Ellipsis and Anaphora in a QUD Model of Discourse

Andrew Kehler

University of California, San Diego

University of Michigan Workshop in Philosophy and Linguistics

1 Introduction

As I was sitting down to begin this paper, I muttered the following to my wife Jill:

(1) I have to start working on a paper. I'm presenting at a workshop at the University of Michigan in October.

A conversation ensued in which she asked me what the workshop was about, what my paper is going to be about, the relationship between the two, and so forth: She clearly inferred that the paper I was going to write was specifically *for* the workshop. Funny, I don't remember saying that.

Of course, Jill did what hearers typically do: They don't generally interpret co-occuring statements as independent facts about the world, they instead assume that they are somehow relevant to one another. In this case, the relevance relation (or 'coherence relation') that is most readily inferred is what various authors (e.g., Hobbs (1990), Asher and Lascarides (2003)) have called EXPLANATION, according to which the second utterance explains (i.e., denotes a cause of) the eventuality described by the first clause. In unexceptional situations, however, the second sentence of (1) can only explain the eventuality described by the first if the paper is being written for the workshop. Although by no means entailed by the passage, this assumption is easily accommodated. Indeed, if I had responded to Jill by saying *Oh, the paper has nothing to do with the workshop, it's for a journal*, she would have been rightfully confused, and perhaps even questioned the sanity of her choice for a partner in life.

According to the approach to discourse coherence just described, establishing coherence is based on making the inferences necessary to meet the constraints imposed by one of a set of coherence relations. In addition to Explanation, other commonly posited relations include RESULT, OCCASION, PARALLEL, ELABORATION, and others (definitions taken or adapted from Hobbs (1990)):

Explanation: Infer P from the assertion of S_1 and Q from the assertion of S_2, where normally $Q \rightarrow P$.

(2) I have to start working on a paper. I'm presenting at a workshop at the University of Michigan in October. (=1)

Result: Infer P from the assertion of S_1 and Q from the assertion of S_2, where normally $P \rightarrow Q$.

(3) I'm presenting at a workshop at the University of Michigan in October. I have to start working on the paper.

Occasion: Infer a change of state for a system of entities from the assertion of S_2, establishing the initial state for this system from the final state of the assertion of S_1.

   (4) Andy walked into his office today. He started working on his paper.

Parallel: Infer $p(a_1, a_2, ...)$ from the assertion of S_1 and $p(b_1, b_2, ...)$ from the assertion of S_2, for a common p and similar a_i and b_i.

   (5) Andy worked on a paper this evening. Jill spent the evening answering email.

Elaboration: Infer $p(a_1, a_2, ...)$ from the assertions of S_1 and S_2.

   (6) Andy worked on a paper this evening. He wrote a draft of the introduction and part of the second section.

The structures of larger discourses are then built recursively: Starting with the assumption that sentences are discourse segments, larger segments (all the way up to the entire discourse) result by relating two smaller segments by a coherence relation. Following the terminology of Moore and Pollack (1992), I will refer to this view as the INFORMATIONAL approach to coherence. Historically, this approach has been applied predominantly to monologues.

   In contrast, other researchers (Grosz and Sidner 1986, inter alia), following work in speech act theory and plan recognition (e.g., Cohen and Perrault (1979), Allen and Perrault (1980)) have taken a different tack. They argued that the role of the utterance in the overall plan underlying the speaker's production of the discourse is the determining factor of coherence. Again following Moore and Pollack, I will refer to this view as the INTENTIONAL approach. In this view, a hearer considers utterances as actions and infers the plan-based speaker intentions underlying them to establish coherence. The intentional approach has been applied predominantly to dialogues; consider, for instance, the following interchange from Cohen et al. (1990):

   (7) Customer: Where are the chuck steaks you advertised for 88 cents per pound?
       Butcher: How many do you want?

A more appropriate information-level response to the customer's question would be *behind the counter*. Yet the butcher's decision to answer her question with another question seems anything but uncooperative. Indeed, an impressive chain of inference has occurred here: the butcher (i) understands that knowing the location of the steaks is a precondition in a plan for coming to obtain steaks, (ii) infers that the customer's goal of obtaining steaks is part of a plan to satisfy the higher-level goal of purchasing the steaks, (iii) realizes that the customer's plan for obtaining the steaks won't work, because she can't get to the steaks herself, (iv) comes up with a different plan, in which the customer obtains the steaks by having him give them to her, (v) realizes that an unsatisfied precondition to *his* plan is knowing the number of steaks she wants, and (vi) asks her a question as a recipe for satisfying this precondition. And, seemingly effortlessly, our customer will presumably respond with an answer, having inferred all of the above in establishing the coherence of the dialog without so much as a blink. Phew!

   Indeed, the informational approach seems to leave us wanting when we consider passages like (7), since no coherence relation posited to apply to the two questions could capture the rich, hierarchical structure of the plan-based inferences that we just witnessed. On the other hand, the intentional approach leaves us wanting when considering passages like (1). For instance, it is not clear (at least to me) what constraints exist on plausible goals, e.g. such that typically hearers aren't happy

with merely assuming that the speaker intends to INFORM the hearer of some facts about himself. Alternatively, we could say that the hearer recognizes the speaker's intention to use the second sentence to perform the action of *explaining* (*justifying, motivating*) the eventuality denoted by the first, but it is only through the coherence-driven reasoning offered by the informational view that Jill was able to figure out what the relevant intention-level action was.

The relationship between these two conceptions of coherence was a topic of some debate in the 1990's (Moore and Pollack 1992, Moore and Paris 1993, Asher and Lascarides 1994, Hobbs 1997, inter alia). Whereas Hobbs (1997), for example, concedes that "the Intentional Perspective is the correct one when we look at things from the broadest possible point of view", he argues that, particularly in written text, "there is a level of detail that is eventually reached at which the Intentional Perspective tells us little". He illustrates his case with an analysis of a dialogue in which the solution to the informational-level coherence problem is a proper subpart of the solution to the intentional-level problem. Moore and Pollack (1992) had likewise argued previously that in fact both levels of analysis must co-exist, illustrating the point with passage (8):

(8) a. George Bush supports big business.

    b. He's sure to veto House Bill 1711.

They point out that at the intentional level, the speaker may be trying to convince the hearer of the claim being made in sentence (8b), and offering sentence (8a) as evidence to support it. At the informational level, she intends that the hearer recognize a Result relationship between the fact expressed in sentence (8a) and the event expressed in sentence (8b). This duality – an unsurprising one, since one way to provide evidence for a proposition is to show that it follows as a consequence of another proposition that the hearer already believes – can allow a hearer to recognize a relation at one level from the recognition of a relation at the other level. For instance, if the hearer knows that House Bill 1711 imposes strong environmental controls on manufacturing processes, but does not know the intentions of the speaker *a priori*, he can infer the intention of providing evidence from having recognized the informational Result relation. Alternatively, if the hearer has no knowledge of the content of House Bill 1711, but has prior reason to believe that the speaker is attempting to provide evidence for the proposition in (8b), he may infer that a Result relation holds, and from this that House Bill 1711 must place undesirable constraints on businesses. As such, there is reason to suggest that both levels co-exist, with links between the two to enable the recognition of relationships on one level from the recognition of relationships on the other.

The question, then, is whether coherence is properly explained by such a two-level approach to reasoning about discourses, or whether there is a more unified way of characterizing these processes. A prospect in this regard is the QUESTION UNDER DISCUSSION (QUD) analysis, proposed by Carlson (1983) and elaborated by researchers including Roberts (1998), Kuppevelt (1995), and Ginzberg (1996). In QUD analyses, discourses are structured by question/answer relationships rather than coherence relations. Roughly speaking, an utterance is coherent insofar as it provides an answer to a (generally implicit) question that is relevant to the preceding discourse.

I will henceforth use Roberts' account as an exemplar. Unlike the informational relations that characterize the coherence relation view, discourses in Roberts' analysis are organized around the conversational goals that interlocutors have and the STRATEGIES OF INQUIRY they employ to satisfy them. In this regard the analysis shares many properties of the intentional approach of Grosz and Sidner (1986) and related works. She suggests, following Stalnaker (1979), that discourse is to be viewed as an attempt by conversational participants to discover and share "the ways things are" (or, to phrase it another way, to answer the question *What is the way things are?*). Thus, by engaging in a conversation, the interlocators agree to jointly adopt goals that center around

finding the answers to this question. This in turn will generally necessitate the adoption and satisfaction of subgoals centered on answering sub-questions, giving rise to a hierarchical discourse structure. Understanding a discourse therefore requires that hearers not only understand the particular utterances in the discourse, but also situate them within the underlying strategy of inquiry.

Whereas the foregoing description highlights the commonalities between her approach and intentional approaches to coherence, Roberts also briefly notes that the coherence relation and QUD analyses are related, in that coherence relations can often be characterized in terms of implicit questions that intervene among utterances. For example, whereas we analyzed example (1) as related by Explanation on its most accessible interpretation, a QUD analysis would presumably posit the implicit question *Why?* as intervening between the two clauses. Note that the constraints on recognizing the coherence of this discourse are similar on the two analyses: Recovering the implicit Explanation relation or the implicit QUD *Why?* both require that the hearer use world knowledge to infer that the second sentence of (1) describes the *reason* for the event described in the first sentence. Other corresponding relationships between coherence relations and questions readily come to mind: *How come?* and *What for?* also correspond to Explanation, *What happened next?* corresponds to Occasion or Result, *Where/when/how?* correspond to Elaboration, and so forth.

The Parallel relation also receives a natural treatment within the QUD analysis, and is worth considering separately because of its special configurational properties, as well as relevance to the rest of the paper. On the informational coherence view, utterances that participate in a Parallel relation are related by a common topic that each utterance instantiates; in the case of (5), repeated below as (10b-c), this topic is roughly *what X did in the evening*. These utterances are equally naturally characterized as providing partial answers to an analogous implicit QUD, as given in (10a). In this case the QUD sits above both utterances rather than intervening between them as in our treatment of (1).[1]

(10) a. { *Who did what in the evening?* }

     b. Andy worked on a paper this evening.

     c. Jill spent the evening answering email.

This treatment captures the oft-cited intuition that the Parallel relation is paratactic, or using the terminology of Rhetorical Structure Theory (Mann and Thompson 1987), 'multi-nuclear'. It also captures the reason why Parallel can relate more than two utterances at once, unlike the other relations, since there is no *a priori* bound on the number of partial answers that a question can have. Finally, Roberts (1998) and Büring (2003) demonstrate that the intonational properties of such passages follow predicted patterns of Question-Answer Congruence under this model (more on this a bit later).

---

[1]Various authors (Jackendoff 1972, Roberts 1998, Büring 2003) have argued that passages like (10) that involve Contrastive Topic marking on parallel entities (using so-called B-accents in the terminology of Jackendoff (1972)) have a more complex structure characterized by the inclusion of subquestions, e.g.:

(9) a. { *Who did what in the evening?* }
     b.   { *What did Andy do in the evening?* }
     c.   Andy worked on a paper this evening.
     d.   { *What did Jill do in the evening?* }
     e.   Jill spent the evening answering email.

We will nonetheless work with the simpler version for now, but revisit this more complex treatment in Section 3.6.

As such, a QUD model integrated within a broader intentional approach potentially provides a synthetic framework for describing discourse coherence in both monologue and dialog. The work of the informational approach is fulfilled by the inference processes used to identify the implicit questions being answered by utterances in a discourse. The work of the intentional approach is fulfilled by the inference processes used to identify how these questions are being used by the speaker in her larger plan, which, importantly, may include a wide variety of subgoals, some of which are satisfied through purely non-linguistic means (as is the case when purchasing steaks, for instance).

This synthesis raises many questions that I will not attempt to address herein. Instead, the remainder of the paper will be devoted to describing several recent results pertaining to ellipsis and anaphora interpretation, cast within a QUD model. These analyses suggest that the QUD model can be used to explain phenomena that have previously required ad-hoc or untenable principles. Further, several experimental studies suggest that comprehenders not only infer latent QUDs, but also have anticipatory expectations about what QUD the next utterance will answer, and that these expectations in turn affect how anaphoric forms are processed. These results, it is hoped, will therefore demonstrate that adequate theories of how particular linguistic expressions are interpreted cannot be developed in the absence of a suitably rich model of discourse coherence establishment.

2 Pronoun Interpretation

My students and I have spent a fair bit of energy during the last several years studying the interpretation of referential third-person pronouns. One reason stems from the paradox of their very existence. After all, in choosing to use a pronoun referentially, a speaker is electing to use a potentially ambiguous expression that, at least on most theories, requires effort to resolve, rather than a less ambiguous or even unambiguous one (such as a uniquely-identifying definite description, or a proper name). Yet pronouns don't seem to get in the way of discourse interpretation very much, if at all, and indeed can even have a facilitating effect as confirmed in the experiments of Gordon et al. (1993, inter alia), which showed that certain discourses tend to be read more slowly if a proper name is used to refer to a focused entity instead of a pronoun.

To my knowledge, Hobbs (1979) was the first to develop a theory of pronoun interpretation specifically based on the establishment of coherence relations. In fact, in his analysis pronoun interpretation is not an independent process at all, but instead results as a by-product of more general reasoning about the most likely interpretation of an utterance. Pronouns are modeled as free variables in logical representations which become bound during these inference processes; potential referents of pronouns are therefore those which result in valid proofs of coherence.

Let us illustrate with passages (11a) and (11b), adapted from an example from Winograd (1972).

(11) The city council denied the demonstrators a permit because...
    a. ...they *feared* violence.
    b. ...they *advocated* violence.

In Hobbs's account, the correct assignment for the pronoun in each case falls out as a side-effect of the process of establishing Explanation (here signalled by *because*), the definition of which is repeated below.

> EXPLANATION: Infer P from the assertion of S_1 and Q from the assertion of S_2, where normally $Q \rightarrow P$.

Oversimplifying considerably, I will code the world knowledge necessary to establish Explanation for (11) within a single axiom, given in (12). (See Hobbs et al. (1993, p. 111) for a more detailed analysis of a similar example.)

(12) $fear(X, V) \land advocate(Y, V) \land enable\_to\_cause(Z, Y, V) \supset deny(X, Y, Z)$

Instantiating it for the current example, the axiom says that if the city council fears violence, the demonstrators advocate violence, and a permit would enable the demonstrators to bring about violence, then this might cause the city council to deny the demonstrators a permit.

The first sentence in (11) can be represented with the predication given in (13).

(13) $deny(city\_council, demonstrators, permit)$

This representation matches the consequent of axiom (12), triggering an inference process that can be used to establish Explanation. At this point, X will become bound to $city\_council$, Y to $demonstrators$, and Z to $permit$.

Each of the follow-ons (11a-b) provides information that can be used to help 'prove' the predications in the antecedent of the axiom, thereby establishing a connection between the clauses. Clause (11a) can be represented with predicate (14), in which the unbound variable T represents the pronoun *they*.

(14) $fear(T, violence)$

When this predicate is used to match the antecedent of axiom (12), the variables T and X are necessarily unified. Since X is already bound to $city\_council$, the variable T representing *they* also receives this binding, and the pronoun is therefore resolved.

Likewise, clause (11b) can be represented as predicate (15).

(15) $advocate(T, violence)$

This predicate also matches a predicate within the antecedent of axiom (12), but in this case, the variables T and Y are unified. Since Y is already bound to *demonstrators*, the representation of *they* also receives this binding.

Thus, the correct referent for the pronoun is identified as a by-product of establishing Explanation in each case. The crucial information determining the choice of referent is semantic in nature, based on the establishment of the relationship between the predication containing the pronoun and the predication containing the potential referents. The fact that coreference came "for free" captures the effortlessness with which people appear to be able to interpret pronouns. Although expressed in an informational theory of coherence, the same line of reasoning presumably applies equally to a QUD model, in which it is used to establish that (11b) is a possible answer to the question *Why?*, in this case cued explicitly by *because*.

2.1   Testing the Coherence Model

Despite the appeal of this example, work in psycholingustics has largely eschewed Hobbs' approach, focusing instead to a large extent on identifying preferences or heuristics that hearers utilize to interpret pronouns. These preferences are often based on linguistic properties of possible antecedent expressions, such as the grammatical and thematic roles that they fill within a sentence. In this section, I describe an experiment that centers on analyzing the question of whether a particular

6

preference exists: the so-called grammatical role parallelism preference. I first discuss the work in terms of an informational theory of coherence (the framework under which the experiment was carried out), and then subsequently recast the results in QUD terms.

Although the purported grammatical role parallelism preference is widely cited across the psycholinguistic literature on pronoun interpretation, the case for its existence has been most forcefully argued by Smyth (1994), who posited an *Extended Feature Match Hypothesis* (EFMH) that characterizes pronoun assignment as a search process that predicts that a "pronoun with two or more grammatically and pragmatically possible antecedents in a preceding clause will be interpreted as coreferential with the candidate that has the same grammatical role" (p. 197). In this analysis, coreference is established by a feature-match process, and due to a priming effect, the identity of the grammatical role filled by the referent is available as one of the criteria for matching, along with other features (e.g. number, gender). This bias is maximized when the clauses containing the pronoun and potential antecedents are fully syntactically parallel; a lack of full syntactic parallelism between the clauses is predicted to prevent syntactic priming and reactivation, resulting in fewer parallel interpretations (pp. 206–207).

We will focus on the two of Smyth's four experiments that are central to our analysis, his Experiments 2 and 3. The Experiment 2 materials were constructed by taking twenty stimuli used Crawley et al. (1990) (who argued that people use a subject assignment strategy in pronoun interpretation, as opposed to a grammatical parallelism strategy) and modifying them so that the clauses were fully parallel syntactically. The nonsubject roles were varied between direct, indirect, and prepositional objects. A sample passage is given in (16).

(16) Mary helped Julie change the tire and then she helped Peter change the oil.

Participants were asked to fill in a blank by writing the name of the person that they understood the pronoun to refer to. The results overwhelmingly favored parallel assignment; 100% of the subject pronouns were assigned to the preceding subject, and 88.12% of the nonsubject pronouns were assigned to the nonsubject referent. His follow-on Experiment 3 tested the prediction that a reduction in the parallelism between the clauses should reduce the number of parallel responses. The results further supported parallel assignment, as the percentage of parallel assignments ranged from 64% to 90% across conditions, but the stimuli in a partial syntactic parallelism condition (in which an adjunct appeared in one clause but not the other) received fewer parallel assignments than those in the fully parallel condition.

As pointed out by Kehler (2002), however, an examination of Smyth's syntactically parallel stimuli suggests that his modifications to Crawley et al.'s examples may have introduced a confound, in that in many cases they appear to have also changed the operative coherence relation from Occasion to Parallel, whereas Occasion appears to be more highly represented in his partially-parallel stimuli. We therefore ran an experiment that controls for and manipulates syntactic parallelism and coherence separately (Kertz, Kehler, and Elman 2006, Kehler et al. 2008). So as to keep the results as directly comparable as possible, our design otherwise followed Smyth's fairly closely, particularly with respect to being an off-line task in which readers are explicitly asked for their pronoun assignments.[2]

In a 2x2x2 design, sixteen stimulus sets were constructed with eight variants as in (17a-d), for a total of 128 experimental stimuli. Each stimulus contains two clauses: an introduction and a follow-on that contains an ambiguous pronoun. Both clauses contain a transitive verb in active voice.

---

[2]Wolf et al. (2004) also ran an experiment to examine the influence of coherence on pronoun interpretation. See Kehler et al. (2008) for discussion and comparison to the experiment described here.

(17)   Samuel threatened Justin with a knife, and

    a. ...Erin blindfolded him (with a scarf).   [Parallel]

    b. ...Erin stopped him (with pepper spray).   [Result]

    c. ...he blindfolded Erin (with a scarf).   [Parallel]

    d. ...he alerted security (with a shout).   [Result]

Each set varied pronoun position (subject/object), sentence structure (fully/partially parallel), and coherence relation (Parallel/Result). Following Wolf et al. (2004), passages participating in Result relationships semantically favored the non-parallel referent, whereas those participating in Parallel relations incorporated no semantic bias. Coherence type was assessed in a norming phase, during which trained judges, blind to our hypothesis, were asked to categorize stimuli as instances of either Parallel or Result coherence. Participants were presented with a two-clause passage like those in (17), and answered a question immediately after that indicated the referent to which they assigned the pronoun.

Smyth's version of the grammatical parallelism preference predicts a strong bias toward interpreting all pronouns to refer to antecedents in a parallel grammatical role when the clauses are fully parallel syntactically, with a weakened effect when the clauses are not fully parallel. The coherence analysis predicts a strong bias toward interpreting pronouns to refer to antecedents in a parallel grammatical role in Parallel relations, and a bias toward the non-parallel antecedent in Result relations. The coherence relation distinction is therefore irrelevant to the grammatical parallelism preference, whereas the full vs. partial syntactic parallelism distinction is irrelevant to the coherence analysis.

The results followed the predictions of the coherence hypothesis, confirming the expected interaction between pronoun position and coherence type. In Parallel relations, 98% of subject pronouns and 90% of object pronouns were interpreted to refer to the previous subject and object respectively, as predicted by both analyses. However, in Result relations, 95% of the subject pronouns were assigned to the previous object, and 94% of object pronouns were assigned to the previous subject. Further, these results were not significantly affected by whether the degree of syntactic parallelism was full or partial. These letter two results are as predicted by the coherence analysis, but contrary to the parallel grammatical role preference.

Although they confirm the coherence analysis, these results still leave us with the question of why we see such a strong parallelism preference when Parallel relations are operative. There has, in fact, always been reason to suggest that this purported preference is not like others that are commonly cited in the literature. For one, the parallelism effect is recalcitrantly strong as compared to these other preferences. For instance, it seems able to withstand strong pragmatic biases toward a non-parallel referent, as in (18):

(18) Condi Rice admires Hillary Clinton, and George W. Bush absolutely worships her.

Assuming a Parallel relation interpretation and a deaccented pronoun *her* (which in turn forces accent on *worships*), informants reliably report that the referent must be Clinton, despite a strong pragmatic bias toward Rice given the political persuasions of the politicians involved. Further, the effect remains even when there is a gender conflict (19):

(19) Condi Rice admires Donald Rumsfeld, and George W. Bush absolutely worships her.

Again, assuming a Parallel interpretation and an unaccented pronoun, informants react as if the speaker is confused about Rumsfeld's gender, even though a female has been mentioned as the

subject of the previous clause. Note that the strength of the parallelism bias cannot be attributed only to the semantics of the Parallel relation, since mentioning either referent with a proper name instead of a pronoun in either of these examples results in a perfectly coherent Parallel passage.

No other preference proposed in the literature is resilient to grammatical and world-knowledge influences in a similar way. Yet, as we have already seen, the effect appears to simply go away when the operative coherence relation is of a different type, such as Result. That is, the coreference pattern that was infelicitous for (18) and (19) is perfectly acceptable on a Result interpretation, per (20a-b) respectively.

(20) a. Condi Rice defeated Hillary Clinton and George Bush congratulated her.   [=Rice]

   b. Condi Rice defeated John Kerry and George Bush congratulated her.   [=Rice]

Lest there be any doubt that these different interpretation patterns are due to the difference in coherence type, we can ask whether passages that are ambiguous between Parallel and Result construals enforce different constraints on the interpretation of unaccented pronouns. This is indeed the case; consider (21):

(21) Powell defied Cheney, and Bush punished him.   (Kehler 2002)

On the Parallel construal of (21) (paraphrase *and* as *and similarly*), *him* can only refer to Cheney if unaccented (i.e., it can refer to Powell only if it receives accent). On the other hand, on the Result construal (paraphrase *and* as *and as a result*), *him* can refer to Powell if it is unaccented. All of these data show a clear pattern whereby relations such as Parallel (those in the Resemblance category per Kehler (2002)) require an unaccented pronoun to corefer with its parallel element, whereby pronouns in non-Resemblance relations like Result are not similarly constrained.

Whereas Kehler et al. (2008) provided an analysis of these facts in terms of informational coherence, here I recast the analysis in terms of a QUD model. Following Roberts, I assume that a speaker produces utterances that are related by implicit QUDs in accordance with her underlying strategy of inquiry, and intones each utterance such that it obeys standard rules of Question-Answer Congruence (Roberts 1998, Büring 2003, inter alia) with respect to the QUD that it answers. I also assume standard rules governing deaccentuation in commonly-held analyses of focus marking and accent placement (Ladd 1980, Selkirk 1996, Schwarzschild 1999, inter alia). Finally, I follow Büring (2003) in noting that, unlike explicit QUDs, implicit QUDs cannot serve to introduce a new element and make it count as 'given' for utterances that follow. With these principles in mind, part of the hearer's job is then to recover these QUDs using the utterance denotations, relevant linguistic cues (e.g., intonational marking), world knowledge, and context.

Let us see how this hashes out for our two types of coherence relation for example (21). On the Result interpretation, we have a discourse structure with an implicit QUD of the following sort:

(22) a. Powell defied Cheney, and

   b. { *What happened as a result?* }

   c. Bush punished him.

The set of alternative answers to (22b) consists of all propositions, and so Question-Answer Congruence requires only that the entire answer be focused. Rules governing deaccentuation tell us that accent on *him* (Given information) is avoided in favor of accent on *punish* (New information). As such, there are no constraints against unaccented *him* referring to the (non-parallel) referent Powell.

Let us now consider the Parallel interpretation, treating the two possible referents for *him* separately. First assume that *him* refers to Cheney. This allows for the following discourse structure:

(23) a. {*Who did what to Cheney?* }

   b. Powell defied Cheney$_c$, and

   c. Bush punished him$_c$.   (*him*=Cheney)

The pronoun *him* remains without accent because (i) Cheney is given as part of the question, and hence is not part of the focus per Question-Answer Congruence, and (ii) can be deaccented because he is mentioned in the first clause.

   Now let's consider the case in which *him* refers to Powell on a Parallel interpretation:

(24) a. { *Who did what to who?* }

   b. Powell$_p$ defied Cheney, and

   c. Bush punished HIM$_p$.   (*HIM*=Powell)

The fact that the events in (24b-c) have different themes means that the theme in the QUD must be a Wh-variable. Since *him* corresponds to this Wh-variable (i.e., it's denotation is part of the QUD's *answer*), Question-Answer Congruence dictates that the referring expression is focal and hence requires accent. The fact that its referent is otherwise Given in the discourse is irrelevant. This scenario is familiar from other cases in which forms expressing Given information nonetheless require accent to establish Question-Answer Congruence, as with the overt question in (25b):

(25) A: Who did John$_j$'s mother praise?

   B: She praised HIM$_j$.   (Schwarzschild (1999), ex. 11)

   Thus, there is no constraint on pronouns referring to non-parallel referents in Parallel relations. Instead, independent rules governing focus/accent placement and Question-Answer Congruence dictate that a pronoun in a Parallel relation can only remain without accent if it is coreferential with its parallel referent.[3]

   This analysis makes a correct prediction: that all of the aforementioned facts concerning coherence and accentuation are actually constraints on *coreference* rather than merely *pronominalization* (Akmajian and Jackendoff 1970, Venditti et al. 2002, de Hoop 2004). That this is the case can be seen by considering variants of our previous examples in which the pronouns are replaced by proper name mentions of their referents. In all of these cases, the requirements on accenting the direct object in Parallel relations are insensitive to whether a full name or pronoun is used:

(26) Condi Rice$_i$ admires Hillary Clinton, and George W. Bush absolutely worships $\left\{ \begin{array}{l} \text{HER}_i \\ \text{RICE} \\ \#\text{her}_i \\ \#\text{Rice} \end{array} \right\}$.

   (cf. 18)

---

[3]The above analysis is restricted to cases in which the QUD sets up a common relation over the utterances. Oehrle (1981) notes that in other "discourse frames" a pronoun can remain deaccented even when not coreferential with its parallel element, as in (i):

(i) A: Can you give me an exact description of Bill's role in the fight?

   B: John hit Bill$_b$ and he$_b$ hit Max.

The difference between this example and the others is that the context sets up Bill as the topic regardless of what thematic role he occupies (e.g. *What about Bill? What was Bill's role?*), as opposed to the question *Who hit who?*. For this QUD, where Bill is not part of the answer in either clause, it is predicted that both references to Bill will remain unaccented, even though they are not grammatically or semantically parallel.

(27) Condi Rice$_i$ admires Donald Rumsfeld, and George W. Bush absolutely worships $\left\{ \begin{array}{l} \text{HER}_i \\ \text{RICE} \\ \#\text{her}_i \\ \#\text{Rice} \end{array} \right\}$.

   (cf. 19)

(28) Powell$_i$ defied Cheney, and Bush punished $\left\{ \begin{array}{l} \text{HIM}_i \\ \text{POWELL} \\ \#\text{him}_i \\ \#\text{Powell} \end{array} \right\}$.

   (cf. 21, *on the Parallel reading*)

Likewise, the lack of accenting on the pronoun in the Result cases remains when a proper name is used instead:

(29) Condi Rice$_i$ defeated $\left\{ \begin{array}{l} \text{Hillary Clinton} \\ \text{John Kerry} \end{array} \right\}$ and George Bush congratulated $\left\{ \begin{array}{l} \text{her}_i \\ \text{Rice} \end{array} \right\}$.

   (cf. 20a-b)

(30) Powell$_i$ defied Cheney, and Bush punished $\left\{ \begin{array}{l} \text{him}_i \\ \text{Powell} \end{array} \right\}$.

   (cf. 21, *on the Result reading*)

Therefore, the information structural constraint at work is one that relates coherence and coreference to accentuation, and is not specific to pronouns. As stated earlier, pronouns are *not* constrained to refer to their parallel elements in Parallel relations. Instead, the information structural constraints imposed by the operative QUD in Parallel relations (but not Result relations) require that the pronoun, like any other referring expression, receive accent when it is not coreferential with its parallel element, and hence unaccented pronouns in Parallel relations will necessarily only corefer with their parallel elements (again, setting aside the case discussed in footnote 3). As such, these data cannot be used to support the existence of a grammatical role parallelism bias, as there is no work left for it to do.

2.2   Toward a Processing Model

The previous sections considered a model in which pronoun interpretation is determined as a by-product of establishing the operative coherence relation (or alternatively, implicit QUD) in the passage. As elegant as that sounds, the true picture is almost certainly not this simple. Pronoun interpretation research in psycholinguistics has been rich source of on-line evidence that interpretation proceeds in a highly incremental fashion (Caramazza et al. 1977, Gordon and Scearce 1995, Stewart, Pickering, and Sanford 1998, Koornneef and van Berkum 2006, inter alia). In the conception we have outlined, however, the inference process operates on complete logical forms of clauses and sentences, leaving unresolved the question of how inference can begin mid-utterance. We thus need to ask how coherence establishment can influence pronoun interpretation in cases in which the pronoun is encountered before the coherence relation is known.

   Recent work (Rohde, Kehler, and Elman 2006, Rohde, Kehler, and Elman 2007, Kehler et al. 2008) has addressed this question, demonstrating that hearers encode probabilistic expectations about how passages are likely to be followed with respect to coherence.[4]   (See Arnold (2001)

---

[4]As before, I will initially discuss the work in terms of an informational theory of coherence, but will ultimately return to its ramifications for QUD analyses.

for a similar proposal, as well as recent work in sentence processing that contends that on-line measurements of interpretation difficulty can be successfully predicted by probabilistic, expectation-driven models, e.g., Hale (2001) and Levy (2008).) The central insight of the model is that any discourse context will give rise to expectations about two types of probabilistic information that are naturally combined: (i) how the discourse is likely to be continued with respect to the ensuing coherence relation, and (ii) the likelihood that a certain referent will get mentioned by a pronoun which, crucially, is *conditioned on those coherence relations*. These come together in the following equation (in which *ante* stands for an antecedent in a particular grammatical or thematic position, and *CR* stands for coherence relation):

$$(31) \quad P(pronoun = ante) = \sum_{CR \in CRs} P(CR) * P(pronoun = ante | CR)$$

For example, to compute the likelihood that a pronoun will corefer with the subject of the previous sentence, we simply sum, over all coherence relations, the likelihood of seeing that coherence relation multiplied by the likelihood of a subject reference given that coherence relation. The equation makes explicit the idea that at any point during comprehension the hearer will have expectations about how the discourse will be continued with respect to coherence, and that the difficulty in interpreting the linguistic material to follow will be conditioned in part on those expectations. These expectations will then evolve based on subsequent linguistic input.

Values for these terms need to be estimated in order to make predictions about on-line interpretation. Sentence completion tasks have become a standard way to estimate biases of this sort (Caramazza et al. 1977, McKoon, Greene, and Ratcliff 1993, Stevenson, Crawley, and Kleinman 1994, Stewart, Pickering, and Sanford 1998, Koornneef and van Berkum 2006, inter alia). Stevenson et al. (1994), for instance, report on a series of such story completion experiments that included passages with a transfer-of-possession context sentence followed by an ambiguous pronoun prompt, as in (32):

(32) John handed a book to Bob. He _____

In such cases, the subject fills the Source thematic role and the object of the preposition fills the Goal role. Participants were asked to provide a natural completion to the pronoun prompt provided in the second sentence, and the pronoun was then categorized as referring to the Source or the Goal. Stevenson et al. found that Goal continuations occurred about as frequently (49%) as Source continuations (51%), an unexpected result in light of the commonly-cited grammatical subject and grammatical role parallelism preferences, since both point to John as the preferred referent.

Stevenson et al. describe two potential explanations for their result: a thematic-role bias which amounts to a heuristic that ranks Goals above Sources, and a bias for focusing on the end state of the previously described event, under the assumption that the Goal is more salient to the end state than the Source. Stevenson et al. ultimately argue for the end-state bias; under this interpretation, the apparent heuristic preference for Goals is an epiphenomenon. Our coherence analysis predicts an end-state bias, but only specifically for passages related by Occasion. That is, among the coherence relations defined in Section 1, only the establishment process for Occasion specifically incorporates a bias toward focusing on the end state of the previous eventuality:

Occasion: Infer a change of state for a system of entities from the assertion of S_2, establishing the initial state for this system from the final state of the assertion of S_1.

As such, the coherence analysis predicts that different pronoun interpretation biases will emerge for different coherence relations, and in particular, that Occasion relations will give rise to a Goal preference.

Rohde et al. (2006) tested these predictions by running an experiment similar to that of Stevenson et al., except that passages like (32) were paired with versions in which the imperfective form of the main verb was used (33).

(33) John was handing a book to Bob. He _____

Crucially, the thematic roles remain the same in examples (32) and (33), but the perfective verb in (32) describes a completed event which is compatible with end-state focus, whereas the imperfective verb in (33) describes an event as an ongoing process, making it incompatible with end-state focus (Moens and Steedman 1988). The thematic role preference thus predicts a similar distribution of Source and Goal interpretations between the two conditions, whereas the event structure hypothesis predicts a greater percentage of Source interpretations in the imperfective condition than in the perfective condition.

The results supported the predictions of the end-state bias. Whereas the perfective condition yielded a near 50-50 split similar to that found by Stevenson et al. (1994), the imperfective condition yielded a much stronger Source bias. To test the predictions of the coherence analysis, judges also categorized the continuations by coherence relation. The results are shown in Table 1;[5] the second and third columns provide estimates of $P(CR)$ and $P(pronoun = source|CR)$ respectively. As can be seen, Occasion, the most common relation, is associated with a strong Goal bias (as predicted), whereas the second and third most common relations, Elaboration and Explanation, are associated with strong Source biases.[6] As such, whereas these numbers combine to yield a near 50-50 split overall, they show that there is nothing 50-50 about the pattern once coherence is taken into account. Each of the coherence relations encodes a considerably stronger bias one way or the other about who will be mentioned next; it is only after the frequencies of coherence continuation are factored in that the biases have a cancelling effect.

Whereas our results support Stevenson et al.'s conclusion that the Goal preference is an epiphenomenon of a bias toward focusing on end states, they further show that the end-state bias is to a large degree an epiphenomenon of the inference processes used to establish Occasion relations. While the context sentences in all of our perfective stimuli describe events with salient end states, the results summarized in Table 1 strongly suggest that it is the coherence relation that dictates the extent to which that endpoint is relevant. Occasion relations exhibit a clear preference for the Goal, as they are precisely the relations that rely specifically on the end state of an eventuality in establishing coherence.

Equation (31) predicts that contextual factors that only influence the distribution $P(CR)$ will nonetheless influence pronoun interpretation biases. Rohde et al. (2007) report on another study

---

[5]The data shown in the table is restricted to the perfective cases since only these are compatible with end-state focus. Analysis of the imperfective condition revealed a different distribution of coherence relations (with fewer Occasions, as expected), but a highly similar relationship between each coherence relation and the corresponding distribution of Source and Goal interpretations, which itself provides further support for the coherence analysis.

[6]We also found a Goal bias for Result relations, but the small set of Result continuations ($< 6\%$; n = 25) was very homogeneous, more than half consisting of the form *X transfers Y to Z. Z thanks X*, making extrapolation difficult. Whereas our coherence analysis would predict that causal inference plays a greater role in establishing Result relations than Occasion relations, the effect described by the second eventuality in a Result sequence is often a direct result of the end state brought about by the first, and thus it would perhaps not be surprising to find an end-state bias for Result relations as well. This notwithstanding, Stewart et al. (1998) show that verbs are highly variable with respect to their biases in Result relations.

Coherence Relation	Percentage of Corpus $P(CR)$	Source Bias $P(pronoun = source|CR)$
Occasion (171)	.38	.18
Elaboration (126)	.28	.98
Explanation (82)	.18	.80
Violated Expectation (38)	.08	.76
Result (25)	.06	.08

Table 1: Probabilities from Rohde et al. 2006 (Perfectives Only)

that elicited passage completions with the same stimuli as Rohde et al. (2006), but added an additional condition that bore only on the instructions. In one version, the participants were asked to have their completion answer the question *What happened next?*, whereas in the other they were to answer the question *Why?*. The idea was that this difference should give rise to different distributions of coherence relations; specifically, more Occasion relations (which, per Table 1, are Goal-biased) in the first case, and more Explanation relations (which are Source-biased) in the second. However, since the manipulation does not pertain to anything specific to pronoun interpretation, it should not affect the probabilities $P(pronoun = ante|CR)$. Note that whereas these experiments were conducted to test an informational, Hobbsian approach to coherence, the manner in which the distribution of coherence relations was manipulated was, essentially, by explicitly suggesting a QUD that the participants' completions should answer.

The outcome was as predicted. On the one hand, the biases toward particular pronoun referents conditioned on coherence relation (i.e., the probabilities $P(pronoun = ante|CR)$) were consistent between the two conditions as well as with Rohde et al. (2006). The distribution of coherence relations varied, however, and in the way predicted by equation (31), the overall pattern of pronoun interpretations shifted dramatically: there were significantly more Source interpretations in the *Why?* condition than the *What next?* condition.

This result is surprising on any theory of pronoun interpretation that is based predominantly on morphosyntactic cues or preferences, since the stimuli themselves were identical between conditions. Note that participants *could* have first assigned the pronoun using morphosyntactic preferences and then written a continuation that accommodated that assignment – it is always possible to answer either QUD with a sentence that begins by mentioning either referent. But apparently this is not what happened; instead the QUD that the participants were answering determined, at least in part, the way they assigned the pronoun.

Since the foregoing experiments were run only with stimuli that contained pronoun prompts, these results are not sufficient to establish that pronoun interpretation preferences are solely attributable to coherence establishment. Indeed, other results from Stevenson et al. (1994) are hard to reconcile with such an assumption. Stevenson et al. paired passages with a pronoun prompt as in (32) with variants that had no pronoun as in (34), in which participants chose their own forms of referring expressions.

(34) John handed a book to Bob. _____

They found that the choice of referring expression was heavily biased towards a pronoun when the referent was the previous subject, and likewise towards a name when the referent was a non-subject. However, the (non-subject) Goal was the next mentioned entity 74.4% of the time. (Arnold (2001) found similarly strong biases using similar contexts, including an 85.6% bias toward mentioning

the Goal next.) The data therefore suggest that the context sentence in (32) and (34) gives rise to a strong *next mention* bias toward the Goal, but that this bias is counteracted somewhat by a subject-oriented bias introduced by the pronoun in (32), which took the Goal bias down to 49%.

If pronoun interpretation is associated with a subject bias, this bias would be predicted to have repercussions for coherence establishment: Because the mere presence of a pronoun – even one whose reference is ambiguous as in (32) – would bias interpretation towards the subject of the last clause, encountering a pronoun would be expected to cause a shift in comprehenders' expectations in favor of the Source-biased coherence relations in Table 1, as compared to the completions for no-pronoun passages like (34) in which participants choose their own referring expressions. This prediction was confirmed in a study by Rohde and Kehler (2008a): significantly more instances of Source-biased coherence relations resulted in continuations with pronoun-prompt contexts like (32) as compared to no-pronoun contexts like (34).

Taken together, these results demonstrate not only that coherence establishment biases influence pronoun interpretation expectancies, but conversely that independent biases in pronoun production influence expectations about ensuing coherence relations. As such, whereas these experiments support the claim that coherence establishment is a critical factor for pronoun interpretation, the situation appears to be more complicated than the model proposed by Hobbs, in which pronouns are represented as variables that are ultimately bound solely as a by-product of pragmatic reasoning. Pronouns are not merely variables, but contribute their own biases as well.

The work described thus far in this section was designed to test a coherence-driven model of pronoun interpretation, posited in terms of an equation (shown in (31)) that captured hearers' expectations regarding the likelihood that different coherence relations might ensue. As we have already discussed, one experiment (Rohde, Kehler, and Elman 2007) demonstrated this using what could be characterized as a QUD-based manipulation, in which participants were asked to complete passages with pronoun prompts in a way that answered either the question *Why?* or *What happened next?*. The different distribution over coherence relations that resulted could therefore similarly be cast in terms of QUDs, that is, as a distribution over implicit questions that the hearer expects the next utterance to answer.

In light of this, Rohde and Kehler (2008b) carried out two experiments that asked whether biases toward particular questions in a dialog continuation condition would mimic analogous biases towards particular coherence relations in a story continuation condition. In the first experiment, the verb in the context sentence was varied between implicit causality (Garvey, Caramazza, and Yates 1976, Caramazza et al. 1977, McKoon, Greene, and Ratcliff 1993, Koornneef and van Berkum 2006, inter alia, henceforth IC) and non-IC types to see whether each type would generate the same percentage of Explanation relations in the story condition as 'Why?' type questions (e.g., 'Why?', 'How come?', and 'What for?') in the dialog condition.[7] Participants were instructed to imagine a conversation with a friend and write natural continuations that represented what the friend was likely to say next (story condition, ex. 35) or the question that they would be likely to pose to their friend (dialog continuation, ex. 36).

(35) John scolded$_{IC}$/saw$_{nonIC}$ Mary. _____

(36) a. Friend: John $_{IC}$/saw$_{nonIC}$ Mary.
    b. You: _____

---

[7] The use of two types of verbs was motivated by an experiment described in Kehler (2008), which demonstrated that sentences containing matrix IC verbs (*praise, amuse, criticize*) were much more likely to be followed by an Explanation continuation than sentences containing non-IC verbs.

Judges annotated the story and dialog continuations for coherence relations and QUD type respectively. Whereas significantly more Explanations were elicited in the IC condition than the non-IC condition, the percentage of Explanation relations in the story continuation condition was statistically indistinguishable from the percentage of 'Why?' type questions in the dialog condition for both verb types, as predicted.

To verify that the effect in Experiment 1 generalizes beyond Explanation and 'Why?' type questions, a second experiment used transfer-of-possession verbs in the context sentence varied by aspect (perfective/imperfective), which, as mentioned earlier, was shown by Rohde et al. (2006) to yield different distributions of coherence relations. The results confirmed that, for both verb types, the percentage of the questions falling into the 'Why?', 'What next?', and 'Where/when/how?' categories in the dialog condition was statistically indistinguishable from the percentage of Explanations, Occasions, and Elaborations in the story condition, as predicted. These results suggest that hearers' contextually-driven probabilistic expectations about what coherence relation will ensue extend to the explicit questions evoked in dialogs, as captured by the QUD analysis.

To sum, the experiments described in this section provide evidence that hearers implicitly track expectations about how the current discourse will be continued with respect to coherence, or alternatively, the implicit QUD that the continuation will answer. Although these expectations exist independently of the existence of linguistic phenomena like pronouns, they nonetheless have considerable impact on how these forms are interpreted. It is difficult to see how a model of pronoun interpretation could be formulated to account for this data without appealing to a sufficiently rich notion of discourse coherence.

This work also demonstrates more generally how the 'top-down' influences of expectations interact with evidence provided by linguistic material in determining discourse interpretations. Roberts (1998) and Büring (2003) discuss how focus marking, communicated in part via accentuation, conspire with Question-Answer Congruence to constrain the space of possible QUDs. The prompt-manipulation experiment of Rohde and Kehler (2008a) described above demonstrates a potentially more subtle effect, in which the occurrence of a pronoun – even an ambiguous one – can influence the expectations that comprehenders have for the QUD being addressed. Assuming that comprehenders maintain expectations not only about QUDs but also the larger strategies of inquiry that those QUDs are situated in, the prediction would then be that the mere occurrence of an ambiguous pronoun could create a shift in expectations at the highest level of reasoning about the speaker's goals and intentions. Carrying out a demonstration of this prediction awaits future work.

3   VP-ellipsis

We now turn to VP-ellipsis, illustrated in (37):

(37) John loves his mother, and Bill does too.

Following standard terminology, we will refer to the clause containing the site of the elided VP as the TARGET clause, and the clause that contains the antecedent as the SOURCE clause. As is well known, a bound pronoun in the source can give rise to ambiguities in the target; the target of example (37), for instance, can mean that Bill loves John's mother (the STRICT reading) or that Bill loves his own mother (the SLOPPY reading).

My focus in the following sections will be on determining the space of strict and sloppy readings such as these in more complex examples. I will first discuss the infamous Dahl puzzles in Sections 3.1–3.4. This work is joint with Daniel Büring, and was described more fully in Kehler

and Büring (2007). I will then discuss a certain class of interesting examples due to Hardt in Sections 3.5–3.6. That work was previously presented in Kehler (2007).

3.1 Dahl Puzzles

Missing readings puzzles, now a well-studied phenomenon in ellipsis, were first discussed by Dahl (1974) with respect to example (38).

(38) John thinks he loves his wife, and Bill does too.

Assuming that both pronouns are bound in the source clause, a naïve theory of strict and sloppy interpretations — in which each pronoun can be interpreted strictly or sloppily — predicts the four readings shown in (39a–39d). As Dahl noted, however, it only has three, missing the reading shown in (39d).

(39) a. Bill thinks John loves John's wife.   (all-strict: BJJ)
          $\lambda x.thinks(x, loves(John, wife(John)))$

       b. Bill thinks Bill loves Bill's wife.   (all-sloppy: BBB)
          $\lambda x.thinks(x, loves(x, wife(x)))$

       c. Bill thinks Bill loves John's wife.   (mixed sloppy–strict: BBJ)
          $\lambda x.thinks(x, loves(x, wife(John)))$

       d. # Bill thinks John loves Bill's wife.   (# mixed strict–sloppy: BJB)
          $\lambda x.thinks(x, loves(John, wife(x)))$

Meaning (39d) is not available as a possible interpretation despite the fact that it, like (39a–39c), generates the meaning of the source clause when applied to the source subject's referent, John (which we take to be the most obvious necessary, though clearly not sufficient, condition on VP ellipsis resolution).

3.2 The Standard Account, and Its Limitations

Various researchers have proposed systems to account for the missing readings facts, all relying on a constraint or mechanism that requires locality of binding (Kehler 1993, Fiengo and May 1994, Fox 2000, Schlenker 2005). I will focus on Fox's analysis as an exemplar, who implements locality of binding using Rule H:

(40) Rule H:
     A pronoun, γ, can be bound by an antecedent, α, only if there is no closer antecedent, ϕ, such that it is possible to bind γ to ϕ and *get the same semantic interpretation.* (p. 115)

Rule H has the effect of forcing transitive binding in a sentence like the source of (38), which offers two otherwise synonymous possible binding configurations:[8]

---

[8]The notation used is based on the system for semantic binding of Büring (2005). The general rule for binding is shown in (41), which is illustrated in the derivation shown in (42).

(41) $[\beta_i X]^g = \lambda x.[X]^{g[i \to x]}(x)$

(43) a. John β_1 thinks he$_1$ β_2 loves his$_2$ wife.

   b. * John β_1 thinks he$_1$ loves his$_1$ wife.

Because both options result in the same interpretation, Rule H dictates that the existence of (43a) renders (43b) ungrammatical.[9] It is worth noting that Rule H is TRANSDERIVATIONAL, in that the grammar has to 'know' about the possibility of the binding configuration in (43a) in ruling out the configuration in (43b).

   Rule H is not enough in itself to capture the missing readings data. Fox therefore introduces an additional constraint on parallel dependencies:

(44) *NP Parallelism* (Fox 2000)

   NPs in the antecedent and elided VPs must either

   a. have the same referential value (Referential Parallelism) or

   b. be linked by identical dependencies (Structural Parallelism)

Clause (44a) allows for pronouns to receive a strict interpretation, and (44b) allows for sloppy interpretations. As Fox notes, this constraint does not follow independently from any other grammatical principle, and hence it needs to be stipulated as an independent one.

   With these constraints in hand, the Dahl facts can be captured:

(45) John β_1 thinks he$_1$ β_2 loves his$_2$ wife.

   a. Bill thinks he$_j$ loves his^J wife. (BJJ)

   b. Bill β_1 thinks he$_1$ β_2 loves his$_2$ wife. (BBB)

   c. Bill β_1 thinks he$_1$ loves his^J wife. (BBJ)

   d. Bill thinks he^J β_1 loves his$_1$ wife. (BJJ again)

   e. # Bill β_1 thinks he^J loves his$_1$ wife. (BJB)

   There are four options with respect to pronoun interpretations: Each of the two pronouns can be interpreted following (44a) or (44b). The three available readings, shown in (45a–45c), are predicted to be possible. Case (45d), however, in which we interpret *he* using (44a) and *his* using (44b), does not result in the missing interpretation. Instead, because the second pronoun is bound to the first in the source, binding it in parallel in the target causes it to receive the 'strict' interpretation, and hence (45d) derives the same all-strict reading as (45a). There is thus no way to recover the missing reading in (45e); doing so would require that the second pronoun be bound directly to the matrix subject in the source, which was the configuration specifically blocked by Rule H.

---

(42) a. John β_1 loves his$_1$ mother.

   b. John $\in$ [β_1 loves his$_1$ mother]^g

   c. John $\in \lambda x.$[loves his$_1$ mother]$^{g[1 \rightarrow x]}(x)$

   d. John $\in \lambda x.x$ loves $g[1 \rightarrow x](1)$'s mother

   e. John loves John's mother.

[9]Büring (2005) presents a slight generalization of Rule H, called Have Local Binding! (or HLB!). Since Büring's HLB! achieves the same effect as Fox's in all examples considered here, we will only refer to Rule H in the discussion. All our comments and criticisms apply to both rules, however.

3.3 Two Observations

As we have seen, the state-of-the-art of our understanding of the Dahl puzzles requires a trans-derivational binding principle and a two-part definition of parallelism that does not appear to follow from anything else. In hopes of improving the situation, let us make two observations. The first one we already know from previous sections: sets of highly-parallel utterances are plausibly analyzed as providing partial answers to a common QUD. The utterances in (10b-c), for instance, repeated again in (46b-c), are naturally seen as each providing an answer to the (explicit or implicit) question in (46a):

(46) a. { *Who did what in the evening?* }
  b. Andy worked on a paper this evening.
  c. Jill spent the evening answering email.

When implicit, this question must be inferred based on the denotations of the 'answers', their intonational properties, and the context. As we have already seen, crucial to this is Question-Answer Congruence, the fact that the meaning of each answer must be included in the set of alternatives representing the question (Roberts 1998, Büring 2003).

The second and more novel observation is that utterances containing certain binding configurations lead to what we call DISJOINTNESS PRESUPPOSITIONS. Consider (47):

(47) a. Who$_1$ thinks that John loves his$_1$ wife?

  b. # John.

Informants report that as an answer to (47a) – crucially, under the interpretation indicated by the noted binding relations, i.e., *Which Xs think that John loves X's wife* – (47b) is strange. Many report the intuition that the respondant is being snarky, as if she should have known that John was not among the intended answers.[10] This intuition contrasts with example (48a), for which the answer (48b) is considerably more natural.

(48) a. Who$_1$ thinks that he$_1$ loves John's wife?

  b. John.

To draw the intuition out a little further, consider (49), which — again on the reading indicated — is judged to be true iff I hope to have John at my apartment and someone else hopes to have him at their apartment (perhaps John is a celebrity, in town for just a day):

(49) Not [only I$_F$]$_1$ am hoping that John will spend the day at my$_1$ apartment.

Crucially, it is not sufficient for the truth of (49) if I hope to have John over, John is hoping to spend the day at his own apartment, and no one else has any hopes about John visiting them. John, in short, is not among the possible alternatives to *I*. The pertinent conclusion, it seems to Büring and I, is that in both (47) and (49) the domains of *who* (possible answers to the question) and *only* (alternatives to the speaker) are presupposed to exclude John.

In Kehler and Büring (2007), we posited that this presupposition is triggered by the fact that a DP referring to John c-commands a pronoun bound by *who/only I*: *his/my* carry a presupposition of disjointness with John. In absence of a mechanism for deriving utterance interpretations that has this as a natural side-effect, we stated the following generalization:

---

[10]Of course, (47) is a perfectly well-formed Q/A sequence if the question is read as 'Who thinks that John loves John's wife', but this is not the reading of interest here (again, as indicated by the indexing in (47a)).

Be Bound or Be Disjoint! (BBBD!): If a pronoun p is free in the c-command domain of a (non-Wh) DP α, p bears a presupposition of disjointness with α (unless α itself binds p)

Under BBBD!, the meaning of example (47a), repeated below as (50a), would be represented as in (50b), which is in turn glossed as (50c).

(50) a. Who$_1$ thinks John loves his$_1$ wife?

 b. $\lambda p. \exists x [person(x) \wedge p = thinks(x, loves(John, wife(x))) \underline{\wedge \; x \not\subseteq John}]$

 c. the set of propositions 'x thinks that John loves x's wife' where x is a person other than John

The underlined part of (50b) is the relevant presupposition, to be read as 'x is disjoint with John'. This excludes the proposition that John thinks that he (himself) loves his (own) wife from the set of propositions denoted by (50a).

Since writing our paper and considering a broader range of examples, some doubt has emerged about whether this is the right explanation of the phenomenon. Thus I will not consider these particular implementation details further. The important thing to bear in mind for the ensuing section is that in questions such as (50a), the proposition denoted by *John thinks he loves his wife* is presupposed to not be in the set of possible answers.

3.4 The QUD Analysis Applied to Dahl's Puzzle

We have observed that sets of clauses bound by parallelism are plausibly analyzed as providing partial answers to a (generally inferred) QUD. Because the source and target clauses in the VPE examples we have been considering are bound by parallelism, we would therefore expect our observation to apply to them, possibly constraining their interpretation. More concretely, Question-Answer Congruence predicts that a source-target pair will only be felicitous under a particular interpretation if a suitable QUD can be inferred to which the source and target each provide partial answers.

We claim that this constraint, combined with the disjointness presuppositions created by BBBD!, explains the Dahl puzzle. We illustrate by stepping through the different readings for example (38). Take the all-strict reading (39a): John thinks John loves John's wife, and Bill thinks John loves John's wife. We are now looking for one question that has both of these propositions as answers. That question is 'Who thinks that John loves his (John's) wife?', as represented in 51:

(51)   { Who thinks that John β_1 loves his$_1$ wife? }

 a. John β_2 thinks that he$_2$ β_3 loves his$_3$ wife, and

 b. Bill does [think that John β_4 loves his$_4$ wife] too.

Because *John thinks that John loves John's wife* (51a) and *Bill thinks that John loves John's wife* (51b) are both in the set denoted by *Who thinks that John loves John's wife*, this question can be accommodated as a QUD to generate the all-strict interpretation of (38).

Example (52) shows the QUD that licenses the all-sloppy reading in (39b):

(52)   { Who β_1 thinks that he$_1$ β_2 loves his$_2$ wife? }

 a. John β_3 thinks that he$_3$ β_4 loves his$_4$ wife, and

 b. Bill does [β_5 think that he$_5$ β_6 loves his$_6$ wife] too.

Again, the Question-Answer Congruence condition is met. The same is true for (39c): (53) shows the QUD that licenses the acceptable mixed reading:

(53)   { Who β_1 thinks that he$_1$ loves John's wife? }

    a. John β_2 thinks that he$_2$ β_3 loves his$_3$ wife, and

    b. Bill does [β_4 think that he$_4$ loves John's wife] too.

Things go wrong, however, when it comes to the QUD that would be needed to license the unattested mixed reading in (39d). The two 'answers' — that John thinks John loves John's wife, and that Bill thinks John loves Bill's wife — point to the question *Who thinks that John loves his wife?* (with *his* bound by *who*):

(54)   { Who β_1 thinks that John loves his$_1$ wife? (and his$_1$ $\not\phi$ John) }

    a. # John β_2 thinks that he$_2$ β_3 loves his$_3$ wife, and

    b. Bill does [β_4 think that John loves his$_4$ wife] too.

But as we have argued, this QUD carries the disjointness presupposition that *his* is not coreferential with John. Whereas the question itself is perfectly felicitous, the proposition *John thinks that John loves John's wife* is not in the answer set it denotes. As such, (54a) does not provide an answer to this question, and the passage is infelicitous under this reading.[11]

To sum, the two observations we made in Section 3.3 combine to explain the Dahl puzzle. (See Kehler and Büring (2007) for a discussion of a broader range of examples.) Unlike previous analyses, the QUD analysis derives the facts without positing a transderivational locality of binding constraint nor an overlaid parallelism constraint, distinguishing it (so far as we know) from all previous approaches.

3.5 Illusory Sloppy Readings

In the large majority of previous approaches to VPE, the anaphoric relations in the target clause are interpreted uniformly from those in the source. We can call such accounts *parallelism-based* analyses (Sag 1976, Williams 1977, Gawron and Peters 1990, Prüst, Scha, and van den Berg 1991, Dalrymple, Shieber, and Pereira 1991, Kehler 1993, Hobbs and Kehler 1997, inter alia). The QUD-based analysis we outlined in the previous section offers a variation on this theme, in which parallelism isn't enforced between the source and target directly, but between each clause and the

---

[11]The foregoing analysis was not entirely explicit about the licensing conditions for eliding VPs in target clauses. For current purposes, we are assuming that ellipsis is possible if the source and target clause denotations are both in the alternative set denoted by the question, i.e., if Question-Answer Congruence is satisfied. Ultimately, one would like the information-structural constraints on VP-ellipsis to reduce to those on deaccenting VPs, and indeed, the variant of (38) in which the target VP is overt but deaccented has only the same three readings as the elided version. However, of the readings analyzed above – in which only binding configurations that obeyed Reinhart's (1983) Coreference Rule were posited – only the target VP for the all-sloppy reading (52) is Given by the source. (While the VP in readings (52a-c) are all Given with respect to the implicit QUD, recall from Section 2.1 that we followed Büring (2003) in noting that implicit QUDs cannot serve to introduce a new element and make it count as Given for utterances that follow.) The problem is actually quite general; even certain simple cases of felicitous ellipsis involve VPs whose interpretations are not Given on any theory that we are aware of, e.g.:

(55) a. John defended Bill because

     b. HE WOULDn't [β_1 defend himself$_1$].

VP-ellipsis is possible here despite the fact that $\lambda x.defend(x,x)$ is not Given. I leave this issue for further research.

inferred QUD to which they are understood to provide partial answers. Nonetheless, common to all of these approaches is that the space of readings for 'reconstructed' pronoun representations in the target are determined in part by some type of parallelism constraint.

In a series of papers, Hardt (1992a, 1992b, 1993, 1999, inter alia) offered a variety of examples of sloppy interpretations under VP-ellipsis that challenged the predictions made by the classic parallelism-based theories, specifically those of Sag (1976) and Williams (1977). Most of these examples, however, are compatible with theories based on a more extended notion of parallelism between antecedent and elided clauses; see, e.g. Dalrymple et al. (1991), Prust et al. (1994), Hobbs and Kehler (1997), inter alia; other examples receive a natural treatment as cases of implicational bridging in Rooth's (1993) analysis. One type of example remains, however, exemplified in (56), which I believe poses a greater challenge for any analysis based on either syntactic or semantic parallelism.

(56) Every boy in Mrs. Smith's class hoped she would pass him. In John's case, I think she will.

The elided clause in this example clearly admits of an apparently sloppy interpretation, that is, *I think she actually will pass him$_j$*. Yet this reading arises without the existence of the type of syntactic or semantic parallelism between the clauses that almost all theories require. While this example presents a clear adequacy criterion for any theory of strict and sloppy interpretations in VP-ellipsis, discussion of such examples are conspicuously absent in much literature on the topic.[12]

Hardt relied on such examples to argue that the interpretations of pronouns in VP representations that are copied from source to target clause representations are not bound by parallelism constraints, but instead are reinterpreted in the target clause using the same discourse principles that apply to overt pronouns. (I will henceforth refer to this as the 'free interpretation hypothesis', or FIH.) For instance, in (56), the semantic representation of the VP *pass him* would be copied to the target representation with the pronoun unresolved, which would then be resolved to *John* via ordinary pronoun interpretation mechanisms, on analogy with the unelided version in (57).

(57) Every boy in John's class hoped she would pass him. In John's case, I think she will pass him.

Kehler and Shieber (1997) subsequently argued against the FIH, claiming that it can only be salvaged if it is assumed that the unelided versions of reconstructed VPs must be fully deaccented. They considered minimal pairs such as (58).

(58) a. Every boy$_i$ was hoping that Mary would ask him$_i$ out, but the waiting is over. Last night at Bob's$_j$ party, she asked him$_j$ out.

b. # Every boy$_i$ was hoping that Mary would ask him$_i$ out, but the waiting is over. Last night at Bob's$_j$ party, she did.

The reading where Mary asked out Bob at Bob's party, while readily available in example (58a), is not available in its elided counterpart (58b).

At first glance, the difference between (58a-b) might be taken to falsify the FIH outright. There is, however, an out: this interpretation appears to require accent within the VP in (58a), which in this case, would fall on the pronoun. Indeed, standard theories of focus marking and accent placement would predict that accent is required; whereas both the verb meaning (*ask X out*) and the object NP meaning *Bob* represent Given information, the result of combining the two (*ask Bob*

---

[12]Fiengo and May (1994, p. 111, fn. 14) being an exception.

out) is not. On theories such as Schwarzschild (1999), inter alia, the VP will therefore require F-marking, which in turn requires that either the verb or object pronoun receive accent. As shown by examples (59a-c), in such cases accent will typically fall on the pronoun:

(59) a. # Isaac likes Mary, and Sue likes Isaac/him.

　　 b. # Isaac likes Mary, and Sue LIKES Isaac/him.

　　 c. Isaac likes Mary, and Sue likes ISAAC/HIM.

The FIH can then be salvaged under the standard assumption that only deaccented VPs can be elided, since, by their nature, elided VPs cannot carry accent. In that case, however, all we have done is reduce the problem of identifying Given predicates that license VP ellipsis to a similar problem, namely, the identification of Given predicates that license the corresponding deaccented VPs (Tancredi 1992, Rooth 1992, inter alia), in which pronoun interpretation will be more restricted than in the general case.[13] So then real mystery is why the VP can be deaccented in the unelided version of (56), shown in (57), but not in the unelided version of (58b), i.e., (58a).

3.6　Ellipsis and Contrastive Topics

The solution to this mystery requires that we take another look at the structure of QUDs in discourses characterized by parallelism. Jackendoff (1972) documented how multiple foci within a single clause may utilize two different intonation contours, an "A contour", which marks the DEPENDENT FOCUS (which we'll refer to simply as the FOCUS), and a "B contour", which marks the INDEPENDENT FOCUS (which we'll refer to as the CONTRASTIVE TOPIC).[14] These accents interact with the discourse context to give rise to implicit questions that structure the discourse. Jackendoff's example goes as follows. Assume a context in which there was a gathering at which there were a number of people and a number of different things to eat, and that various people ate various things. Speaker A in the discourse is asking questions of the form *Who ate what?*. This speaker can break down the answers in one of two ways: by person or by food. Answering by person requires the intonation pattern in (60b), which is compatible with the subquestions given in (60a).

(60) a. { Well, what about FRED? What did HE eat? }

　　 b. $FRED_B$ ate the $BEANS_A$.

Alternatively, answering by food requires the intonation pattern in (61b), which is compatible with the subquestions given in (61a).

(61) a. { Well, what about the BEANS? Who ate THEM? }
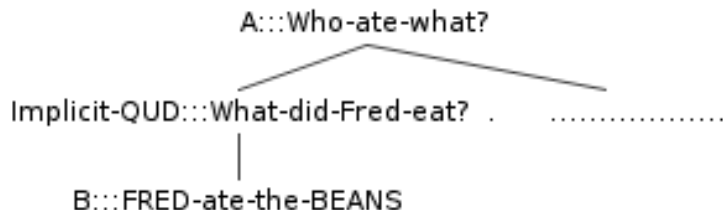
　　 b. $FRED_A$ ate the $BEANS_B$.

Note that these intonation patterns and contexts cannot be switched: (61b) cannot be used to answer (60a), and similarly for (60b) and (61a). Likewise, a list of partial answers to *Who ate what?* could not freely alternate between the two patterns.

Roberts (1998) extends this insight within a more formal analysis. In her framework, B-accented Contrastive Topics *presuppose* a particular Question-Under-Discussion (QUD) – indeed, possibly

---

[13]Here I set aside the fact that deaccented VPs, in containing overt linguistic material, allow for a greater range of *inferred* sources than elided VPs, e.g., *John called Bill a republican and MARY insulted him TOO.* (Lakoff 1971)

[14]Roberts (1998) characterizes both contours as involving an L+H* pitch accent followed by an L- phrase accent, with an L boundary tone for A-contours and an H boundary tone for B-contours.

a complex question/subquestion strategy – which may need to be accommodated (see also Bur-ing (2003)). A strategy for (60) would look like:

```
                    A:::Who-ate-what?
                   ╱           ╲
Implicit-QUD:::What-did-Fred-eat?  .    .................
                   │
            B:::FRED-ate-the-BEANS
```

The intuition becomes particularly clear when considering discourses in which a B-accent is used unexpectedly, with the effect of altering the strategy of inquiry. Robert offers example (62):

(62) A. [When are you going to China]$_F$?　　(= Roberts ex. 47)

　　B. Well, I'm going to CHINA$_B$ in APRIL$_A$.

　　A. Oh? Where else are you going, and when?

A's first utterance offers an explicit QUD for B to respond to. Strictly speaking, B's response is an infelicitous answer, since a B-accent is placed on China, which is already Given by the QUD. To have B's response be made felicitous, A must accommodate the new, more general QUD in (63):

(63) [QUD: When are you going to which place?]

which is then placed atop a more complex discourse structure:

```
        Implicit-QUD:::When-are-you-going-to-which-place?
                   ╱                        ╲
A:::When-are-you-going-to-China?    A:::Oh?-Where-else-are-you-going-and-when?
                   │
    B:::Well-I'm-going-to-CHINA-in-APRIL
```
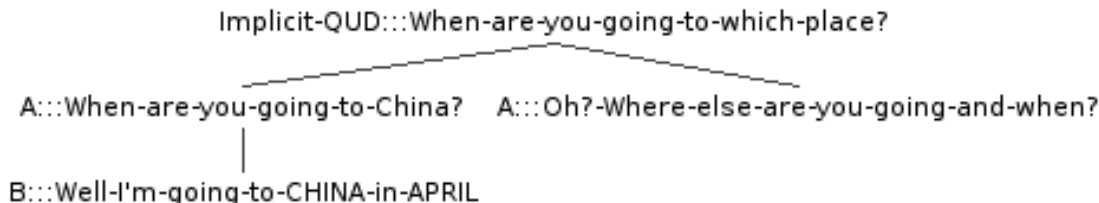
This leads to the implication that B's response is only a partial answer to the superquestion in (63), leading A to assume there are other places that B plans to visit, hence her follow-up question.

Returning to the VP-ellipsis examples at hand, Jackendoff (1972) notes that phrases such as *As for X* (to which I add *in X's case, regarding X*, etc.) are usually intoned with a B-accent on X:

> ...the idea of the B-accented focus as topic accords with the intonation of preposed (or "topicalized") phrases, which almost always receive a B accent. [...]

> As for FRED$_B$, I don't think HE$_B$ can MAKE$_A$ it. (p. 263)

Combining the above insights leads me to the following analysis: that the phrase *In X's case* (and others in its class) introduces a QUD created by substituting X for a parallel entity within the meaning of an anaphorically-identified, contextually-salient referent. The remainder of the clause answers (and is intoned with respect to) this QUD, and both question and answer are situated in a more complex QUD structure akin to that for example (62).

For (56), repeated below as (64), *In John's case* (with B-accent on *John*) takes the embedded clause in (64a) as its antecedent, with John being parallel to the bound pronoun. The result is represented in (65), with the QUD shown in (65b):

24

(64) a. Every boy in Mrs. Smith's class hoped she would pass him.

    b. In John$_j$'s$_B$ case, I think she will$_F$. [pass him$_j$]

(65) a. Every boy in Mrs. Smith's class hoped she would pass him.

    b. *Will Mrs. Smith pass John?*

    c. (Yes,) I think she will$_F$. [pass him$_j$]

The treatment of example (64) is now straightforward: as indicated above, the remainder of the clause in (64b), i.e., (65c), answers the QUD in (65b). As such, *the QUD (65b) is the source clause for the ellipsis, not (65a)*. As such, the interpretation in question is actually a *strict* reading. There is no re-binding of a pronoun involved. The difference between this case and (58) is that *Last night at Bob's party* in (58) is simply a run-of-the-mill adverbial; it does not contain a B-accent, nor introduce a QUD as *In John's case* does.

Further evidence for the QUD semantics for *In X's case* is provided by the fact that phrases that commonly serve as short answers to questions – e.g., *forget it, no way, no chance*, and *PLEEEAAASSSE* – serve as natural follow-ons:

(66) I think Mrs. Smith will pass most of the students in the class. But in John's case, forget it / no way / no chance / PLEEEAAASSSE!

If *in John's case* introduces a QUD of the form shown in (65b), the felicity of such phrases is entirely expected.

Also expected is the fact that the relevant facts remain constant with a variety of other event referential expressions which do not involve ellipsis.

(67)   Every boy in Mrs. Smith's class hoped she would pass him.

    a. In John's case, I'm sure she'll do it.

    b. In John's case, I think it's inevitable.

    c. In John's case, I'm optimistic.

These data demonstrate that any analysis of (64) in which the facts result specifically from the reconstruction of elided material misses an important generalization.

Finally, consider the following two variants of (64), shown in (68–69):

(68) I think Mrs. Smith will pass most of the students in the class. In John's case, however, I don't think she will.

(69) George and John are both worried about their grades. I'm almost certain that Mrs. Smith will pass George. In John's case, however, I don't think she will.

A strong majority of informants find that the elided clauses in these passages have the same "sloppy" interpretations – i.e., *I don't think she will pass John* – for reasons that are intuitively parallel to those in (64). Yet in neither of these examples does the source clause for the ellipsis contain a pronoun. Instead, John is parallel to the direct object of the antecedent clause: in (68) John is parallel to *most of the students* (or the variable left behind if the quantifier is raised), and in (69) John is parallel to *George*. But with this realization, we see that the pronoun in (64) has nothing to do with the ability to get (what appears to be) a sloppy interpretation: because John is parallel

to the pronoun *him*, this pronoun representation is not copied to the target in the way required for sloppy readings. Presumably, any analysis that captures the facts for (64) should apply equally to (68–69). The analysis presented here does, whereas the FIH does not.

To sum, examples like (64) are interesting, but not because they provide support for the FIH. Instead, the mystery associated with their readings is why the VPs can be deaccented in such cases, but not in cases such as (58). As I have argued, the facts are predicted from more general properties of contrastive topic phrases and their role in evoking QUDs in discourse. The examples are also interesting in that they add to the current suite of evidence against approaches to VP-ellipsis that are based on a syntactic reconstruction mechanism, since the required antecedent on such theories – *pass John* – does not occur as a syntactic VP in the contexts of any of these examples.

These data also fit in well with the QUD-based analysis of Dahl's puzzles described in Sections 3.1–3.4, moving us toward a more comprehensive analysis of strict and sloppy interpretations in a QUD framework. Not covered here, however, is another potential application of QUDs to ellipsis: the reconcilation of mysteries involving mismatches in syntactic form. In Kehler (2000, 2002), for instance, I argued that source and target clauses participating in Parallel relations require matching syntactic form (70–71):

(70) # This problem was looked into by John, and Bob did too. [look into the problem]

(71) # This theory was expressed using SDRSs by Smith, and Jones did too. [express this theory using SDRSs]

whereas those participating in Cause-Effect (here, Violated Expectation) relations do not (72–73):

(72) This problem was to have been looked into, but obviously nobody did. [look into the problem] (Vincent Della Pietra, in conversation)

(73) Of course this theory could be expressed using SDRSs, but for the sake of simplicity we have chosen not to. [express this theory using SDRSs] (from text of Lascarides and Asher (1993))

Since these sets of examples are related by different QUD structures, it stands to reason that an analysis in terms of QUDs instead of coherence relations might be possible. I will not pursue this here, but see Kertz (2008) for arguments and experiments that point to an information structural explanation.

4 Conclusions

We have surveyed several recent results pertaining to ellipsis and anaphora interpretation, analyzed from the purview of a QUD model of discourse. These analyses suggest that the QUD model can be used to explain phenomena that have previously required ad-hoc or untenable principles, including the parallel grammatical role preference in pronoun interpretation, constraints on local binding and NP parallelism to explain the Dahl puzzles in VP ellipsis interpretation, and the Free Interpretation Hypothesis to explain illusory sloppy readings in VPE. Further, several experimental studies suggested that comprehenders not only infer latent QUDs, but also have anticipatory expectations about what QUD the next utterance will answer, and that these expectations in turn affect how anaphoric forms are processed. These results motivate extensions to dynamic theories of QUD-structured discourse interpretation to not only represent the space of possible implicit questions, but also probability distributions over them (and for that matter, their possible answers),

and ultimately over the space of strategies of inquiry of which those questions are components. One lesson of this work, it is hoped, is that it is folly to attempt to study discourse-dependent linguistic phenomena – including not only focus and accent placement, but also ellipsis and anaphora – without a suitably rich model of discourse coherence establishment firmly in hand. In light of the inherent discourse-dependence of these phenomena, it would have indeed been surprising if it had turned out any other way.

References

Akmajian, Adrian and Ray S. Jackendoff. 1970. Coreferentiality and stress. *Linguistic Inquiry*, 1(1):124–126.

Allen, James F. and C. Raymond Perrault. 1980. Analyzing intention in utterances. *Artificial Intelligence*, 15(3):143–178. Reprinted in Grosz et al. 1986: 441-458.

Arnold, Jennifer. 2001. The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, 21(2):137–162.

Asher, Nicholas and Alex Lascarides. 1994. Intentions and information in discourse. In *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics (ACL-94)*, pages 34–41, Las Cruces, June.

Asher, Nicholas and Alex Lascarides. 2003. *Logics of Conversation*. Cambridge University Press, Cambridge.

Büring, Daniel. 2003. On D-trees, beans, and B-accents. *Linguistics and Philosophy*, 26:511–545.

Büring, Daniel. 2005. *Binding Theory*. Cambidge University Press, Cambridge.

Caramazza, A., E. Grober, C. Garvey, and J. Yates. 1977. Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behaviour*, 16:601–609.

Carlson, Lauri. 1983. *Dialogue Games: An Approach to Discourse Analysis*. Reidel, Dordrecht.

Cohen, Philip R., Jerry Morgan, and Martha E. Pollack. 1990. Introduction. In Philip R. Cohen, Jerry Morgan, and Martha E. Pollack, editors, *Intentions in Communication*. Bradford, MIT Press, pages 1–13.

Cohen, Philip R. and C. Raymond Perrault. 1979. Elements of a plan-based theory of speech acts. *Cognitive Science*, 3:177–212. Reprinted in Grosz et al. 1986: 423-440.

Crawley, R. A., R. J. Stevenson, and D. Kleinman. 1990. The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, 19:245–264.

Dahl, Östen. 1974. How to open a sentence: Abstraction in natural language. Technical report, University of Göteborg. Logical Grammar Reports, No. 12.

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

de Hoop, Helen. 2004. On the interpretation of stressed pronouns. In Reinhard Blutner and Henk Zeevat, editors, *Optimality Theory and Pragmatics*. Palgrave Macmillan, New York, pages 25–41.

Fiengo, Robert and Robert May. 1994. *Indices and Identity*. MIT Press, Cambridge, MA. Linguistic Inquiry Monograph Twenty-Four.

Fox, Danny. 2000. *Economy and Semantic Interpretation*. MIT Press, Cambridge, MA.

Garvey, C., A. Caramazza, and J. Yates. 1976. Factors underlying assignment of pronoun antecedents. *Cognition*, 3:227–243.

Gawron, Mark and Stanley Peters. 1990. *Anaphora and Quantification in Situation Semantics*. CSLI/University of Chicago Press, Stanford University. CSLI Lecture Notes, Number 19.

Ginzburg, Jonathan. 1996. Dynamics and the semantics of dialog. In J. Seligman, editor, *Language, Logic and Computation, Volume 1*. CSLI, Stanford.

Gordon, Peter C., Barbara J. Grosz, and Laura A. Gilliom. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, 17(3):311–347.

Gordon, Peter C. and Kimberly A. Scearce. 1995. Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory and Cognition*, 23(3):313–323.

Grosz, Barbara J. and Candace L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

Hale, John. 2001. A probabilistic earley parser as a psycholinguistic model. In *Proceedings of the Annual Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 159–166.

Hardt, Daniel. 1992a. Verb phrase ellipsis and semantic identity. In *Semantics and Linguistic theory II*, Ohio State University, May.

Hardt, Daniel. 1992b. VP ellipsis and contextual interpretation. In *Proceedings of the International Conference on Computational Linguistics (COLING-92)*, pages 303–309, Nantes, July.

Hardt, Daniel. 1993. *Verb Phrase Ellipsis: Form, Meaning, and Processing*. Ph.D. thesis, University of Pennsylvania.

Hardt, Daniel. 1999. Dynamic interpretation of verb phrase ellipsis. *Linguistics and Philosophy*, 22(2):187–221.

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

Hobbs, Jerry R. 1990. *Literature and Cognition*. CSLI Lecture Notes 21, Stanford, CA.

Hobbs, Jerry R. 1997. On the relation between the informational and intentional perspectives on discourse. In Eduard H. Hovy and Donia R. Scott, editors, *Computational and Conversational Discourse*. Springer-Verlag, Berlin, pages 139–157.

Hobbs, Jerry R. and Andrew Kehler. 1997. A theory of parallelism and the case of VP ellipsis. In *Proceedings of the 35th Conference of the Association for Computational Linguistics (ACL-97)*, pages 394–401, Madrid, Spain, July.

Hobbs, Jerry R., Mark E. Stickel, Douglas E. Appelt, and Paul Martin. 1993. Interpretation as abduction. *Artificial Intelligence*, 63:69–142.

Jackendoff, Ray S. 1972. *Semantic Interpretation in Generative Grammar*. MIT Press.

Kehler, Andrew. 1993. A discourse copying algorithm for ellipsis and anaphora resolution. In *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics (EACL-93)*, pages 203–212, Utrecht, the Netherlands, April.

Kehler, Andrew. 2000. Coherence and the resolution of ellipsis. *Linguistics and Philosophy*, 23(6):533–575.

Kehler, Andrew. 2002. *Coherence, Reference, and the Theory of Grammar*. CSLI Publications.

Kehler, Andrew. 2007. Contrastive topics and illusory sloppy interpretations in VP-ellipsis. Presented at the Annual Meeting of the Linguistic Society of America, Los Angeles, January.

Kehler, Andrew and Daniel Büring. 2007. Be bound or be disjoint! In *Proceedings of the 38th Meeting of the North East Linguistics Society (NELS-38)*, Ottawa, Canada.

Kehler, Andrew, Laura Kertz, Hannah Rohde, and Jeffrey L. Elman. 2008. Coherence and coreference revisited. *Journal of Semantics (Special Issue on Processing Meaning)*, 25(1):1–44.

Kehler, Andrew and Stuart Shieber. 1997. Anaphoric dependencies in ellipsis. *Computational Linguistics*, 23(3):457–466.

Kertz, Laura. 2008. Focus structure and acceptability in verb phrase ellipsis. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, pages 283–291.

Kertz, Laura, Andrew Kehler, and Jeffrey L. Elman. 2006. Grammatical and coherence-based factors in pronoun interpretation. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, Vancouver, Canada.

Koornneef, Arnout W. and Jos J. A. van Berkum. 2006. On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye-tracking. *Journal of Memory and Language*, 54:445–465.

Ladd, D. Robert. 1980. *The Structure of Intonational Meaning*. Indiana University Press, Bloomington.

Lakoff, George. 1971. The role of deduction in grammar. In Charles J. Fillmore and D. Terence Langendoen, editors, *Studies in Linguistic Semantics*. Holt, Rinehart, and Winston, New York.

Lascarides, Alex and Nicholas Asher. 1993. Temporal interpretation, discourse relations, and common sense entailment. *Linguistics and Philosophy*, 16(5):437–493.

Levy, Roger. 2008. Expectation-based syntactic comprehension. *Cognition*, 106(3):1126–1177.

Mann, William C. and Sandra A. Thompson. 1987. Rhetorical structure theory: A theory of text organization. Technical Report RS-87-190, Information Sciences Institute, University of Southern California, Marina del Rey, California.

McKoon, Gail, Steven B. Greene, and Roger Ratcliff. 1993. Discourse models, pronoun resolution, and the implicit causality of verbs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18:266–283.

Moens, Marc and Mark Steedman. 1988. Temporal ontology and temporal reference. *Computational Linguistics*, 14(2):15–28.

Moore, Johanna D. and Cecile L. Paris. 1993. Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational Linguistics*, 19(4):651–694.

Moore, Johanna D. and Martha E. Pollack. 1992. A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics*, 18(4):537–544.

Oehrle, Richard T. 1981. Common problems in the theory of anaphora and the theory of discourse. In Herman Parret, Marina Sbisà, and Jef Verschueren, editors, *Possibilities and Limitations of Pragmatics*. John Benjamins, Amsterdam, pages 509–530. Studies in Language Companion Series, Volume 7.

Prüst, Hub, Remko Scha, and Martin van den Berg. 1991. A formal discourse grammar tackling verb phrase anaphora. University of Amsterdam manuscript.

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

Reinhart, Tanya. 1983. *Anaphora and Semantic Interpretation*. University of Chicago Press, Chicago.

Roberts, Craige. 1998. Information structure in discourse: Towards and integrated formal theory of pragmatics. In J. H. Yoon and Andreas Kathol, editors, *OSU Working Papers in Linguistics 49: Papers in Semantics (1996)*. Ohio State University, pages 91–136. Revised version as of Sept. 1998.

Rohde, Hannah and Andrew Kehler. 2008a. The bidirectional influence between coherence establishment and pronoun interpretation. Poster presented at the 21st Annual CUNY conference on Human Sentence Processing, March.

Rohde, Hannah and Andrew Kehler. 2008b. Coherence-driven expectations in discourse and dialog. Poster presented at the 21st Annual CUNY conference on Human Sentence Processing, March.

Rohde, Hannah, Andrew Kehler, and Jeffrey L. Elman. 2006. Event structure and discourse coherence biases in pronoun interpretation. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, pages 697–702, Vancouver, Canada.

Rohde, Hannah, Andrew Kehler, and Jeffrey L. Elman. 2007. Pronoun interpretation as a side effect of discourse coherence. In *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, pages 617–622, Memphis, TN.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics*, 1:75–116.

Rooth, Mats. 1993. Ellipsis redundancy and reduction redundancy. In S. Berman and A. Hestvik, editors, *Proceedings of the Stuttgart Workshop on Ellipsis: Arbeitspapiere des Sonderforschungsbereich 340, Bericht Nr. 29-1992, SFB 340*. University of Stuttgart, University of Tuebingen, and IBM Germany.

Sag, Ivan. 1976. *Deletion and Logical Form*. Ph.D. thesis, Massachusetts Institute of Technology.

Schlenker, Philippe. 2005. Non-redundancy: Towards a semantic reinterpretation of binding theory. *Natural Language Semantics*, 13:1–92.

Schwarzschild, Roger. 1999. Givenness, AvoidF, and other constraints on the placement of accent. *Natural Language Semantics*, 7:141–177.

Selkirk, Elisabeth O. 1996. Sentence prosody: Intonation, stress and phrasing. In John A. Goldsmith, editor, *The Handbook of Phonological Theory*. Blackwell, London.

Smyth, R. 1994. Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research*, 23:197–229.

Stalnaker, Robert. 1979. Assertion. In Peter Cole, editor, *Syntax and Semantics 9: Pragmatics*. Academic Press, New York, pages 315–332.

Stevenson, Rosemary J., Rosalind A. Crawley, and David Kleinman. 1994. Thematic roles, focus, and the representation of events. *Language and Cognitive Processes*, 9:519–548.

Stewart, Andrew J., Martin J. Pickering, and Anthony J. Sanford. 1998. Implicit consequentiality. In *Proceedings of the 20th Annual Conference of the Cognitive Science Society*, pages 1031–1036.

Tancredi, Christopher. 1992. *Deletion, Deaccenting, and Presupposition*. Ph.D. thesis, Massachusetts Institute of Technology. Distributed by MIT Working Papers in Linguistics.

van Kuppevelt, Jan. 1995. Discourse structure, topicality, and questioning. *Journal of Linguistics*, 31:109–147.

Venditti, Jennifer J., Matthew Stone, Preetham Nanda, and Paul Tepper. 2002. Discourse constraints on the interpretation of nuclear-accented pronouns. In *Proceedings of the 2002 International Conference on Speech Prosody*, Aix-en-Provence, France.

Williams, Edwin. 1977. Discourse and logical form. *Linguistic Inquiry*, 8(1):101–139.

Winograd, Terry. 1972. *Understanding Natural Language*. Academic press, New York.

Wolf, Florian, Edward Gibson, and Timothy Desmet. 2004. Discourse coherence and pronoun interpretation. *Language and Cognitive Processes*, 19(6):665–675.