to get these sorts of discourses, such as the discourses in (42). Unsurprisingly, the same discourses without ellipsis in (43) are also difficult.

(42) a. # Martha and Irv planned to nominate each other, just like Alice did.
b. # Harry is a great speaker, just like John will.

(43) a. ? Martha and Irv planned to nominate each other, just like Alice planned to nominate Irv.
b. # Harry is a great speaker, just like John will speak.

3 Voice and Argument Structure Mismatches

In this section I will look at the interaction between argument structure mismatches. Both Parallel and Cause-Effect coherence can be established between two clauses that disagree in voice, even when one clause contains an elided VP. Often acceptability drops dramatically when there is a voice mismatch between and elided VP and its antecedent. Similarly, two clauses with change-of-state verbs as the main predicate may be Parallel or in a Cause-Effect relation even if one is causative and the other is inchoative. And their acceptability also degrades sharply when one is elided, although for some speakers this can be repaired to an extent. I will explore how the coherence theory of VPE might explain both the availability of these mismatches and their degraded acceptability.

The theory I will present here is very powerful, and may potentially generate unattested readings for VPE sentences. In the last part of this section I will look at these cases and suggest some plausible restrictions that will rule out such cases.

3.1 Passive/Active Voice Alternation

Because active voice and passive voice paraphrases have different discourse properties and are preferred in slightly different situations, I assume that they have different logical forms. I introduced the problem of active/passive mismatches in the introduction using (44). For this section, I will illustrate my analysis of these examples on the simpler example (45). (45a) shows the logical form of the sentence *I sent the set*. (45b) shows its passive counterpart, and (45c) shows the logical form of the clause *it was* in (45).

(44) Steve asked me to send the set by courier through my company insured, and it was. [sent by courier through my company insured](Kehler, 2002, p 53)

(45) Steve asked me to send the set, and it was.

a. I sent the set.
   lf: sent′(set′)(I′)
b. The set was sent.
   lf: PV(sent′)(set′)
c. $PV(\underline{\text{VP}})(\underline{\text{IT}})$

PV is a function of type $\langle\langle e,\langle e,t\rangle\rangle,\langle e,t\rangle\rangle$ that "strips off" the outermost argument of the verb[6].

(45) is coherent because the second clause is the Result of the first. Establishing Result presupposes that someone asking someone else to do something leads us to expect that that person did what was asked of them. This is encoded in (46a). But the actual presupposition that is needed to unify with the logical forms of the two clauses is (46b).

(46) a. $\text{ask}'(\underline{\text{VP}_1}(\underline{\text{HUMAN}_2}))\ (\underline{\text{HUMAN}_3}) \rightsquigarrow \underline{\text{VP}_1}(\underline{\text{HUMAN}_2})$
b. $\text{ask}'(\underline{\text{VP}_1}(\underline{\text{HUMAN}_2}))\ (\underline{\text{HUMAN}_3}) \rightsquigarrow PV(\underline{\text{VP}_1})$

We could assume that the matching presupposition (46b) is used in this case, but this seems to miss a generalization, and it will fail to account for the next example where the two clauses are Parallel. Rather than using (46a) directly, then, we can either assume that coherence establishment may proceed by unification through a second presupposition, namely (47a), or we can state the specificity axiom (47b).

(47) a. $\underline{\text{TV}_1}(\underline{\text{ENTITY}_2})\ (\underline{\text{AGENT}_3}) \rightsquigarrow PV(\underline{\text{TV}_1})(\underline{\text{ENTITY}_2})$
b. $(\alpha(\beta))(\gamma) \sqsubseteq (PV(\alpha))(\beta)$

Because (47b) does not require any revisions to any of the discourse relations I've presented, I will adopt it for this paper. The derivation of (45) using this axiom is (48).

(48)

$$\begin{bmatrix} \text{lf} & \boxed{1} \wedge \boxed{2} \\ \text{context} & \boxed{1} \wedge \boxed{2} \\ \text{presup} & \text{ask}'(\text{send}'(\text{set}')(\text{I}'))(\text{steve}') \rightsquigarrow \text{send}'(\text{set}')(\text{I}') \\ \text{rel} & \textsc{Result} \end{bmatrix}$$

$$\begin{bmatrix} \text{lf} & \boxed{1} & \text{ask}'(\text{send}'(\text{set}')(\text{I}'))(\text{steve}') \\ \text{context} & & \boxed{1} \end{bmatrix} \quad \begin{bmatrix} \text{lf} & & PV(\underline{\text{VP}})(\underline{\text{IT}}) \\ \text{context} & \boxed{2} & \text{sent}'(\text{set}')(\text{I}') \end{bmatrix}$$

[6]This simplifies the situation, since PV can apply to ditransitives as well. In these cases PV is a function of type $\langle\langle e,\langle e,\langle e,t\rangle\rangle\rangle,\langle e,\langle e,t\rangle\rangle\rangle$. PV is really a function of polymorphic type $\langle\langle e,\alpha\rangle,\alpha\rangle$.

An interesting consequence of this analysis is that the contextual understanding of the second clause is that I sent the package. This seems like the correct default understanding, but the assumption that it was me who sent the package can be suspended or denied without contradicting the discourse as it stands.

Adding (47b) allows voice mismatches to occur in Parallel clauses as well. The modified version of (3c) in (49) is parsed as shown in (50).

(49) Speaker A: Someone mugged Tom yesterday.
Speaker B: Sandy was, too.

(50)

$$\begin{bmatrix} \text{lf} & \boxed{1}\wedge\boxed{2} \\ \text{context} & \boxed{1}\wedge\boxed{2} \\ \text{MSCD} & PV(\text{mug}')(\underline{\textsc{human}}) \\ \text{rel} & \textsc{Parallel} \end{bmatrix}$$

$$\begin{bmatrix} \text{lf} & \boxed{1} & \text{mug}'(\text{tom}')(\text{someone}') \\ \text{context} & & \boxed{1} \end{bmatrix} \qquad \begin{bmatrix} \text{lf} & & PV(\underline{\text{VP}})(\text{sandy}') \\ \text{context} & \boxed{2} & PV(\text{mug}')(\text{sandy}') \end{bmatrix}$$

Axiom (47b) predicts the correct reference possibilities for pronouns in clauses related by Parallel as well. We can see in (51a), from Kehler (2002, Chapter 7), that Parallel only allows pronouns to refer to the same thematic argument, and not the same surface syntactic position or even the most plausible preceding referent, as predicted by Parallel and (47b). If we revise (47b) to handle agents introduced with a *by*-phrase, it would also predict (51).

(51) a. ? Margaret Thatcher admires Hillary Clinton, and she is worshipped by George W. Bush. [she = Hillary Clinton, she ≠ Margaret Thatcher]

b. ? Margaret Thatcher is admired by Hillary Clinton, and she worships George W. Bush. [she = Hillary Clinton, she ≠ Margaret Thatcher]

(52) # Margaret Thatcher admires Ronald Reagan, and George W. Bush absolutely worships her. [her ≠ Margaret Thatcher]

The previous two examples have shown an active antecedent for a passive elided VP. The more common case is a passive antecedent for an active elided VP. Axiom (47b) correctly handles this configuration, too, shown in (53), which presupposes (54). The parse tree showing this derivation is (55)

(53) This problem should have been looked into, but nobody did.

(54) should-have′($\underline{\text{VP}}_i$)($\underline{\text{HUMAN}}_j$) $\rightsquigarrow$ did′($\underline{\text{VP}}_i$)($\underline{\text{HUMAN}}_j$)

(55)

$$\begin{bmatrix} \text{lf} & \boxed{1} \wedge \boxed{2} \\ \text{context} & \boxed{1} \wedge \boxed{2} \\ \text{presup} & \text{should-have′(look-into′(this-problem′))(somebody′)} \\ & \rightsquigarrow \text{did′(look-into′(this-problem′))(somebody′)} \\ \text{rel} & \textsc{Violated Expectation} \end{bmatrix}$$

$$\begin{bmatrix} \text{lf} & \text{should′}(PV(\text{look′}))(\text{problem′}) \\ \text{context} & \boxed{1} \ \text{should′(look′(problem′))(some′)} \end{bmatrix} \qquad \begin{bmatrix} \text{lf} & \neg\text{did′}(\underline{\text{VP}})(\text{somebody′}) \\ \text{context} & \boxed{1} \ \neg\text{did′(look′(problem′))(some′)} \end{bmatrix}$$

(53) contrasts with the previous two simplified examples. It is much more felicitous than the other two examples, without being much longer or more complex in any obvious way. This might be explained by the fact that it is a transition from the passive, a marked form, to the active, the unmarked form. If the function of the passive is to change the salience levels of the verb arguments, then one might expect a move from more marked salience assignments to less marked salience assignments to the arguments will always be more felicitous than a change from the default to a more marked salience assignment. Or, to put it more concretely, in (53), it's not relevant who was supposed to look into the problem, but it's very relevant that nobody looked into the problem. However, in (49), if it's not relevant who mugged Sandy in the second clause, then it's not clear why the agent shouldn't also be demoted in the first clause as well.

Considerations of the naturalness in the relative prominence of the arguments can also explain why the longer examples in (56) are much better than the truncated examples in (57). The longer examples gradually move the focus away from the agent, making the switch to passive a much less drastic change in the relative prominence of the arguments.

(56) a. Steve asked me to send the set by courier through my company insured, and it was. [sent by courier through my company insured](Kehler, 2002, p 53)

b. Speaker A: Someone mugged Tom yesterday.
Speaker B: Oh yeah?
Speaker C: You know, the same thing happened to Mary.
Speaker B: Wow!
Speaker A: You know, now that I think of it, Sandy was, too. [mugged]
(Sag, 1976, 75, footnote 2)

(57) a. # Steve asked me to send the set, and it was.

b. # Speaker A: Someone mugged Tom yesterday.
Speaker B: Sandy was, too.

This sort of explanation is confounded by the fact that the short examples in (57) are far more felicitous when there is no ellipsis. The examples (58) are not perfect, but they are much better than their counterparts in (57).

(58) a. ? Steve asked me to send the set, and the set was sent.
b. ? Speaker A: Someone mugged Tom yesterday.
Speaker B: Sandy was mugged, too.

One explanation for this might be that VPE makes it more difficult to adjust the prominence of the objects in the discourse model, especially if that change in prominence is to a more marked assignment of prominence to the thematic roles of the arguments. This is admittedly very speculative and vague, and that a lot more work needs to be done to explain these data.

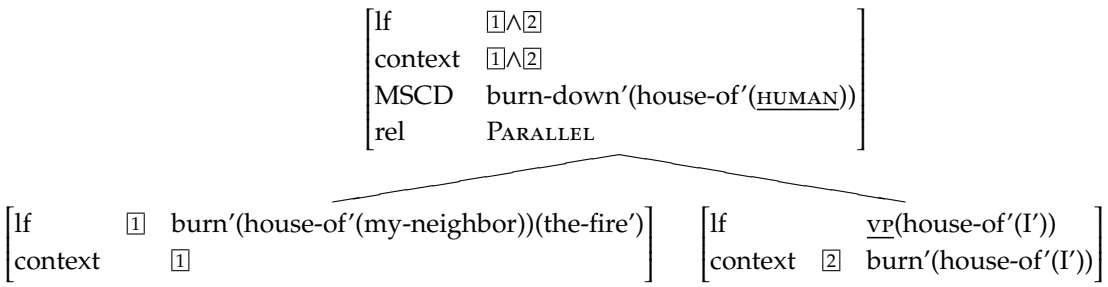
3.2 Causative/Inchoative Alternation

The analysis of voice alternations can be straightforwardly extended to allow coherence to be established between causative and inchoative clauses like those in (59). The specificity axiom needed by the grammar to handle this is (60). The parse for (59) is (61).

(59) ? The fire burned my neighbor's house down just like my house did.

(60) $\alpha(\beta)(\gamma) \sqsubseteq \alpha(\beta)$ iff $s(\alpha) \preceq \Delta$-STATE, where Δ-STATE is a sort denoting the change-of-state predicates.

(61)



One might be wondering whether it's a good thing for the grammar to accept (59). At least some English speakers I've asked reject it completely. This difference among speakers

can be modeled as speakers willingness to use (60), or how salient the implicational knowledge it encodes is to the user. In fact, of the people I've ran this example by, those who accept it have been linguists who have had the causative/inchoative alternation explicitly pointed out to them in the past. One person who rejected it may not have had such experience. One might expect the implication implicit in (60) to be more salient for those who have studied the causative/inchoative pattern.

But then one wonders why speakers would find (62a) acceptable but (62b) unacceptable. Even those who accept (59) tend to reject (62b).

(62) a. Bill melted the copper vase, and the magnesium vase melted, too.
b. # Bill melted the copper vase, and the magnesium vase did, too. [melt]

Frazier and Clifton (2006) argue that certain discourse factors promote an expectation for syntactic parallelism *per se*, and not as a side effect of coherence relations. Discourse factors they identify as promoting coherence include the use of the presuppositional element *too*, use of *and* to conjoin sentences, rather than *or* or *just like*. VPE must also promote parallelism *per se* as well, making (62b) far less felicitous than (62a).

Another way to attack this issue is to split change-of-state verbs into two lexical items: one for the causative, one for the inchoative. If we add a licensing condition to VPE requiring the elided VP to be e-Given (Merchant, 2001), then the unacceptability of a causative/inchoative mismatch under VPE is explained. Givenness constraints and other restrictions on discourses is the topic of the next section.

3.3 The Problem of Overgeneration

The theory of ellipsis I've laid out so far is very powerful. It predicts that any salient knowledge can be used to fill in the referent of the elided VP. Here I will show that other independently needed discourse properties can be used to restrict VPE reference possibilities.

3.3.1 No Optional Bridges

Example (63) from Rooth (1992) shows two things that are important for the analysis presented so far. First, to establish coherence in (63a), the MSCD calculation that establishes Parallel must be allowed to procede through an implicational bridge or must make use of

a specificity axiom relating *tell* statements to *heard about* statements. Adding such a specificity axiom would be *ad hoc* and would not capture parallel cases with other predicates. Generalizing the axiom to cover parallel cases amounts to making the $\sqsubseteq$ relation identical to entailment. I suggest modeling (63a) by presupposing an implicational bridge rather than making specificity identical to entailment, because this makes the presupposition generated by (64) that calling someone a Republican is an insult an explicit property of the discourse, rather than implicit in the entailments the speakers and hearers accept.

(63) a. Someone told Mary about the budget cuts, and Sue heard about them, too.
b. Someone told Mary about the budget cuts, and Sue did, too.

(64) John called Mary a Republican, and then Harry insulted her.

(63b) shows that implicational bridging can only be used when necessary to establish coherence. Given what I've said about (63a), one might expect an implicational bridge could optionally be used to get a reading for (63b) where Sue heard about the budget cuts. However, the lf of the second clause can unify directly with the lf of the first clause and produce a reading where Sue told Mary about the budget cuts. This shorts out any reading requiring an implicational bridge[7]. Another way to think of this is that any use of an implicational bridge necessarily makes the Common Denomenator of the two matched clauses at most as specific, and often less specific, so the lack of optional bridges is a consequence of calculating the Most Specific Common Denomenator.

This proposal can be seen as one way of spelling out Hardt (2005)'s notion of a "semantically visible" violation triggering an implicational bridge. A violation (divergence from syntactic parallelism) is "semantically visible" just in case the incoming lf fails to unify with any available part of the context.

3.3.2 Kill/Die

Restricting the use of an implicational bridge to only instances where its necessary for establishing coherence might still allow us to derive unattested readings. Most literal uses of the verb *kill* entails a statement using the verb *die*. However, examples (65) and (66) show that this entailment cannot be used to resolved the ellipsis[8].

(65) a. Sam died because Max caused him to. [die]

[7] See Hardt (2005) for evidence that implicational bridging is similarly restricted in full unaccented VPs.
[8] Many thanks to Jason Merchant for pointing out the examples in this section.

b. * Sam died because Max did. [kill him].

(66) a. * Sue was killed because her sister did. [die]

b. * Sue killed Alice because Tom had. [died]

The grammar so far explains why (65a) is coherent and has the interpretation indicated. The bridging implication is (67a), which is obviously easily accommodated. A problem here is that (67b) should be equally easily accommodated for (65b).

(67) a. cause'(die'($\underline{\text{ANIMATE}_1}$))($\underline{\text{ENTITY}_2}$) $\rightsquigarrow$ die'($\underline{\text{ANIMATE}_1}$)

b. kill'($\underline{\text{ANIMATE}_1}$)($\underline{\text{ENTITY}_2}$) $\rightsquigarrow$ die'($\underline{\text{ANIMATE}_1}$)

But the unattested reading for (65b) violates Information Structure constraints orthogonal to discourse coherence. Büring (2005) shows that information that is new to the discourse must bear a pitch accent. He defines "new information" using Schwarzschild (1999)'s notion of Givenness, roughly definable as (68).

(68) GIVEN
An expression E is Given iff the existential closure of E is entailed by the existential closure of some previously uttered expression A.

The occurrence of *kill* in the ellipsis site in (65) is not Given, so it must receive a pitch accent. But it cannot get an accent if it is not pronounced.

Givenness does not explain the impossibility of (66). But the presuppositions to be accommodated to accept the asserted Explanation are not plausible either, so it's not obvious that this is really a counterexample. In most circumstances, Sue's sister's death would not explain why Sue would be killed, and similarly Tom's death would not normally explain why Sue killed Alice without more information.

But this response faces a problem: we are perfectly happy to accommodate these explanations when the VPs are pronounced. To save the analysis of VPE, we might try to tell a processing story that prevents discourse processing from considering these VPE possibilities. The presence of an underspecified element such as an elided VP dramatically increases the possible bridging implications that could be used to establish an Explanation compared to a similar utterance without the underspecified element. Such a story could have other benefits. For instance, I expect it would predict that possibilities would be more restricted when larger parts of the sentences are left underspecified: NP deletion and pronouns should be least restricted, while VP ellipsis would be more restricted and

IP ellipsis (stripping and sluicing) would be the most restricted. But it is well beyond the scope of this paper to provide such a story or explore its possible consequences.

Another approach is to follow Merchant (2001) and say that the licensing conditions on the use of VPE are more restricted than Givenness. For a VP to be left unpronounced, the VP must be e-Given, as defined in (69).

(69) e-GIVEN
An expression E is e-Given iff there exists some previously uttered expression A such that:
1. The existential closure of A entails the existential closure of E and,
2. The existential closure of E entails the existential closure of A.

The readings indicated for (66) are not e-Given, so those readings are not felicitous. This e-Givenness analysis of (66) is compatible with a processing story, if such an account can be formulated, and a processing account of this restriction can be used to explain the requirement that an elided VP must be e-Given.

These two forms notions of the "givenness" of an expression form a newness hierarchy which corresponds straightforwardly to a phonological prominence hierarchy.

Newness/Prominence Hierarchy

A constituent that is ...	may be ...	must be ...
New (not Given)	–	pronounced with an accent
Given (not e-Given)	pronounced without accent	pronounced
e-Given	unpronounced (if licensed by syntax)	–

4 Split Antecedents and Ellipsis Containing Antecedents

In the introduction I mentioned two cases of "split antecedents" in (9), which are unaccounted for on the PSV analysis. With the implicational bridge mechanism I've outlined in the last two sections, I can now account for these and similar cases. I will first cover the Webber (1978) cases of split antecedents similar to (9b). Then I will cover cases of ellipsis containing antecedents in (9a).

(9) a. When John had to cook, he didn't want to. When he had to clean, he didn't either.

b. Bob wants to sail round the world and Alice wants to climb Kilimanjaro, and they both will, if they can afford to.

4.1 Split Antecedents

Before I look at some more complicated cases, I will first take a closer look at (70). PSV assert (70c) only has a meaning where Saskia sings, and not one where Saskia sings and dances.

(70) a. Maaike dances.
b. Brigitte sings.
c. Saskia does too. (Prüst et al., 1994)

The theory presented so far fails to predict any coherent readings for this discourse. To get the reading where Saskia dances, it must combine with *Brigitte sings* to form a unit. This complex DCU is combined with the atomic DCU for the clause *Maaike dances*. These two DCUs are intuitively Parallel because *Maaike dances* is parallel to each of the daughters of the second DCU. This can be modeled by allowing Parallel to compute the parent MSCD over either the lf values of the two daughter DCUs or the MSCD values of the two daughter DCUs, whichever is more specific. (71) is a refinement of Parallel that accomplishes this.

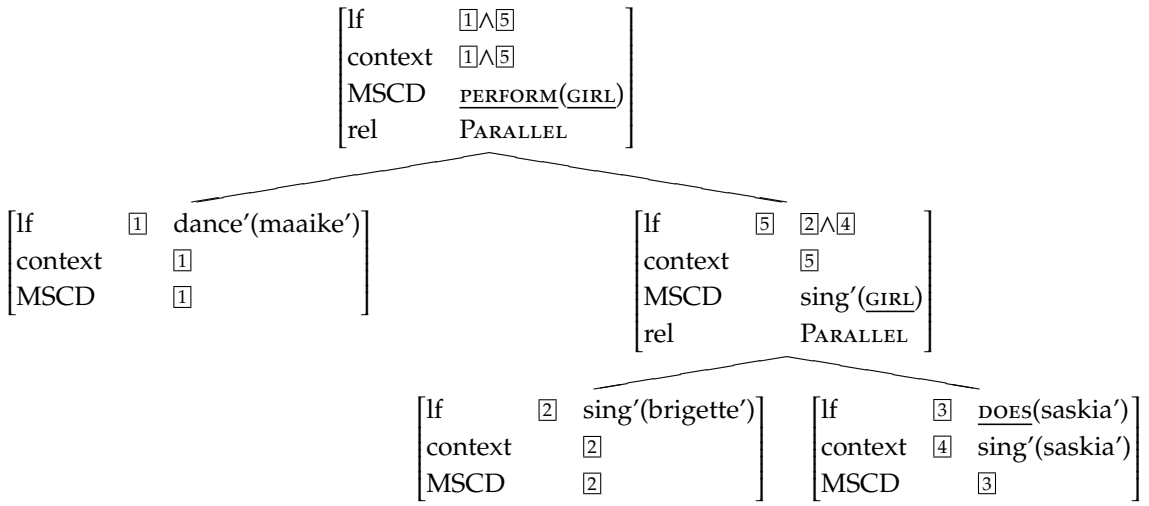
(71) PARALLEL:

$$\begin{bmatrix} \text{lf} & & \boxed{1}\mathfrak{R}\boxed{2} \\ \text{context} & & \boxed{1}\mathfrak{R}\boxed{2} \\ \text{MSCD} & \boxed{6} & (\boxed{1} \,¢\, \boxed{3}) \sqcap (\boxed{4} ¢ \boxed{5}) \\ \text{rel} & & \textsc{Parallel} \end{bmatrix} \rightarrow \begin{bmatrix} \text{lf} & \boxed{1} \\ \text{context} & \boxed{1} \\ \text{MSCD} & \boxed{4} \end{bmatrix}, \begin{bmatrix} \text{lf} & & \boxed{3} \\ \text{context} & \boxed{2} & \boxed{6} \sqcap \boxed{3} \\ \text{MSCD} & & \boxed{5} \end{bmatrix}$$

Constraints:
$\boxed{1}$ has no wildcards.
$\boxed{6}$ has no trivial wildcards.
$\mathfrak{R}$ is the propositional relation expressed by a clausal connective. The default relation is propositional conjunction when no other relation is expressed.

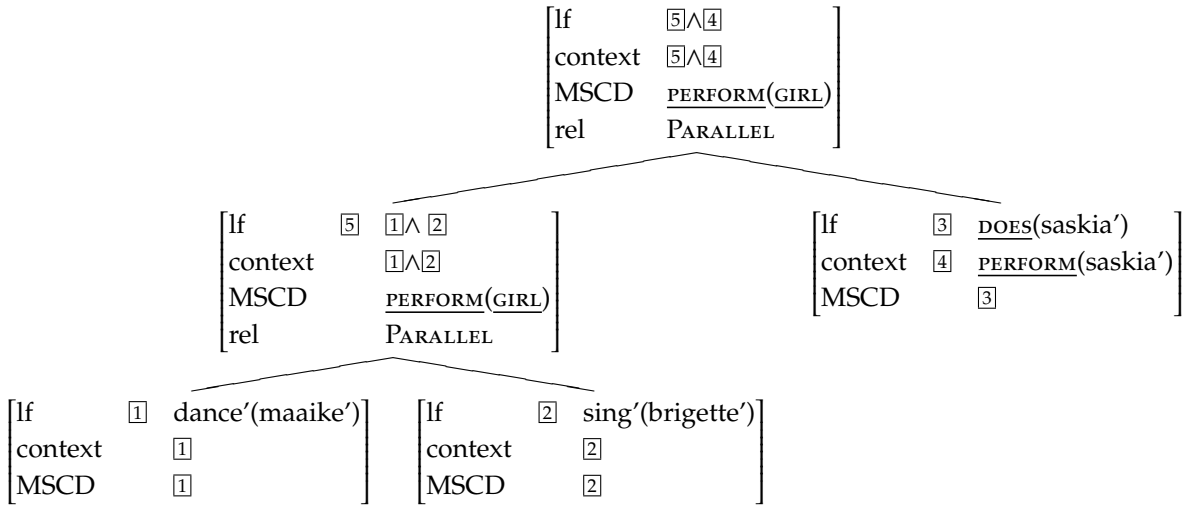
With this version of Parallel, the grammar now derives the desired reading for (70). The discourse parse tree showing that derives this reading is (72).

(72)



The grammar still predicts the "spit antecedent" meaning of (70c) to be infelicitous. (73) shows the discourse tree that results from trying to understand (70c) as meaning that Saskia sings and dances. We get a logical form for the discourse containing an unresolved wildcard. This causes the discourse to be infelicitous. Since (72) is a felicitous parse of this discourse, it will always be preferred.

(73)

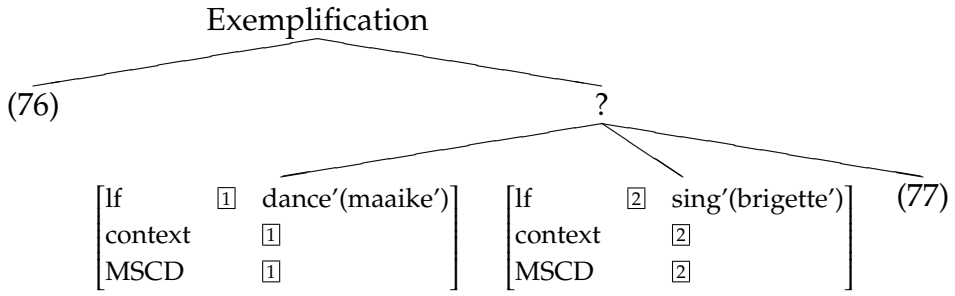


But (74) seems to show that this discourse can have the split antecedent reading, where Saskia sings and dances.

(74) Saskia, being a competitive type, has managed to acquire all the skills Maaike and Brigitte possess.
   a. Maaike dances.
   b. Brigitte sings.
   c. Saskia does too. (Elbourne, to appear)

But the context leading in to (74a–c) is crucial. Now, (74a–c) are no longer coherent in virtue of being Parallel. Instead, together they are an Exemplification of the previous context. The PSV grammar does not model this discourse because they never propose a model for computing Exemplification coherence calculation. I will not present a formal model of this either, but (75) is an informal sketch of how this example might be resolved in the present system.

(75)



(76) Saskia, being a competitive type, has managed to acquire all the skills Maaike and Brigitte possess. $\models$
$\boxed{5}$ [$\underline{\text{ABILITY}_1}$(maaike')$\wedge\underline{\text{ABILITY}_2}$ (brigitte')]$\rightsquigarrow$ λx[$\underline{\text{ABILITY}_1}(x)\wedge$ $\underline{\text{ABILITY}_2}(x)$](saskia')

(77) $\begin{bmatrix} \text{lf} & \boxed{3} & \underline{\text{DOES}}(\text{saskia'}) \\ \text{context} & \boxed{4} & \lambda x[\text{dance'}(x)\wedge\text{sing'}(x)](\text{saskia'}) \\ \text{MSCD} & & \boxed{3} \end{bmatrix}$

In this case, if we unify $\boxed{1}$, $\boxed{2}$, and $\boxed{3}$ with the appropriate subterms of $\boxed{5}$, the wildcard $\underline{\text{DOES}}$ in $\boxed{3}$ will be resolved, producing the correct contextual logical form $\boxed{4}$. How we get from the sentence in (76) to the formula $\boxed{5}$, how to precisely model Exemplification coherence, and what we should fill in for the question mark in the tree are all non-trivial problems that I will not address here.

Coming back to (9b), repeated here in (78), this system faces two challenges. It must get the two antecedent VPs inside the ellipsis site, and then it must somehow derive the "respectively" reading that (78) has, where Bob is the one who will sail round the world and Alice is the one who will climb Kilimanjaro, and not some other reading where they both do both activities, or where Alice does both and Bob does neither.

(78) Bob wants to sail round the world and Alice wants to climb Kilimanjaro, and they both will, if they can afford to.

Start with the first two clauses. *Bob wants to sail round the world* and *Alice wants to climb Kilimanjaro* are Parallel, represented in the subtree (79). These first two clauses are connected to the next two clauses by Result. The presupposition used to establish this relation is (80). (81) shows the logical form of these DCUs. As it is currently stated, Result, repeated in (32), cannot be established.

(79)

$$\begin{bmatrix} \text{lf} & \boxed{1}\wedge\boxed{2} \\ \text{context} & \boxed{1}\wedge\boxed{2} \\ \text{MSCD} & \text{want-to}'(\underline{\text{ACCOMPLISH-FEAT}})(\underline{\text{HUMAN}}) \\ \text{rel} & \textsc{Parallel} \end{bmatrix}$$

$$\begin{bmatrix} \text{lf} & \boxed{1} & \text{want-to}'(\text{sail-round-world}')(\text{bob}') \\ \text{context} & & \boxed{1} \\ \text{MSCD} & & \boxed{1} \end{bmatrix} \quad \begin{bmatrix} \text{lf} & \boxed{2} & \text{want-to}'(\text{climb-k}')(\text{alice}') \\ \text{context} & & \boxed{2} \\ \text{MSCD} & & \boxed{2} \end{bmatrix}$$

(80) $[\text{want-to}'(\underline{\text{VP}}_j)(\underline{\text{HUMAN}}_i)] \rightsquigarrow [\text{can-afford-to}'(\underline{\text{VP}}_j)(\underline{\text{HUMAN}}_i) \rightarrow \text{will}'(\underline{\text{VP}}_j)(\underline{\text{HUMAN}}_i)]$

(81) a. want-to′(sail-round-world′)(bob′)∧want-to′(climb-k′)(alice′)

b. (can-afford-to′(VP))(THEY) →(will′(VP))(THEY)

(32) RESULT:

$$\begin{bmatrix} \text{lf} & \boxed{1}\,\mathfrak{R}\,\boxed{2} \\ \text{context} & \boxed{1}\,\mathfrak{R}\,\boxed{2} \\ \text{presup} & ((\phi \wedge \psi) \rightsquigarrow \chi) \sqcap ((\boxed{3} \wedge \psi) \rightsquigarrow \boxed{4}) \\ \text{rel} & \textsc{Explanation} \end{bmatrix} \rightarrow \begin{bmatrix} \text{lf} & & \boxed{3} \\ \text{context} & \boxed{2} & (\phi \sqcap \boxed{3}) \end{bmatrix}' \begin{bmatrix} \text{lf} & & \boxed{4} \\ \text{context} & \boxed{1} & (\chi \sqcap \boxed{4}) \end{bmatrix}$$

Constraints:

$((\phi \wedge \psi) \rightsquigarrow \chi) \sqcap ((\boxed{3} \wedge \psi) \rightsquigarrow \boxed{4}) \neq \bot$

$\phi \wedge \psi \rightsquigarrow \chi$ is salient world knowledge or can be accommodated by the participants in the conversation.

One thing we may try is creating another rule for establishing Result (and other Cause-Effect coherence as well). A rule that might be used is (82). For convenience, I've suppressed the second conjunct in the antecedent of the bridging implication, which may be needed to establish coherence and explain the presuppositions that are accommodated to establish the Result. I could add it to the rule, but for the examples I look at it will just be bound to $\top$ which will have no effect on the presuppositions.

(82) RESULT:

$$\begin{bmatrix} \text{lf} & \boxed{1}\,\mathfrak{R}_1\,\boxed{2} \\ \text{context} & \boxed{1}\,\mathfrak{R}_1\,\boxed{2} \\ \text{presup} & \left\{\begin{array}{l}\phi,\\ (\psi \rightsquigarrow \chi) \sqcap (\boxed{3} \rightsquigarrow \boxed{5}),\\ (\xi \rightsquigarrow \eta) \sqcap (\boxed{4} \rightsquigarrow \boxed{5})\end{array}\right\} \\ \text{rel} & \textsc{Explanation} \end{bmatrix} \rightarrow \begin{bmatrix} \text{lf} & \boxed{3}\,\mathfrak{R}_2\,\boxed{4} \\ \text{context} & \boxed{1} \quad (\boxed{3} \sqcap \psi)\mathfrak{R}_2(\boxed{4} \sqcap \xi) \\ \text{rel} & \textsc{Parallel} \end{bmatrix}, \begin{bmatrix} \text{lf} & \boxed{5} \\ \text{context} & \boxed{2} \quad (\boxed{5} \sqcap \chi)\mathfrak{R}_2(\boxed{5} \sqcap \eta) \end{bmatrix}$$

Constraints:
$(\psi \rightsquigarrow \chi) \sqcap (\boxed{3} \rightsquigarrow \boxed{5}) \neq \bot$ and $(\xi \rightsquigarrow \eta) \sqcap (\boxed{4} \rightsquigarrow \boxed{5}) \neq \bot$
ϕ is salient world knowledge or can be accommodated by the participants in the conversation.
$\psi \rightsquigarrow \chi$ and $\xi \rightsquigarrow \eta$ instantiate ϕ

This additional rule for establishing a Result allows the Result to connect two parallel causes to their results. There is an issue with building more rules for establishing the same relation. As of now I've given no principled reason for this second version of Result to exist in addition to the first. The two rules are doing basically the same thing in only slightly different contexts, and the account of split antecedent cases is much less attractive if we need a new rule to handle every case.

The problem here is not just conceptual. These sorts of elided VPs can have more than two antecedents.

(83) a. Bob wants to sail round the world, Alice wants to climb Kilimanjaro, and Zaphod wants to build a fancy new space ship, and they each will, if they can afford to.

b. Bob wants to sail round the world, Alice wants to climb Kilimanjaro, Zaphod wants to build a fancy new space ship, and Yolanda wants to swim the English channel, and they each will, if they can afford to.

c. ...

This problem may be rectified by generalizing the rules for Cause-Effect establishment. The two versions of Result, call them Result^1 and Result^2, would be special cases of a more general Result^n. I will not give the more general rule, which would also require stating a more general version of Parallel that allowed greater than binary branching. Instead I will proceed using the Result rule in (82) to handle the cases with two antecedent verb phrases.

Using this rule, the discourse parse of (78) is (84).

(84)
$$\begin{bmatrix} \text{lf} & (\boxed{1} \wedge \boxed{2}) \wedge \boxed{3} \\ \text{context} & (\boxed{1} \wedge \boxed{2}) \wedge \boxed{3} \\ \text{presup} & \{(80), (85a), (85b)\} \\ \text{rel} & \textsc{Explanation} \end{bmatrix}$$

(79)
$$\begin{bmatrix} \text{lf} & (81b) \\ \text{context} & \boxed{3} \quad (86) \end{bmatrix}$$

(85) a. [want-to′(sail-round-world′)(bob′)]
⇝[can-afford-to′(sail-round-world′)(bob′) →will′(sail-round-world′)(bob′)]

b. [want-to′(climb-kiliman′)(alice′)]
⇝[can-afford-to′(climb-kiliman′)(alice′) →will′(climb-kiliman′)(alice′)]

(86) [can-afford-to′(sail-round-world′)(bob′) →will′(sail-round-world′)(bob′)]
∧[can-afford-to′(climb-kiliman′)(alice′) →will′(climb-kiliman′)(alice′)]

The final contextually determined value of the two ellipsis clauses (86) correctly predicts that the clause *they both will, if they can afford to* means that Bob will sail round the world if he can afford to and Alice will climb Kilimanjaro if she can afford to.

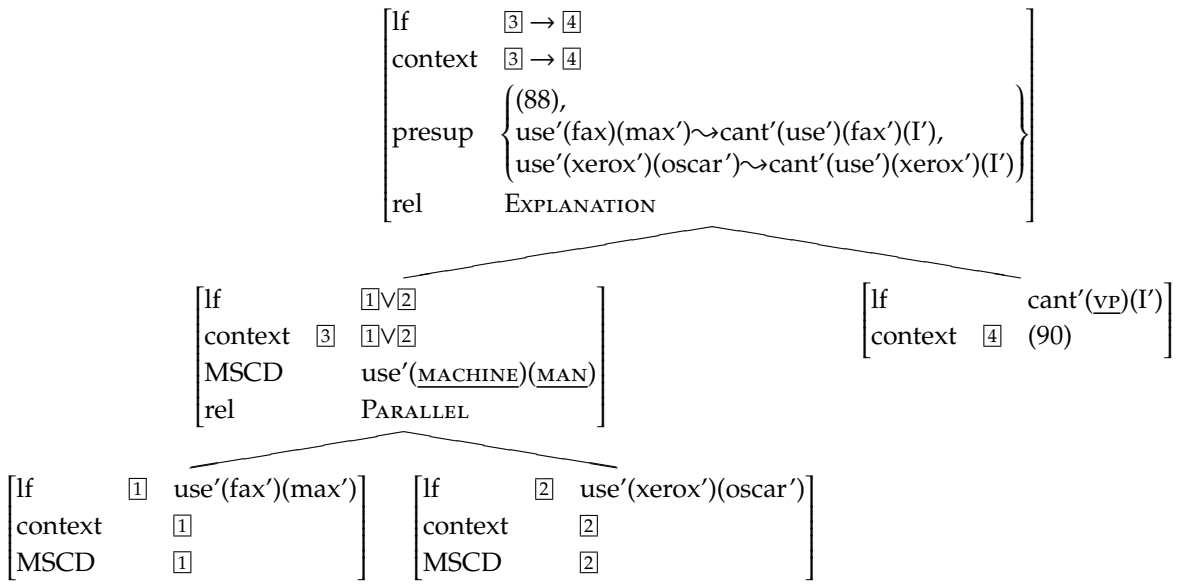
(82) generates the correct interpretation of (87) as well. In this case, the VPE clause means that I can't use the machine that is being used by Oscar or Max. The bridging implication for this example is (88).

(87) Whenever Max uses the fax or Oscar uses the Xerox, I can't. (Fiengo and May, 1994)

(88) use′($\underline{\textsc{machine}}_j$)($\underline{\textsc{human}}_i$) ⇝ ¬(can′(use′))($\underline{\textsc{machine}}_j$)($\underline{\textsc{human}}_k$)

The parse tree for this discourse is (89).

(89)



(90) cant′(use′)(fax′)(I′)∨ cant′(use′)(xerox′)(I′)

The truth derived for this discourse assert that if Max uses the fax or Oscar uses the Xerox, then I cannot use the fax or I cannot use the Xerox. This is true in the situation where Max is using the fax, and I can't use the Xerox but I can use the fax, a situation in which we would intuitively say that this discourse is false. But this situation violates a presupposition of the discourse, namely that if Max uses the fax, then I cannot use the fax. So for this discourse, the correct interpretation of the ellipsis clause where I cannot use whichever machine is being used at the time is modeled in the presuppositions of the discourse, not the truth conditions. The derived truth conditions are more general than our natural understanding. The derived presuppositions eliminate the unnatural interpretations and leave only the attested readings of the discourse.

4.2 Ellipsis Containing Antecedents

Let's return to (9a). I will present two approaches to account for this case. The first approach posits a new specificity axiom. The second approach involves processing the discourse top-down, rather than incrementally, essentially holding off on resolving the first case of VP ellipsis. The first approach has many advantages, but a closer look at some other ellipsis examples will argue against it.

I repeat (9a) here as (91). The issue is how we can get the second clause to mean *John*

didn't want to clean, rather than *John didn't want to cook* or *John didn't clean*.

(91) When John had to cook, he didn't want to. When he had to clean, he didn't either.

The two sentences are Parallel, but their subclauses are both related by an Occasion relation. Since working out an analysis of Occasion relations is beyond the scope of this paper, I will simply stipulate (92) to resolve pronoun and VPE reference in this case. The hooked arrow, $\hookrightarrow$, denotes the propositional relation expressed by *when*.

(92) $\begin{bmatrix} \text{lf} & \boxed{1}\hookrightarrow\boxed{2} \\ \text{context} & \boxed{1}\hookrightarrow\boxed{2} \\ \text{rel} & \textsc{Occasion} \end{bmatrix} \rightarrow \text{When}, \begin{bmatrix} \text{lf} & \boxed{1} & \phi_{\langle\langle e,t\rangle,\langle e,t\rangle\rangle}(\beta_{\langle e,t\rangle})(\alpha_e) \\ \text{context} & & \boxed{1} \end{bmatrix}, \begin{bmatrix} \text{lf} & & \psi_{\langle\langle e,t\rangle,\langle e,t\rangle\rangle}(\underline{\text{VP}})(\gamma_e) \\ \text{context} & \boxed{2} & \psi_{\langle\langle e,t\rangle,\langle e,t\rangle\rangle}(\beta_{\langle e,t\rangle})(\gamma_e \sqcap \alpha_e) \end{bmatrix}$

(91) is not straightforwardly accepted by the discourse grammar just by adding (92) and calculating coherence purely bottom-up like we have been. If we just apply (92) to the two clauses, and then combine them using Parallel, the result is (93).

(93)

$$\begin{bmatrix} \text{lf} & \boxed{1}\wedge\boxed{2} \\ \text{context} & \boxed{1}\wedge\boxed{2} \\ \text{MSCD} & \text{had-to}'(\underline{\textsc{chore}})(\text{j}') \hookrightarrow \neg\underline{\text{VP}}(\text{j}') \\ \text{rel} & \textsc{Parallel} \end{bmatrix}$$

$$\begin{bmatrix} \text{lf} & \boxed{1} & \text{had-to}'(\text{cook}')(\text{j}') \\ & & \hookrightarrow \neg\text{want}'(\text{cook}')(\text{j}') \\ \text{context} & & \boxed{1} \\ \text{rel} & & \textsc{Occasion} \end{bmatrix} \quad \begin{bmatrix} \text{lf} & & \text{had-to}'(\text{clean}')(\underline{\textsc{he}}) \hookrightarrow \neg\ \text{clean}'(\underline{\textsc{he}}) \\ \text{context} & \boxed{2} & \text{had-to}'(\text{clean}')(\text{j}') \hookrightarrow \neg\text{clean}'(\text{j}') \\ \text{rel} & & \textsc{Occasion} \end{bmatrix}$$

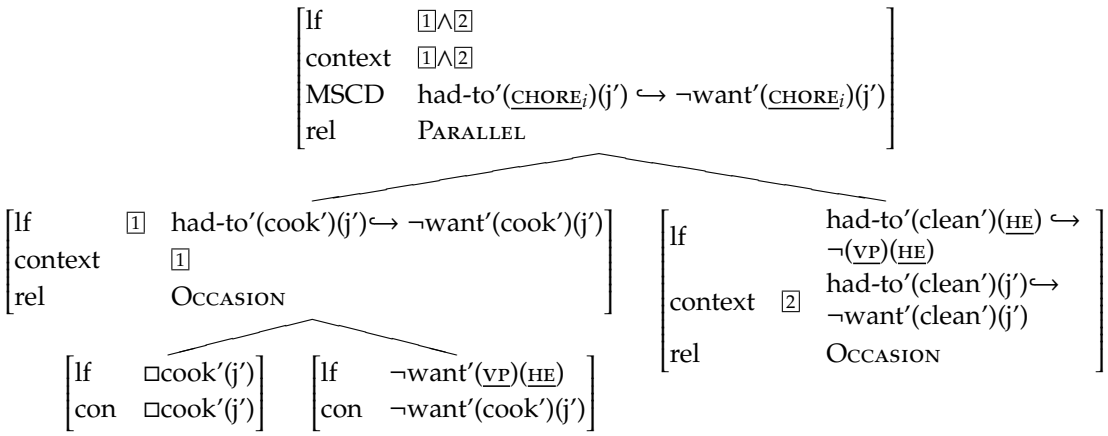
The MSCD for the Parallel node contains the trivial wildcard $\underline{\text{VP}}$, rendering the discourse infelicitous. This explains why the final clause can't be interpreted as *John didn't clean*.

Example (91) will be accepted by the discourse grammar if we add the specificity axiom (94) to the grammar. This axiom says that if two instances of the same wildcard are made to co-vary (making them no longer wildcards but the same named variable), then the result is more specific than the original formula where the wildcards vary independently. The axiom is sound. The multple wildcards could corefer, but need not. The multiple named variables must corefer, so the set of cases that satisfies the formula with named variables is a subset of the cases with wildcards.

(94) $\phi(\underline{\sigma}_i) \sqsubseteq \phi(\underline{\sigma})$ if $\underline{\sigma}$ occurs more than once in ϕ

The axiom is used to parse the discourse in (95). In this parse, the first two clauses are combined by (92). The second two clauses are recognized as parallel. Rather than combining them by (92), they are treated as an atomic DCU that combines by Parallel with the Occasion DCU from the first sentence. This Parallel relation requires (94) to arrive at the MSCD for the two Parallel sentences.

(95)



Adding (94) to the grammar allows the grammar to generate more than one non-equivalent MSCD for two clauses. This ambiguity in the MSCD of two clauses can result in an ambiguity in the interpretation of the discourse, which leads to the strict/sloppy ambiguity discussed prominently in Sag (1976); Williams (1977) and many, many other times in the ellipsis literature. (96) is an example, where the second clause can be interpreted as *Bill loves Bill's wife* or *Bill loves John's wife*. The logical forms of the two clauses are shown in (97), which leads to the two different possible MSCDs in (98). (99) shows the two predicted readings.

(96) John loves his wife, and Bill does too.

(97) a. love′(wife-of′(j′))(j′)

b. DOES(b′)

(98) (97a)¢(97b) =

a. love′(wife-of′(j′))(MAN)

b. love′(wife-of′(MAN$_1$))(MAN$_1$)

(99) a. (97b)⊓(98a) = love′(wife-of′(j′))(b′)

b. (97b)⊓(98b) = love′(wife-of′(b′))(b′)

But this approach overgenerates. It predicts that (100) should have the same reading as (91). I'm not sure what reading (100) most naturally has, but it seems infelicitous to

me. Likewise, (101) should have a sloppy reading where the second clause is interpreted as *7 is obviously less than or equal to 7*. It does not have this reading.

(100) ? When John had to cook, he didn't want to cook. When he had to clean, he didn't either.

(101) # 5 is obviously less than or equal to 5, and 7 is too. (Rooth, 1992, p 5)

It's possible that this approach could be restricted to rule out (100) and (101), or that some other pragmatic principle could make the sloppy readings infelicitous. It's hard to see how such an explanation would go, though.

The second approach is basically the analysis of Elbourne (2005). First, we resolve the second ellipsis site, matching it up to the clause with the first ellipsis site. This copies in *want VP* into the ellipsis site, which still needs to be resolved. The two elided VPs are then resolved as normal, copying the VP from their *when* clauses into their ellipsis site.

To see how this would work in the current framework, let's think about what sort of process a hearer might undergo when (91) is uttered. The hearer might begin to resolve the first ellipsis site when it is encountered. However, shortly after the ellipsis site is encountered, the second sentence begins in a way that very tightly parallels the first. The hearer might then entertain a second parse of the discourse that does not resolve the coherence relation of the first two clauses immediately, but instead first attempts to establish Parallel over the whole utterance. This first step is illustrated in (102).

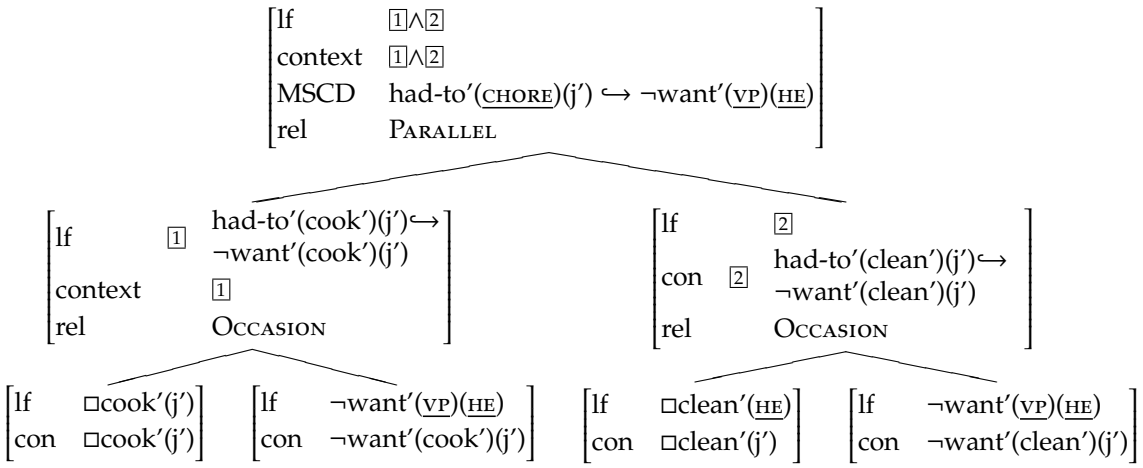
(102)

$$\begin{bmatrix} \text{lf} & \boxed{1}\wedge\boxed{2} \\ \text{context} & \boxed{1}\wedge\boxed{2} \\ \text{MSCD} & \text{had-to}'(\underline{\text{CHORE}})(\text{j}') \hookrightarrow \neg\text{want}'(\underline{\text{VP}})(\underline{\text{HE}}) \\ \text{rel} & \textsc{Parallel} \end{bmatrix}$$

$$\begin{bmatrix} \text{lf} & \boxed{1} & \Box\text{cook}'(\text{j}') \hookrightarrow \neg\ \text{want-to}'(\underline{\text{VP}})(\underline{\text{HE}}) \\ \text{context} & & \boxed{1} \end{bmatrix} \qquad \begin{bmatrix} \text{lf} & & \text{had-to}'(\text{clean}')(\underline{\text{HE}}) \hookrightarrow \neg\underline{\text{VP}}\ (\underline{\text{HE}}) \\ \text{context} & \boxed{2} & \Box\text{clean}'(\text{j}') \hookrightarrow \neg\text{want}'(\underline{\text{VP}})\ (\underline{\text{HE}}) \end{bmatrix}$$

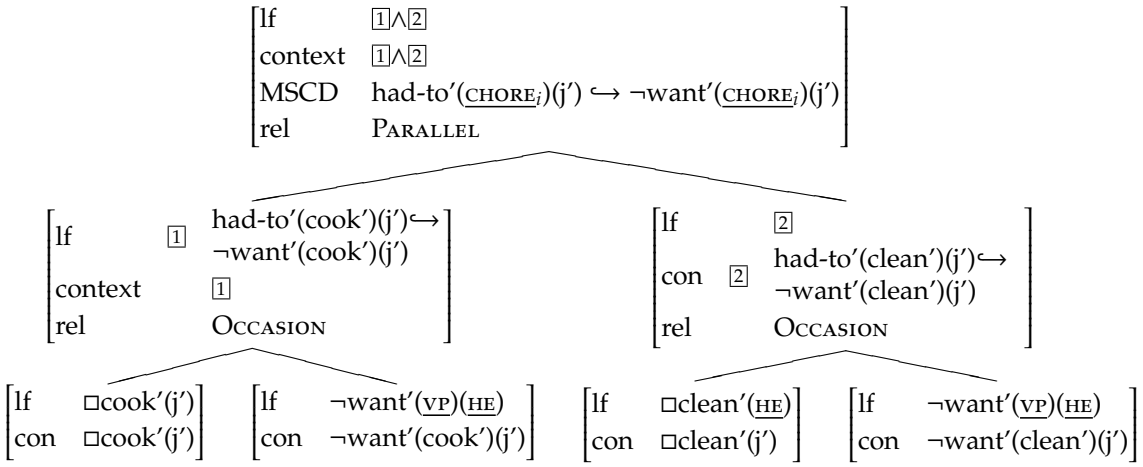
Parallel fails here, first because the MSCD contains a wildcard and the lf and context values of the root node are not yet grounded. But this is ok, because the two arguments of the Parallel relation are each complex, requiring coherence to be established between their atomic clauses. Enforcing the constraints of Occasion in (92) to the discourse built in (102) leads to the discourse parse (103).

(103)



We still aren't done. Enforcing Occasion altered the lf attribute of the second the two arguments to Parallel, making them more specific. This makes nearly every attribute of the Parallel DCU more specific as well. Updating the Parallel node now meets all the Parallel constraints, resolves all wildcards, and establishes the discourse as coherent.

(104)



The last leaf node of (104) is interesting, since it does not match the DCU the sentence grammar would give for the clause, (105b).

(105) a. $\begin{bmatrix} \text{lf} & \neg\text{want}'(\underline{\text{VP}})(\underline{\text{HE}}) \\ \text{con} & \neg\text{want}'(\text{clean}')(\text{j}') \end{bmatrix}$

b. $\neg\underline{\text{VP}}(\underline{\text{HE}})$

But the lf value supplied by the sentence grammar is merely a constraint that must be respected by the discourse grammar. (105) is consistent with this constraint. The lf value is a specification of the lf value derived by the sentence grammar.

5 Conclusions

I have advanced a theory that models the inferences required to establish discourse coherence in a small set of cases. This model is inadequate in many ways: it refers to salience without using a specific model of salience, it requires a model of expectation generation/selection that has not been provided, it does not model non-context-free discourse relations, it does not model any occasion/narration discourse relations, and it does not provide a model for subclausal coherence.

But this model accomplishes something important: it resolves pronoun and VPE reference as a side effect of discourse calculations. This draws together some very distinct ideas and discoveries made in seemingly independent areas of inquiry and connects them in a systematic way. If VPE is resolved by calculating a discourse coherence calculation, then we expect that an elided VP must take its antecedent from a clause that it stands in some discourse relation with. If an elided VP is a unification variable at logical form and VPE resolution involves forcing that variable to refer to a previous structure, then we expect the cost of processing VPE will not vary with the size of the ellipsis referent. Further, if VPE resolution occurs as a side effect of coherence establishment, then no search algorithm for an antecedent is needed beyond the search algorithms already needed to find a way of connecting the clause to the current discourse in a coherent way.

But the use of ellipsis does complicate that search for a coherent way to integrate the clause with the discourse. It provides less information to use as "hooks" to attach the clause to the discourse, which may explain both the restricted inferences that can be used to resolve the ellipsis and what Kehler describes as the signal use of an ellipsis construction sends to a hearer, that the missing material is easily recoverable. If the missing material is not easily recovered, then the coherence calculations could quickly become intractable, and the discourse will be infelicitous.

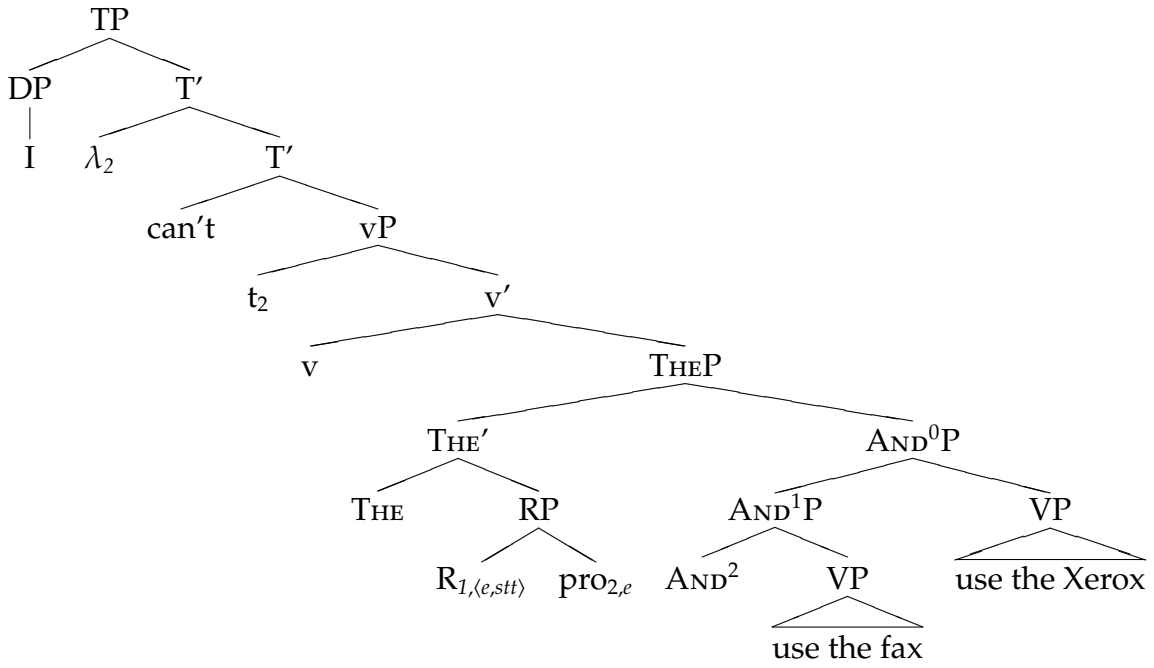
This investigation leads to another important theoretical result. It resolves many cases of apparent syntactic mismatches between the ellipsis and antecedent VPs without requiring a highly articulated, abstract syntactic structure. It is consistent with such syntactic structures, should we find a great deal of independent evidence supporting such structure. For example, Heim et al. (1991) argues that reciprocals in object position scope

the *each* outside of the VP, leaving *the other* inside the VP, so that *Martha and Irv planned to nominate each other* is LF identical to *Martha and Irv each planned to nominate the other*. There are other arguments supporting this analysis of reciprocals, but a requirement for LF identity for VPE is not one of them.

Another important instance of building articulated, abstract structure into the sentence grammar is the invisible restriction variables posited by Elbourne (to appear). In his analysis of split antecedents, an ellipsis site involves a definite determiner with a restrictor variable and possibly many conjoined VPs. For Elbourene, the LF of the second clause in (87) is (106). He paraphrases the interpretation as (107).

(87) Whenever Max uses the fax or Oscar uses the Xerox, I can't.

(106)



(107) Whenever Max uses the fax or Oscar uses the Xerox, I cant perform the particular action or actions out of using the fax and using the Xerox that are being performed.

On this account, the variable $R_{1,\langle e,stt\rangle}$ gets a value like "λX.perform the action or actions out of X that are being performed". He leaves it up to the discourse grammar to assign this value to $R_{1,\langle e,stt\rangle}$. On my account, no variable is necessary in the syntax. The elided VP is resolved to Elbourne's AND^{0}P in (106), with the restriction occurring as a result of the inferences that are used to establish the coherence of the two clauses.

Elbourne points out that the restricting variable $R_{1,\langle e,stt\rangle}$ is similar to abstract restrictor

variables on quantification advanced by von Fintel (1994), Stanley (2000), and Stanley and Szabó (2000). But these domain restrictions may also no longer be necessary when we have adequately modeled the inferences that establish coherence.

This leaves two big points as the take home message from this paper.

1. Ellipsis is economical because discourse coherence calculations allow speakers and hearers to reuse a structure uttered earlier in the discourse.
2. Modeling discourse coherence inferences allows us to model VPE without positing abstract structures in the syntax.

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 — 246.

Asher, Nicholas. 1993. *Reference to abstract objects in discourse: A philosophical semantics for natural language metaphysics*. Kluwer Academic Publishers.

Büring, Daniel. 2005. Intonation, semantics and information structure. In *Interfaces*, ed. Gillian Ramchand and Charles Reiss. URL `http://semanticsarchive.net/Archive/GQ0YjgxM/buring.information.structure.v2005.pdf`.

Chomsky, Noam. 1995. *The minimalist program*. MIT Press.

Culicover, Peter, and Ray Jackendoff. 2005. *Simpler syntax*. Oxford University Press.

Dalrymple, Mary, Stuart M. Shieber, and Fernando C. N. Pereira. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy* 14:399 – 452.

Dowty, David, Robert Wall, and Stanley Peters. 1981. *Introduction to montague semantics*. Reidel Publishing Company.

Elbourne, Paul. 2005. The semantics of ellipsis. In *Interdisciplinary studies on information structure 03*, ed. S. Ishihara, M. Schmitz, and A. Schwarz.

Elbourne, Paul. to appear. Ellipsis sites as definite descriptions. URL `http://www.qmul.ac.uk/~mlw115/LIsplit.pdf`, in Linguistic Inquiry.

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

von Fintel, Kai. 1994. Restrictions on quantifier domains. Doctoral Dissertation, University of Massachusetts, Amherst.

Forbes, Katherine, Eleni Miltsakaki, Rashmi Prasad, Anoop Sarkar, Aravind Joshi, and Bonnie Webber. 2003. D-ltag system: Discourse parsing with a lexicalized tree-adjoining grammar. *Journal of Logic, Language, and Information* 12:261 – 279.

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

Frazier, Lyn, and Charles Clifton. 2001. Parsing coordinates and ellipsis: Copy α. In *Syntax*, 4, 1 – 22.

Frazier, Lyn, and Charles Clifton. 2006. Ellipsis and discourse coherence. *Linguistics and Philosophy* 29:315–346. URL `http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1533872`.

Hankamer, Jorge, and Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7:391 — 428.

Hardt, Daniel. 1993. Verb phrase ellipsis: Form, meaning, and processing. Doctoral Dissertation, University of Pennsylvania. URL `http://www.id.cbs.dk/~dh/papers/93-23.pdf`.

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

Hardt, Daniel. 2003. Ellipsis and the structure of discourse. In *The interfaces: Deriving and interpreting ommitted structures*, ed. Kerstin Schwabe and Susanne Winkler. John Benjamins Publishing Company.

Hardt, Daniel. 2005. Inference, ellipsis, and deaccenting. In *Proceedings of the Amsterdam Colloquium*. URL `http://www.id.cbs.dk/~dh/papers/amst05.pdf`.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar*. Blackwell Publishers.

Heim, Irene, Howard Lasnik, and Robert May. 1991. Reciprocity and plurality. *Linguistic Inquiry* 22.

Hobbs, Jerry. 1979. Coherence and coreference. *Cognitive Science* 3:67 – 90.

Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford: CSLI Publications.

Kennedy, Christopher. 2003. Ellipsis and syntactic representation. In *The interfaces: Deriving and interpreting ommitted structures*, ed. Kerstin Schwabe and Susanne Winkler. John Benjamins Publishing Company.

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

Merchant, Jason. 2004a. A deletion solution to the sloppy ellipsis puzzle. LSA handout.

Merchant, Jason. 2004b. Fragments and ellipsis. *Linguistics and Philosophy* .

Merchant, Jason. 2007. Voice and ellipsis. Unpublished draft.

Merchant, Jason. to appear. An asymmetry in voice mismatches in vp-ellipsis and pseudogapping. URL `http://home.uchicago.edu/~merchant/pubs/VPE.pseudogapping.voice.pdf`, in progress.

Prüst, Hub, Remko Scha, and Martin van den Berg. 1994. Discourse grammar and verb phrase anaphora. *Linguistics and Philosophy* 17:261–327.

Rooth, Mats. 1992. Ellipsis redundancy and reduction redundancy. In *Proceedings of the Stuttgart Ellipsis Workshop*, ed. S. Berman and A. Hestvik. University of Stuttgart. URL `ftp://ftp.ims.uni-stuttgart.de/pub/papers/mats/ellipsis.ps.gz`.

Sag, Ivan. 1976. Deletion and logical form. Doctoral Dissertation, Massachusettes Institute of Technology.

Schwarz, Bernhard. 2000. Topics in ellipsis. Doctoral Dissertation, University of Massachusetts.

Schwarzschild, Roger. 1999. Givenness, avoidf and other constraints on the placement of accent. *Natural Language Semantics* 7.

Stanley, Jason. 2000. Context and logical form. *Linguistics and Philosophy* 23:391 — 434.

Stanley, Jason, and Zoltan Szabó. 2000. On quantifier domain restriction. *Mind and Language* 15:219 —261.

Steedman, Mark. 2000. *The syntactic process*. Cambridge, MA: MIT Press.

Tancredi, Christopher. 1992. Deletion, deaccenting, and presupposition. Doctoral Dissertation, Massachusettes Institute of Technology.

Webber, Bonnie, Matthew Stone, Aravind Joshi, and Alistair Knott. 2003. Anaphora and discourse structure. *Computational Linguistics* 29.

Webber, Bonnie Lynn. 1978. A formal approach to discourse anaphora. Doctoral Dissertation, Harvard University.

Wiebe, Janyce. 1993. Issues in linguistic segmentation. In *Workshop on Intentionality and Structure in Discourse Relations, Association of Computational Linguistics*, 148 – 151. Ohio State University.

Williams, Edwin S. 1977. Discourse and logical form. *Linguistic Inquiry* 8:101 – 139.