

May, 2022

Coherence, Saliency and Anaphora: The role of the QUD
Craig Roberts, The Ohio State University and New York University

Anaphora is simple, but its resolution is complicated because discourse is complex.

I. Introduction¹

It has long been recognized that anaphora is a probe for linguistic structure. In the case of Binding Theory, anaphoric potential is a probe for syntactic structure: in order to yield a bound interpretation, an anaphoric trigger and its antecedent must stand in a particular relationship in the syntactic structure in which they occur (see Büring 2005 for overview and discussion).² Kamp's (1981) Discourse Representation Theory, Heim's (1982) Context Change Semantics and other theories of dynamic semantics show how anaphora can be used to probe the logical structure of a discourse. And coherence theories like those of Kehler (2002) and Asher & Lascarides (2003) argue that coherence relations help to establish another type of structure over the utterances in a discourse, with anaphora constrained by the resulting structure. An anaphoric trigger, such as a pronoun, must stand in certain type of structural relationship to a potential antecedent in order for the anaphora to be resolved to that antecedent.

I would argue that the structures in question aren't strictly linguistic in character or function. They reflect more general constraints on how human interlocutors structure shared information in an on-going exchange, with a view to addressing issues of joint interest. Underlying this is what it means to be a rational agent engaged in cooperatively pursuing joint goals with other rational agents. In this respect, the theory I will propose, along with other theories based on the notion of the Question Under Discussion (Roberts 1996, Ginzburg 2012), are Gricean in spirit. Not all approaches to the study of discourse are of this character.

Many contemporary approaches to the study of discourse emphasize the role of coherence relations and more generally, the rhetorical structure of discourse—Mann & Thompson, Asher & Lascarides. These accounts do not always make clear the respect in which these relations and the structures they represent are related to Gricean theory.

¹ This paper was originally given at a workshop on coherence and anaphora at NASSLLI 2016 at Rutgers University. Thanks to the workshop organizers, Sam Cumming and Daniel Altshuler, and to the great audience there, for their stimulating feedback. This work was presented in preliminary form in the OSU Pragmatics Group in 2015, and I am especially grateful to Micha Elsner, Marie-Catherine DeMarneffe and other participants in the Group for their very useful comments and criticisms. And it was presented at the University of São Carlos, Brazil, where the audience also offered useful feedback. This work was supported in part by NSF Grant #0952571, awarded to David Beaver, Craig Roberts, Mandy Simons and Judith Tonhauser; and from NSF Grant ##1452674, 2015-2020, also to Beaver, Roberts, Simons & Tonhauser.

² Schlenker (2005) argues that Binding Theory is best understood in largely semantic terms; cf. the earlier argument due to Reinhart (1983) that Principle C is pragmatic. But this issue is independent of the claim that anaphora is an excellent probe for syntactic structure, since compositional semantics is a function of that structure.

Claim:

Anaphora is simple, but its resolution is complicated by virtue of the fact that it is constrained by the structure of the discourse in which the trigger occurs. Discourse itself may be very complexly structured, bringing to bear multiple simultaneous constraints on interpretation.

An anaphoric trigger must stand in a certain type of relationship in discourse to a potential antecedent in order for the anaphora to be resolved to that antecedent. This generalization is intended to cover a wide range of anaphoric triggers, including:

- null anaphora
- ellipsis
- pro-forms
- demonstratives and demonstrative descriptions
- definite descriptions
- proper names

Over the years (see Roberts 1989, 2002, 2003, 2004b, 2010, 2015, 2019), I've developed an integrated view of anaphora in which anaphora resolution is constrained as follows across all types of trigger:

- (1) Anaphora requires an antecedent (discourse referent) which
 - a) is **weakly familiar**, and hence logically accessible—entailed to exist in its local context of interpretation;
 - b) yields a contextually **plausible** interpretation of the utterance in which it occurs, and hence one which is
 - consistent with the trigger's descriptive content and what is predicated of it, with common sense and with the interlocutors' common ground, and
 - coherent with the rest of the discourse;and
 - c) is the uniquely most **salient** discourse referent satisfying constraints (a) and (b)

Weak familiarity (1a) builds on the familiar logical constraints on accessibility built into Kamp's Discourse Representation Theory and Heim's Context Change Semantics.³ It requires that the interlocutors share a discourse referent for the anaphoric item. Intuitively, a discourse referent is a body of information, purportedly about a single individual, which may or may not correlate with an individual in the actual world. We often talk about something which doesn't actually exist, so that the corresponding discourse referent, which locally licenses use of a pronoun or other anaphoric expression, is an informational entity, and not a real referent: *Suppose there WERE a unicorn in the garden. Maybe we could tame it.* When interlocutors share such a body

³ I do not adopt all features of these older accounts. Crucially, it is now clear that unselective binding doesn't yield the correct truth conditions, e.g. for donkey sentences. More generally, I would argue that it is not semantics that is dynamic, but pragmatics—the context that should bear on interpretation may change throughout the course of a single utterance, in ways tightly constrained but not fully determined by the compositional semantics. Crucially for anaphora, the changes in context track the merely local introduction of new discourse referents, roofed by any quantificational operators that take wider scope than the NP that triggers introduction, as in Kamp and Heim.

of information, we say that they have a familiar discourse referent. The familiarity is weak (Roberts 2003), because this discourse referent needn't have been introduced by an explicit utterance, but might correlate with some entity whose existence is simply evident to all: [A dog wanders into the classroom, and the instructor says:] *Does he belong to any of you?*

Technically, a discourse referent antecedent is weakly familiar (Roberts 2003) if the corresponding entity is entailed to exist in the **local context** of utterance, in the sense of Heim (1982,1983) and as realized in a variety of theories of dynamic interpretation (Kamp 1981, Groenendijk & Stokhof 1991, Dekker 2012, Zeevat, Muyskens, 1996, AnderBois et al. 2015, Martin 2016, etc.). Such accessibility is a *sine qua non* for anaphora, as argued especially by work on modal subordination (Roberts 1989, 1995, 1996b, to appear). Briefly, logical accessibility limits accessibility of discourse referents introduced under quantification to the scope of the relevant operator, scope which I take to be limited to the sentence in which it occurs.

With respect to (1b), I agree with the authors whose work on coherence is cited above that coherence relations are not *determined* as a function of independently resolved anaphora, but that plausibility of coherence relation and anaphora resolution are simultaneous constraints on interpretation. Thus, we should **think of interpretation in discourse as the resolution of a simultaneous equation in multiple variables, which must satisfy multiple constraints** (Roberts 2017).

As we'll discuss and illustrate below, (1c) does *not* require that the antecedent be maximally salient *simpliciter*, but rather that it be the uniquely most salient discourse referent that satisfies other constraints, as given by the anaphoric trigger itself and the context of utterance.

Under this characterization, you can see why anaphora serves as a probe for discourse structure, and for coherence in particular, for we see in (1b) that coherence is an important constraint on the plausibility of the resulting interpretation. But I will argue that coherence relations and the structure they give rise to over discourse come to bear on salience (1c) as well. Intuitively, the antecedent for an anaphoric trigger in utterance *u* is most likely to be found in those parts of preceding discourse to which *u* coheres most closely—those to which it is **most closely related** under coherence.

In order to characterize the way in which discourse is structured, as reflected in anaphoric potential, two broad claims must be addressed:

- a) There are relations over discourse segments which render it coherent: relevance, rhetorical relations, etc. A given discourse is coherent to the extent that relations of the relevant sort(s) connect the utterances which comprise it.
- b) These coherence relations constrain anaphora resolution, and vice versa.

To give substance to claims (a) and (b) above, two questions must be addressed:

- i. What is the character of the relation(s) that structure discourse, making it coherent?
- ii. How does that structure constrain anaphora resolution?

Here I will argue for the following answers to these questions:

- i. The fundamental structure of a discourse is the hierarchical structure established by the evident joint goals, plans and intentions of the interlocutors, including the understood question(s) under discussion (QUD) and the strategies that subserve those goals and questions. This is the **intentional structure of discourse**.⁴ It subsumes the strategies of inquiry of Roberts (1996), the D-trees of Büring (1999), the coherence relations of Kehler (2002), and the SDRS discourse structures of Asher & Lascarides (2003). None of those others by themselves is sufficient to account for the ways in which intentions bear on anaphora resolution. But the intentional structure can capture all the ways in which other kinds of coherence relations and discourse structures bear on anaphora.
- ii. A QUD-based version of the **Right Frontier constraint** of Polanyi (1985), here defined over the intentional structure of discourse, plays a crucial role in defining an appropriate notion of relative salience, which in turn is crucial for anaphora resolution generally, and for pronominal anaphora, zero-anaphora and ellipsis in particular.

The **maximally salient** entities in a given discourse context are those that are (a) logically accessible along the Right Frontier of the intentional structure of that discourse and (b) pertain to the most immediate goal along that frontier. Entities pertaining to other superordinate goals along the frontier are still salient, but somewhat less so.

Thus, syntactic constraints, logical accessibility and the intentional structure of the discourse (with its Right Frontier) together tightly constrain the potential antecedents for a given anaphoric trigger.

There is one other factor which seems to bear on anaphora resolution: Production studies have shown that when speakers refer to an entity that was referred to by the subject of the preceding sentence, there is a very strong tendency to pronominalize the second mention. Kehler & Rohde (2017:17) summarize some of the literature clearly:

...recent work has shown that speakers pronominalize references to the subject referent considerably more often than other grammatical roles, and further that, rather counter-intuitively, this bias appears to be insensitive to the semantic and pragmatic factors that comprehenders bring to bear during interpretation (H. Rohde, 2008; Fukumura & Gompel, 2010; H. Rohde & Kehler, 2014; cf. Arnold 2001). Indeed, the failure to pronominalize a subject referent mention can lead to a so-called repeated name penalty whether or not the preferred referent for a pronoun would have been the subject (Gordon, Grosz, & Gilliom, 1993; Gordon & Scearce, 1995; Almor, 1999). Fukumura and van Gompel (2010) explicitly link the repeated-name penalty with the strong production bias to pronominalize a re-mention of the previous subject: The processor is so surprised to see a name instead of a pronoun for a referent in subject position that it causes a reading time delay. If this is correct, it stands to reason that the processor would therefore *not* be surprised to see a pronoun used to refer to a subject referent even if anticipatory semantic and pragmatic biases were pointing it away from that referent, hence leading to a lack of processing delay for bias-incongruent subject references.

⁴ The general notion that discourse has an intentional structure based on the interlocutors' goals and plans is originally due to Grosz (1977), and has its roots in Planning Theory from Artificial Intelligence.

However, as their experimental work over several years reflects, this does not result in a bias toward resolution to a subject in all contexts. Rather, it depends both on the verb-type in the first utterance and on the understood coherence relation between the two utterances, two factors which, they argue, are themselves related. We'll return below to illustrate this phenomenon and how it might bear on the question of salience.

To informally illustrate the Right Frontier constraint and understand why we need to define it over the intentional structure of discourse, consider this classic example of anaphora in a naturally occurring task-oriented dialogue, due to Grosz (1977). In (2), an Expert and an Apprentice are working together to disassemble an engine. They're not in the same room, but the Expert can see the feed from a camera trained on the Apprentice's work space:

- (2)
- a. A: One **bolt_v** is stuck.
 - b. I'm trying to use both the pliers_y and the wrench_z to get **it_v** unstuck, but I haven't had much luck.
 - c. E: Don't use the pliers_y.
 - d. Show me what you're doing.
 - e. A: I'm pointing at the bolts_x.
 - f. E: Show me the ½" combination wrench_w, please.
 - g. A: OK.
 - h. E: Good, now show me the ½" box wrench_u.
 - i. A: I already got **it_v** loosened.

The feature of interest is the anaphoric relation highlighted with red type. As Grosz discusses, the intended resolution of the anaphoric pronoun in (2.i) cannot be justified on the grounds of recency of mention of the antecedent, which hasn't been mentioned since (b). In the meantime, two other entities have been mentioned, the ½" combination wrench in (f) and the ½" box wrench in (h), both of which could felicitously be referred to with a singular neuter pronoun. Of course, one can readily construct a variant on (i) in which *it* is used to refer to the box wrench: e.g. *Is **it_u** the one with the open end?* In (i) as it is, the crucial clue is *loosened*, implicating that the intended referent is the bolt that was *stuck*. The question of interest, then, is why the resolution indicated by coindexation with *v* is one possible way of resolving the pronoun, given the long span of talk between it and the mention of its intended antecedent *bolt_v*.

If we merely concentrate on the evident rhetorical relations between the explicit utterances in (2), that by itself will also fail to correctly predict the target anaphoric relation.⁵ Move (a) identifies a problem, which is elaborated on in (b). Moves (c), (d), (f), and (h) are all Directives by the Expert intended to guide the Apprentice in solving her problem; one might say that each is a Continuation of the preceding move, but that isn't obvious since there are intervening contributions by the Apprentice. And what is the relationship between (i) and the preceding contributions?

⁵ Here, as below, I capitalize the names of rhetorical relations, to distinguish them from the ordinary uses of these words.

The way to make sense of this dialogue is in terms of its intentional structure,⁶ as suggested partly by the task at hand. Some of the crucial aspects of this structure are implicit in the dialogue, as laid out in Figure 1, where the labels on the lines of association between nodes are (with one exception, noted) discourse relations drawn from Asher & Lascarides (2003):⁷

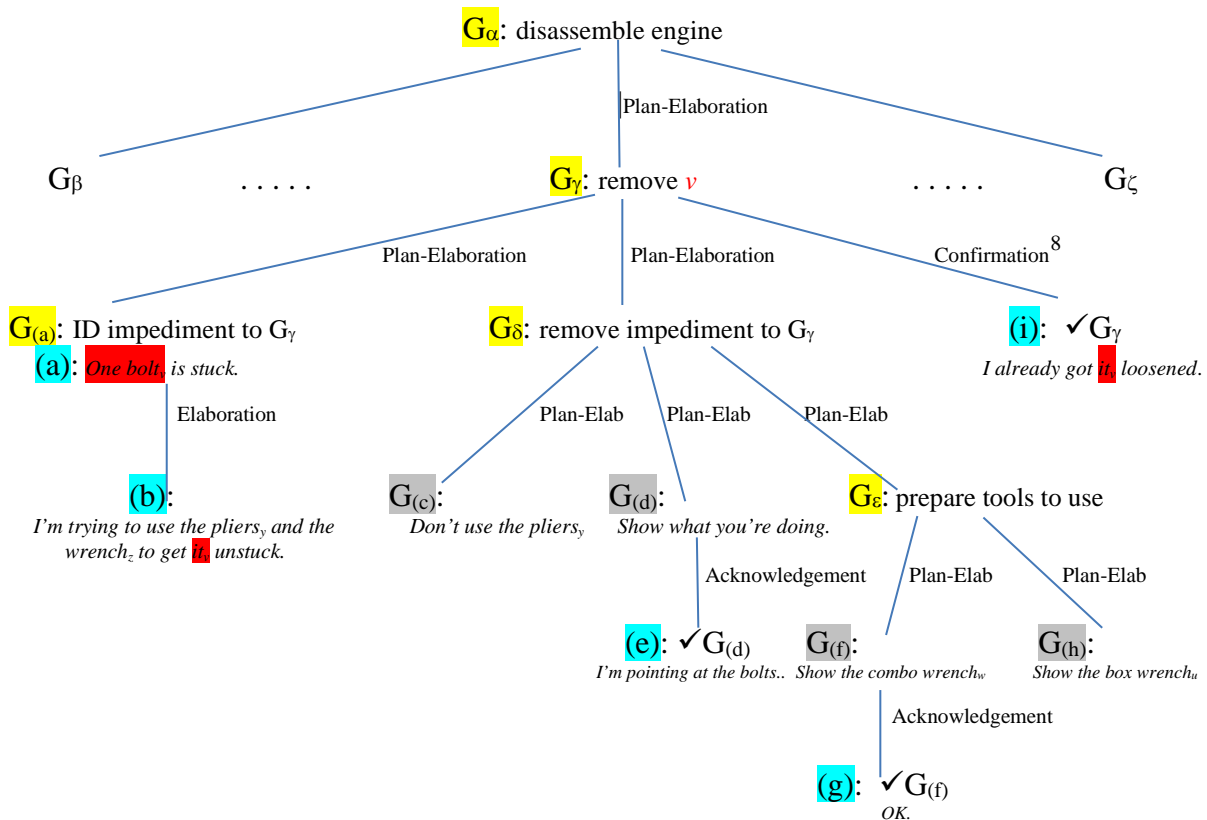


Figure 1: Intentional Structure of Grosz' (1977) dialogue (2)

Key: Nodes introduced by explicit utterances, with associated *uttered content*:

uttered by **Apprentice**: (a), (b), (e), (g), (i)

uttered by **Expert**: (c), (d), (f), (h)

SARG (Speech-Act Related Goal), implicitly introduced; joint goals of A and E.

Introduction of a letter-indexed goal is triggered by the co-indexed explicit move, while those with Greek indices are contextually inferred.

Anaphoric relation of interest

✓G: indicates that the associated goal G is achieved.

⁶ This is close to what Grosz called the *linguistic structure* of the discourse, a term I find misleading as discussed in Roberts (1998b).

⁷ I could have used relations from another inventory, e.g. that of Kehler. But I use A&L's because their coherence relations and resulting discourse structures include a variety of merely implicit moves and relations that are crucial for building a fully integrated discourse structure; and these, I will argue, end up being central components of the intentional elements of the structure. Actually, I think this argues that their inventory is functionally heterogeneous in a way that Kehler's is not, as I'll discuss briefly below.

⁸ This is close to Asher & Lascarides' Acknowledgement, but the latter requires that the superordinate move be that of a different agent, which is not the case here.

In this dialogue there are no conjoined relations—which would be indicated by horizontal association lines, as in Asher & Lascarides, but only subordinating relations—indicated by vertical association lines. A single goal defines the whole interchange of which (2) is part: the over-arching task is to disassemble an engine. One thing required to accomplish this task is to remove bolt *v*. The goal of this discourse segment is to develop and carry out a plan to successfully do this, in view of the impediment that the Apprentice notes at the outset in (a): *v* is stuck. The anaphoric relation of interest is between the denotation of the NP *one bolt* introduced in the first move (a) and *it* in the last move (i). But neither (a) nor the goal $G_{(a)}$ that (a) subserves has a direct rhetorical relationship to (i); their only connection is via G_γ , the implied goal of removing *v*, which goal is presupposed by *stuck* in (a).

Setting aside for the moment the question of the character of the structure in (2), assume that it is dynamically constructed: Each node in the tree either represents the speech act associated with an utterance (the nodes in blue and grey in Figure 1) or one of the associated Speech Act Related Goals or Questions (the nodes in yellow) addressed by such a speech act. At any point in discourse, the intended coherence relation(s) between an utterance and the pre-existing nodes in the structure are inferred and a new node for the utterance is added by one or more association lines, horizontal for coordination, vertical for subordination. Intentional structures are built down and to the right: i.e. the graphic convention is that new nodes, corresponding to new moves in the discourse and their corresponding goals, etc., are introduced rightward.

Take the Right Frontier in a discourse structure like (2) to be defined as follows at any given point in the corresponding discourse:

Right Frontier of a discourse structure:

At any point in the construction of a hierarchical discourse structure like that in (2), when a new node *x* is introduced by some utterance or an (implied or explicit) discourse relation, the Right Frontier of that structure at the time of introduction of *x* consists of the following nodes:

- any pre-existing node *y* to which *x* is immediately conjoined as sister (i.e. conjoined without intervening conjoined nodes), and
- any node(s) *z* that dominate *x* in the hierarchical structure

Then here is a version of the **Right Frontier constraint** on anaphora resolution (*not* that of previous authors). This version is defined in terms of the intentional structure of discourse, and it only constrains anaphora indirectly, via salience:

Right Frontier Constraint:

The discourse referents that are **maximally salient** at a given point in discourse are those that pertain to the most immediate goal along the Right Frontier at that point.

i.e., those that are arguments of that goal or of any immediately preceding sister.

As in Roberts (2003, 2004), the antecedent to a pronominal anaphoric trigger must be maximally salient. Hence, it must lie on the Right Frontier for the node in which the trigger is introduced. Note that even if the antecedent is in an immediately preceding conjunct, then typically the two

conjuncts will both be subordinate to some SARG or QUD; thus, the antecedent must pertain to the most immediate goal along that frontier.

To see an example where the antecedent is given by a coordinated sister node, replace (h) in (2) with (h') to yield:

- f. E: Show me the ½" combination wrench_w, please.
- g. A: OK
- h'. E: and put it next to the screwdriver on the table.

Here instead of (h) introducing a separate element in the plan to identify the appropriate tools, (h') introduces a node coordinated with (f) as part of the Plan-Elab involved in preparing the appropriate tools. This coordination puts node $G_{(f)}$ as the immediate goal on the Right Frontier for the new coordinated node (h'), correctly predicting the felicitous resolution of *it* to the maximally salient *the ½" combination wrench*.

In the actual structure in Figure 1, node (i) has no coordinated sisters, so the only nodes on its Right Frontier are its (immediately dominating) mother G_γ and grandmother G_α . Thus, the anaphoric relationship between the v associated with G_γ and *it_v* in (i) is licensed by the second, dominance clause of the definition of the Right Frontier, and the fact that G_γ immediately dominates, satisfying the Constraint. G_γ itself pertains to the target bolt v introduced in the subordinate $G_{(a)}$ (its far left branch), and it is this that licenses use of *it_v* in G_γ 's subordinate node (i). The Right Frontier constraint does not license direct access of *it_v* in (i) to *one bolt_v* in $G_{(a)}$, since the latter is neither conjoined with nor dominant over (i).

To appreciate the requirement of immediate dominance in the Constraint, consider the following:

First, note that one could replace (2g) with (g'):

- f. E: Show me the ½" combination wrench_w, please.
- g'. A: It/the wrench is too heavy for me to lift.

Here either *it* or *the wrench* would be correctly predicted to refer to the ½" combination wrench in the dominating goal $G_{(f)}$ introduced by the Directive (f). Use of the definite description would be less than optimal when the pronoun is felicitous—we typically use the least descriptive content compatible with correctly retrieving the intended antecedent (cf. Heim's Maximize Presupposition)—but it seems to be acceptable here. At that point in the discourse, though there may be other weakly familiar wrenches in the tool box, there is no other salient wrench as defined in terms of the Right Frontier.

But if the speaker did utter *it* in a node subordinate to $G_{(h)}$, as we saw in (h') above, the structure would make available two potential antecedents along the Right Frontier: the box wrench in the immediately dominating (mother) node $G_{(h)}$ and the v in the great-great grandmother node G_γ . For example, replace (i) with the same clause we saw in (g'):

- f. E: Show me the ½" combination wrench_w, please.

- g. A: OK
- h. E: Good, now show me the ½" box wrench_u.
- i'. A: It/the wrench is too heavy for me to lift.

While *it* would be felicitously understood to refer to the ½" box wrench in the dominating goal G_(h) introduced by Directive (h), *the wrench* is non-optimal here in the context following (f) and (g). Definite descriptions not only do not require that their antecedents be maximally salient (unlike pronouns), but are typically used either when the intended antecedent is *not* maximally salient or when it isn't the only maximally salient entity which could be referred to with a pronoun of appropriate gender and number. So the use of a definite description when a pronoun could have been used tends to suggest that the intended antecedent is not the maximally salient potential antecedent. In (i'), use of *the wrench* suggests that the antecedent may not be the maximally salient box wrench in the most immediate goal along the Right Frontier—the box wrench in G_(h), but perhaps some other wrench along the Frontier. But then there is a problem that did not arise in (g'): Because the combination wrench mentioned in G_(f) and the box wrench in G_(h) are at this point both relevant tools for the dominant goal G_ε ('prepare tools to use'), both are available along the Right Frontier. I.e., there are actually *two relatively salient wrenches*, both within conjuncts of the immediately dominating goal G_ε and both equally salient at that level. So either we have to understand *the wrench* to be used for some reason instead of *it* to refer to the maximally salient box wrench, a motivation which fails here because we could have used *it* to refer to that wrench, or the anaphora is infelicitous at G_ε because of lack of a **unique** appropriate antecedent *at that level*. So we can explain why use of *it* is preferable here to use of *the wrench* to refer to the box wrench.

All this illustrates how the antecedent of a definite description should be *unique among entities satisfying its descriptive content at the same level of salience*. And a discourse referent at an immediately dominating goal is more salient than others farther up along the Right Frontier.

In a follow-up to (h), we only get resolution to the bolt *v* in node G_γ when this is consistent with evidence that the goal associated with (h), G_(h), has either been achieved or, as in the actual (2), is no longer relevant because the goal that G_(h) subserved has been achieved in some other fashion. Thus, consider this revision of the last move in (2):

- h. E: Good, now show me the ½" box wrench_u.
- i''. A: **It's** stuck so bad that I don't think the box wrench will help.

The revised (i'') is awkward—one is at first inclined to take *it* to be the box wrench, but that doesn't make sense, e.g. with the predicate *stuck*. One can reconsider and accommodate the stuck bolt, but it takes work: it is anaphorically infelicitous that the only plausible antecedent for the pronoun is non-maximally salient, especially since there are other entities (the box wrench) which are maximally salient and could be referred to with *it*. Here, if the bolt is intended, it would be better to say (i'''):

- i''': A: The bolt's stuck so bad that I don't think the box wrench will help.

This example and its variants illustrate how interpretation is the resolution of a simultaneous equation in multiple variables, a process repeatedly emphasized by authors who work on complex corpora, including Asher & Lascarides, Ginzburg, and Kehler, and their associates. The resolution of anaphora is not independent of coming to an understanding of the underlying structure of the discourse.

In what follows, we'll:

§II. define the intentional structure of discourse more precisely, basing this on the QUD theory of Roberts (1996/2012), with its central coherence relation of RELEVANCE to the QUD and evident domain goals. We'll relate the intentional structure so-defined to the discourse structures proposed in Büring (1999), Kehler (2002) and Asher & Lascarides (2003), showing how these are consistent with the proposed understanding of discourse structure.

§III. consider other pragmatic factors that have been proposed in the literature on anaphora resolution; and argue that to the extent that they do play a role they can be shown to follow from the structure in II; and motivate why intentional structure should underlie the notion of relative salience in discourse, both intuitively and with (all too brief) appeal to relevant literature in cognitive science and experimental psycholinguistics

§IV. offer conclusions.

II. The intentional structure of discourse

As background to this discussion, I offer a very brief introduction to QUD theory. This is intended as a general theory of the notion of a **Context of Utterance**, represented as the scoreboard for a language game. See Appendix A for a formal characterization of the discourse scoreboard of Roberts (1996, 2013, 2015, 2022). Its key features are:

- a set of goals G , those evident to the interlocutors and understood to guide their behavior in the discourse game. Goals are themselves organized hierarchically, some subserving others or organized into complex plans;
- a set of Questions Under Discussion (QUDs), constrained by logical relations s.t. once a question is under discussion, another question can be posed only insofar as it subserves those already under discussion. This captures the intuition that once interlocutors opt to address a question, they are committed to resolving it insofar as possible, modulo only the continuing relevance to their over-arching goals and plans. All questions q on the QUD correspond to common goals in G : the goal of addressing q ;
- a Common Ground consisting of propositions that characterize the propositional information that the interlocutors (purport to) commonly believe; and
- a set of moves, some implicit and others explicit (speech acts), consisting of assertions, interrogations and directives. Assertions contribute to the CG, interrogations to the QUD (and hence indirectly to G), directives to G (adding, if accepted, the addressee's goal of performing the directed action).

The intentional structure of a discourse is then defined in terms of these elements of the mutually evident scoreboard.⁹

The notion of RELEVANCE is that of Roberts (1996), but now extended as in Roberts (2013) to include not only to RELEVANCE to a question under discussion (QUD), but to an over-arching goal.¹⁰

A move *m* is **RELEVANT** to a goal *g* iff *m* constitutes or proposes part of a strategy for achieving *g*.

Since in that framework a QUD corresponds to a shared goal of the interlocutors—the goal of satisfactorily resolving the question, this definition subsumes the original 1996 definition of RELEVANCE:

A move *m* is RELEVANT to the question under discussion *q* iff *m* addresses *q*, directly or indirectly yielding a partial answer to *q*.

So a move that answers a QUD is RELEVANT at that juncture in discourse in case it addresses the question (contextually entails a partial answer), because doing so helps to achieve the domain goal of answering the question, and hence any domain goals subserved by answering it.

Intentional structure of a discourse:

a static snapshot of a segment of discourse which consists of a structure whose nodes are:

- those elements of G, QUD which were on the scoreboard at some point during that segment of discourse
 - organized according to the subservience relations over them, and
- any other moves made during the segment, along with
 - subservience relations between those moves and G and QUD, and
 - any other understood coherence relations over those moves (e.g., among those in the inventory of Content-level relations of Asher & Lascarides, or the coherence relations of Kehler)

The intentional structure thus consists of a hierarchical organization over (a) G, (b) the discourse goals in the QUD, and (c) the discourse moves which subserve them. Each node in such a structure either serves to:

- | | |
|--|-----------------|
| (a) make goals explicit/entail them, | [set-up moves] |
| (b) subserve these goals by | |
| (bi) establishing sub-goals (to form strategies) or | [set-up moves] |
| (bii) directly or indirectly (helping to) achieve those goals,
(answering questions, suggesting how to achieve goals) | [pay-off moves] |
| (c) serve the auxiliary goal of maintaining the scoreboard
(clarifying, accepting, rejecting, connecting, etc.) | [referee play] |

⁹ Compare Asher & Lascarides' (2003:138) discourse structure.

¹⁰ Again, if we followed Kaufmann & Kaufmann (2012) in characterizing goals as decision problems, they *are* questions, and the proposed revision would follow from the earlier definition.

Moreover, moves themselves may be complex, with strategies for each type of move (a) – (c) above:

- (a) strategies for achieving goals generally, and strategies of inquiry in particular
- (b) strategies for conveying information (including explicit and implicit rhetorical relations between utterances) (Kehler 2002, Asher & Lascarides 2003, Ginzburg 2012, etc.)
- (c) complex policing strategies (some discourse particles play a role here)

Such complex strategies can be captured as strategies for achieving goals. This is defined over the goals in G on the discourse scoreboard, as a generalization of the notion of **strategies of inquiry** in Roberts (1996) (given in Appendix B):¹¹

The *strategy of goal-achievement* which aims at achieving goal g , $\text{Strat}(g)$:

For any goal $g \in G$, $\text{Strat}(g)$ is the ordered pair $\langle g, S \rangle$, where S is the set such that:

If there are no $g' \in G$ such that g' subserves g , then $S = \emptyset$.

Otherwise, for all $g' \in G$, g' subserves g iff $\text{Strat}(g') \in S$.

We can understand the ordered pair which Strat yields for a given goal g , $\langle g, W \rangle$ as ‘the strategy to achieve g by carrying out the set of sub-plans in S ’. So in Figure 1, the discourse fragment represented by G_γ , removing the bolt, is part of a larger strategy to achieve the goal G_α of disassembling the engine. In turn, identifying the impediment to that goal ($G_{(a)}$), and removing it ($G_{(\delta)}$), are part of a strategy to remove the bolt, to achieve G_γ ; the tree rooted in G_δ represents the strategy to achieve that goal; etc. Hence, the tree structure represents the strategy for goal-achievement in this discourse. Each goal-node in the tree dominates a strategy to achieve it. Büring’s (1999) D-trees are an instance of these intentional structures: a super-question and its sub-questions related by contrastive Topics, along with the answers to those sub-questions.

How does this bear on the intentional structure (Figure 1) of the discourse in (2)?

In Figure 1, each goal node represents a shared goal of the interlocutors. In some cases these goals correspond to explicit utterances in the actual discourse: $G_{(c)}$, $G_{(d)}$, $G_{(f)}$ and $G_{(h)}$ are goals introduced by directive speech acts, expressed by utterances in the imperative mood (see Portner 2007, Roberts 2018, 2022 for accounts in which the directives canonically expressed by imperative clauses add to the interlocutors’ evident goals, typically those of the addressee). But the remaining goals are introduced either as Speech-Act-Related-Goals—like $G_{(a)}$, implied by utterance (a), or as goals implied by the understood task in which the interlocutors take themselves to be engaged: G_α , G_γ , G_δ and G_ϵ . These SARGs and other implicit goals are crucial to the coherence of the structure: without $G_{(a)}$, G_γ , and G_δ , it is difficult to understand the relationship between (a) and (c); but because it is understood that the goal is to remove the bolt, it is understood in (c) that the Expert intends that the Apprentice should not use the pliers *to remove the bolt*; she is not prohibiting all uses of the pliers, but only their use in this particular task. And it is only because all of the content under G_δ was understood to pertain to the stuck bolt that the final utterance (i) can felicitously involve pronominal anaphora to the bolt, despite the

¹¹ A strategy for achieving a goal can be understood as a plan, both characterized and constrained in terms of Planning Theory (see the work in Cohen et al. 1990, and references therein) and the related approach to philosophy of action in Bratman (1987). Space precludes exposition here, but that should be kept in mind: there is a background theory for what it means for a plan to be well-formed, rational and coherent.

fact that its last-mention was not recent. The rhetorical relations involved between explicit utterances—the Elaboration between (a) and (b), the Acknowledgement of (d) in (e) and of (f) in (g), as well as the confirmation of G_γ in (i), and the various Plan-Elaboration relations of (c), (d), (f) and (h) to their superordinate goals, are all important building blocks used in the structure. But the core, and the driving force, are the goals and the interlocutors' cooperative commitment to addressing them in the interchange. Remove them, and the discourse lacks overall coherence.

This is this not merely a reflection of the fact that (2) is a task-oriented dialogue. We see something quite similar in (3), an information-seeking dialogue from Asher & Lascarides (1998a), with intentional structure in Figure 2:

- (3) a. **A**: I need to catch the 1:20 to Philadelphia.
 b. Where's it leaving from?
 c. **B**: Platform 7.
 d. **A**: Where do I get a ticket?
 e. **B**: From the booth at the far right end of the hall.

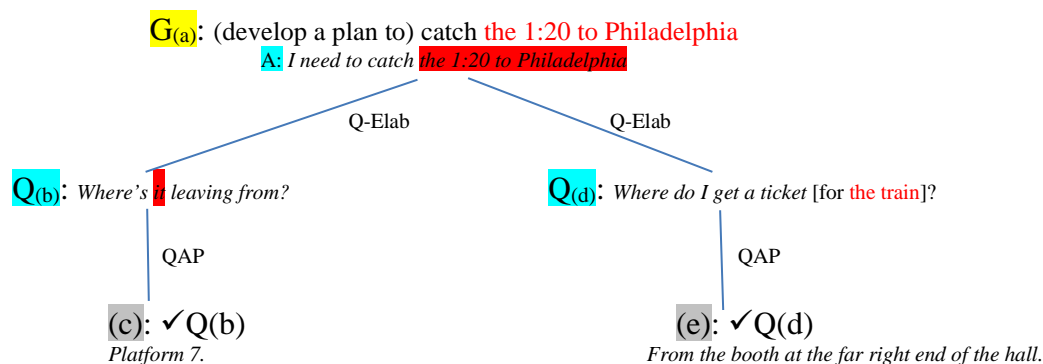


Figure 2: Intentional structure of Asher & Lascarides (1998a) (3)

Again, I use Asher & Lascarides' (2003) discourse relations to develop this structure. As in Figure 2, the blue and grey highlighted nodes are introduced directly by utterances in the discourse, the yellow node is an implicit SARG, here introduced by A's utterance (3.a), wherein her need to catch the train implies that this is her goal in engaging in conversation. This SARG is the basis for taking the two questions (b) and (d) to stand in the Q-Elab relation to $G_{(a)}$. Q-Elab is a relationship in which a question serves as part of the elaboration of a plan to achieve an understood SARG; so here the pair of questions serve as a strategy of inquiry to help develop a plan to achieve the goal of catching the train. These questions are resolved by the utterances (c) and (e), presumably resulting in a satisfactory plan. Again, without the understood SARG $G_{(a)}$ the discourse would lack coherence. Not only does this structure license the red-highlighted anaphora, but also (d) is understood to mean 'where do I get a ticket *for the train*', the implicit purposive modification arising from an enrichment implicature based on RELEVANCE to the overarching goal.

Again, lest it be claimed that this type of intentional structure is limited to task-oriented and information-seeking dialogues, note the following features of the inventory of discourse relations assumed by Asher & Lascarides (2003) (see Appendix A):

- the Content-level relations are those that relate pairs of explicit utterances.
 - But several of these (Background, Continuation, FocusBackgroundPair, and Narration) crucially presuppose that the pairs so related are subordinate to a shared Topic. A Topic in the sense they intend is arguably a QUD (van Kuppevelt 1995,1996, von Fintel 1994, Ginzburg 1994, Roberts 1996, 2012a). QUDs are discourse goals, and goals represent intentions. So all of these Topic-subordinate relations do involve subordination to an often implicit intention.
 - The Content-level relations Topic and QAP involve subordination to an explicit Topic or question, so are intentional, as well.
 - The remaining Content-level relations are either logical (Alternation/ \vee , Consequence/ \Rightarrow , DefeasibleConsequence) or involve causal relations (Explanation, Result). Utterance pairs so-related, along those related with the subordinating relation of Elaboration, usually serve *as a unit* as just part of a larger strategy of inquiry. Thus, they subserve a larger intentional structure, given at least by the over-arching topic of conversation (QUD). Note in this connection that the answer to a question is quite often represented by a sequence of utterances, rather than by utterance of a single sentence. In other words, a move may be multi-sentential, and such complex moves themselves are related under Coherence.
- the Text Structuring relations (Contrast, Parallel) themselves presuppose a contrasting or common theme. But again, a theme is a Topic (Halliday 1967, Roberts 2012b, Westera 2017).
- The Cognitive-level relations all involve either subordination to an explicit or implicit SARG or question. Thus, again, these relations all subserve intentions.

This leaves only Divergent relations (Correction, Counterevidence, Dis(R)) and Metatalk. But again, a pair of utterances related in one of these ways does not exist in an intentional vacuum, but serves to contribute to or to guide, correct and referee the play in an intentionally structured interchange. For extensive research and discussion about how such meta-relations play a role in constraining discourse play, see Ginzburg (2012) and the large body of related work in his general framework. I take such moves to be part of the crucial policing practices interlocutors' engage in—crucial to maintaining a good general grasp of the scoreboard for all interlocutors, but orthogonal to the intentional structure that is the primary driver of the inquiry being conducted.

The hypothesis is, then, that most of the Content-level relations of Asher & Lascarides (2003) and the coherence relations of Kehler (2002) are pre-compiled types of building blocks that play roles in the construction of complex strategies of inquiry and goal resolution. They are often (semi-)conventionally signaled, especially by particular connectives or discourse particles. And they put their own constraints on anaphora resolution. But there the intention here is not to reduce such discourse relations to goals or questions. Rather, it is to claim that the central structure *in which such relations play a role* is intentional in character, constrained by features of what it is to be a rational agent cooperatively pursuing joint goals with another agent.

Kehler and colleagues (Kertz, Kehler & Elman 2006; Rohde, Kehler & Elman 2006 and 2007; Kehler, Kertz, Rohde & Elman 2008; Rohde & Kehler 2008a and 2008b; Kehler & Rohde 2014, 2017, Westera & Rohde 2019) have conducted a long suite of experiments arguing that

coherence relations guide and constrain anaphora resolution, and that even different conjoining relations (e.g. Parallel vs. Explanation) yield different patterns of resolution. Kehler (2009) points out that in their experimental materials Rohde et al. (2006) used different types of questions to bias to different coherence relations—e.g., *What happened next?* to bias to the relation Occasion, or *Why?* to bias to Explanation. He agrees with Roberts (2004) that we can understand different coherence relations as reflecting different strategies of inquiry in a QUD-based discourse structure, which sets up “expectations about how the discourse will be continued with respect to coherence”.

In Kehler & Rohde (2017) the authors present even more compelling experimental evidence that interlocutors’ grasp of the intended coherence relations between two utterances and of the QUD are closely related, each constraining the other.

I take Ginzburg’s (2012) discourse relations in KoS, and especially his use of the QUD, to also be broadly compatible with what I’m proposing here; but Ginzburg tends to focus more on moves that referee play (type (c) above), including especially strategies for correction.

III. Intention, attention, and salience in anaphora resolution

A variety of pragmatic factors have been taken to be involved in anaphora resolution in discourse, prominent among them:

- recency
- grammatical parallelism
- Centering principles
- visual properties in a scene
- prosodic focus

What role do these factors play in anaphora resolution? How, if at all, are they related to the intentional structure of discourse?

1. recency: there is good reason to think that recency of potential antecedents is a relatively minor factor in anaphora resolution. See Grosz’s discourse (2) above for an illustration, wherein the initially mentioned bolt was the preferred antecedent for the pronoun, over the more recently mentioned wrenches. Terken & Hirschberg (1994) offer experimental evidence that recency is less important than grammatical parallelism.
2. grammatical parallelism: Besides Terken & Hirschberg, Smyth (1994) offers experimental evidence for a preference for antecedents with the same grammatical role as the pronominal. But Kehler (2002) points out confounds in Smyth’s (1994) materials, raising doubts about whether parallelism is an independent factor in anaphora resolution. Instead, in the suite of experiments cited above, he and his associates provide evidence that coherence, as reflected in felicitous rhetorical relations, is more successful than grammatical role parallelism in predicting the preferred resolution. Parallelism is thus just epiphenomenal, reflecting certain common rhetorical relations, but can be readily over-ridden when other kinds of relations are brought to bear. For example, consider the following from Kertz, Kehler & Elman (2006):

- (4) Samuel threatened Justin with a knife, and
- | | | |
|--|------------|-------------------------------|
| a. ...Erin blindfolded him (with a scarf) | [Parallel] | [parallel thematic roles] |
| b. ...Erin stopped him (with pepper spray) | [Result] | [non-parallel thematic roles] |
| c. ...he blindfolded Erin (with a scarf) | [Parallel] | |
| d. ...he alerted security (with a shout) | [Result] | |

When the follow-up stands in a Parallel rhetorical relation with the first conjunct, in (a) or (c), we find the expected parallel thematic roles. But when the (just as easily processed) Result relation is more plausible, in (b) or (d), subjects strongly prefer to resolve the object pronoun *him* to the non-parallel prior subject *Samuel*. Kehler (2009:8) summarizes:

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.

3. Centering Theory (Grosz, Joshi & Weinstein 1995; see Walker et al. 1998 for overview, different realizations and critical discussion) arose in computational linguistics as a set of heuristics for determining the likely occurrence and intended resolution of pronominal anaphora in an uttered sentence as a function of (a) the syntactic structure of the target sentence (including word order and the grammatical or thematic role of a target pronoun) and (b) that of the immediately preceding sentence uttered (and in particular, the word order and/or grammatical role(s) of any potential antecedent NPs in that preceding sentence). Arguments in an utterance are ranked as a function of these syntactic factors; for example, the sentence-initial topic or subject of a sentence is typically ranked higher than other arguments both as a potential antecedent (in the preceding sentence) and as most likely to be pronominalized (in the target sentence). Principles are proposed which predict the likelihood of pronominalization and coreference resolution as a function of the relations between the ranked arguments (“Centers”) in the two utterances, relating the highest-ranked “Backward Looking Center” of the second utterance to the set of “Forward Looking Centers” of the previous sentence (its ranked potential antecedents).

But there is empirical evidence that to the extent that such principles are applicable, they are instead subordinate to a requirement of relevance to task. Gordon, Grosz & Gillion (1993) argue that there is no psychological or empirical evidence for the claims of Centering Theory about preferences for certain types of transitions (pronominal coreference relations) between utterances in discourse, e.g. for Continuations (wherein a subject argument is more likely than others to be pronominalized if it is coreferential with some argument of the preceding sentence) to be preferred over other kinds of transitions. Poesio et al. (2004) argue that rhetorical relations are more important in determining pronominal relations between utterances than Centering principles *per se* (“an analysis in terms of underlying semantic connections between events or propositions is more perspicuous than one in terms of entity coherence”, p.80), and that “Topic Continuity” in particular—whereby supposedly there is a preference for same-Topic from one utterance to the next, is not robust. Since Topic Continuity for subject-initial languages like English is a way of encoding a preference for the subjects of adjacent sentences to be coreferential, this is an argument that there is no strong preference for subject antecedents. And while Tetreault & Allen (2004) conclude that some essentially semantic information (about events and situation types, object types, and other

content that could be automatically retrieved) significantly improved their pronoun resolution algorithm, Tetreault (2005) found that “naive versions of Grosz and Sidner's [1986] theory and Kameyama's intrasentential centering theories” did not, concluding that “Our results show that incorporating basic clausal structure into a leading pronoun resolution [algorithm] does not improve performance.”

Finally, Poesio & Rieser (2011, especially §5.4,261ff) offer a sophisticated, integrated computational model of anaphora resolution. This model takes into account the relevant psycholinguistic evidence that anaphora resolution is incremental, which is to say that it tends to take place in real-time, prior to the completion of utterance interpretation. A central element of their system is the modeling of incremental shifts in joint focus of the interlocutors. These shifts take place as a function of joint tasks, e.g. interlocutors moving together through areas on a map (the TRAINS corpus, Allen et al. 1995) or in visual world studies as a consequence of instructions like *Pick up the cube. Put it in...*, where attention is thereby focused in the visual array on the set of containers into which the cube would fit (Brown-Schmidt et al. 2005). The expectations established through such joint tasks affect what Brown-Schmidt et al. call a “rapid restriction of referential domains”, limiting the set of potential antecedents for any anaphoric elements. As part of their system, Poesio & Rieser do use a version of Centering theory as one among many tools, but the Centering principles are only invoked as a last resort: “The establishment of (Centering-guided) bonding [anaphoric] links is one trigger for further inference processes that hypothesize dominance/satisfaction-precedes relations between the core speech acts generated by the two utterances, **if they haven't already been established by coherence assumptions, or by previous intention recognition processes**” [my emphases, CR]. So on this model Centering principles only come into play if coherence (rhetorical relations) and/or the joint attention restriction [presumably including the sort observed in the eye-tracking studies] have failed to resolve the anaphoric relation in question, and even then are at best a default (over-rideable) feature of anaphora resolution.

What other factors might we appeal to in order to characterize what it is for a potential antecedent to be salient?

4. visual properties in a scene: Visual salience clearly does matter in anaphora generation and resolution. Clarke, Elsner & Rohde (2015) offer experimental evidence that “visual properties (salience, clutter, area, and distance) influence R[eferring] E[xpression] G[eneration] for targets embedded in images from the *Where's Wally?* books. Referring expressions for large targets are shorter than those for smaller targets, and expressions about targets in highly cluttered scenes use more words. We also find that participants are more likely to mention non-target landmarks that are large, salient, and in close proximity to the target.”

Similarly, Gleitman et al. (2007) conducted production experiments involving perspective predicates (*give/take*) and active/passive pairs; they used subliminal attention-capture manipulation of entities in a visual field—for example “a sudden onset, which is undetectable to the speaker but nevertheless influences initial saccades to characters...” (550) to manipulate visual salience of potential referents. Then they immediately asked the speaker to describe the depicted action. Subjects showed “a reliable relationship between initial looking patterns [induced by the subliminal attention-capture manipulation] and speaking

patterns [which of two arguments would be more likely to be chosen as subject of the subsequent description of the scene].”

But I would argue that high visual/physical salience won't suffice if the entity in question isn't understood to be RELEVANT to the topic of conversation. Consider (5):

- (5) [Context: You and I are sitting in a café discussing how to understand Sperber & Wilson's (1985) definition of Relevance, and I say:]
I see it now!
[Even though I'm holding a coffee mug by the handle right under your nose and shaking it for emphasis, you don't take *it* to refer to the mug.] (Roberts 2010)

This argues that in the *Where's Wally?* experiments it was crucial that the visual scene was the topic of discussion. In (5), the coffee mug is clearly *not* among those things relevant to the discussion of Sperber & Wilson, and so is not among those things the interlocutors would be likely to consider as possible referents.

This suggests that visually salient entities are not generally available as possible referents *unless* it is already clear that they are potentially RELEVANT to the QUD (as they are in the visual arrays used in experiments like *Where's Wally?*).

5. prosodic focus: Focused constituents are those associated with prosodic prominence. In English and many other languages, this involves some kind of conventionally marked pitch excursion over the course of a tune; in Japanese, it is marked by pitch range expansion, also conventional in character and import; there are other, rhythmic ways one might mark one constituent in such a way as to make it appear more prominent than its surroundings. Focused constituents are more perceptually salient than those which are not (and are thereby backgrounded). This is a function of the signal itself. And there seems to be a grammar of accompanying gesture which must be congruent with aspects of prosody, including phrasing and emphasis (Cassell et al. 1994).

But the role of Focus in context has been taken to be quite different than that of making the denotation of a focused constituent itself more salient: it distinguishes some constituent in the utterance, a member of a contextually relevant set of alternatives, whose denotation is asserted to be the correct answer to the QUD or to contrast in some other way with other elements of that alternative set, etc.¹²

¹² Bock & Mazzella 1983, Terken & Nootboom 1987, and many other psycholinguistic studies ignore the literature in pragmatics and semantics on Focus, instead inquiring into the role of prosodic prominence in marking “new” vs. “given” information or referential NPs. But there are at least three different notions of “givenness” in discourse. The notion which is arguably relevant for understanding Focus is this: the distinction between what is thematic with respect to the QUD (thematically-given—part of the question but not the answer), vs. what is rhematic (rhematically-new—roughly, that part of the utterance intended to be the answer). Moreover, pragmatic Focus as reflected in prosodic prominence constrains phrasing: under most theories of prosodic constituency and its relationship to Focus (e.g. Selkirk 1996), there can be no more than one pragmatic Focus per intermediate intonational phrase. Hence, the complement of prosodically reflected Focus, i.e. prosodic backgrounding is also phrasally constrained. So taking the Focus to be the rheme, and backgrounded content to serve as the theme with respect to the understood QUD, both prominence and phrasing play a role in reflecting the QUD addressed by an utterance (Roberts 1998; Féry & Samek-Lodovici 2006; Beaver & Clark 2008).

Rooth (1992) argues that the prosodic focus associated with an utterance conventionally presupposes the relevant set of alternatives. Roughly, we abstract on the focused element(s) in the constituent uttered and then take the presupposed set of alternatives to be those we derive by permitting the variables to range over all the contextually relevant values of the appropriate type. Roberts (1996/2012) argues that for a given focused utterance, the alternative set resulting from this abstraction must be congruent with the QUD. Semantically, a question is itself a set of alternatives—the possible answers to the question. Then congruence requires that the focally determined alternative set is the set of answers to the QUD. We see evidence for this in the following felicity judgments, where the (in)felicity of the answer is purely a function of focus:

- (6) What does Alex like to eat?
 - (i) Alex likes PASTA.
 - (ii) #ALEX likes pasta.

- (7) Who likes pasta?
 - (i) #Alex likes PASTA.
 - (ii) ALEX likes pasta.

Salience in this sense—prosodic prominence in the signal—plays no direct role in anaphora resolution. I.e. by itself it is not used to make one potential antecedent more salient. But it may be used to indirectly help resolve anaphora by indicating (a) the understood QUD, or (b) contrast across structure. Consider the well-known example-types illustrated by (8) and (9) (Lakoff 1971):

- (8) Julie said Alice was a Republican, and then she INSULTED her.
- (9) Julie said Alice was a Republican, and then SHE insulted HER.

One gets different truth conditions for the second conjuncts of these two string-identical utterances, with *she* in (8) coreferential with *Julie*, *her* with *Alice*, the opposite resolution in (9). We can explain this difference straightforwardly on the basis of the different prosodic prominences in the two—with the pronouns unaccented in (8), accented in (9)—and the independently motivated semantics and pragmatics of prosodic focus. In the second conjunct of (8), via Rooth’s Focus abstraction we derive the set {*she R-ed her*, *R* a two-place relation}, while in (9) we get {*x insulted y*: *x* and *y* individuals}. We take these alternative sets to be the questions presupposed/ For (8) this is the question *What did she do to her?*. Since both conjuncts in a conjunction must address the same question, it is natural to resolve *she* in (8) to Julie, *her* to Alice, yielding the question ‘what did Julie do to Alice?’. But in (9) the presupposed question is ‘who insulted who?’; then this implies that calling someone a Republican is an insult (in order for the first conjunct to constitute a partial answer), and in turn, implies that the order of the referents is reversed in the second conjunct (in order for it to be informative, given the first conjunct). So the role of prosodic focus in anaphora resolution, when it comes to bear, is conventionally triggered and very robust, but again, it is

itself a function of the QUD.¹³ Since the QUD reflects the discourse task at hand, relevance to task again is argued to be a central factor in anaphora resolution.

Summarizing, the factors we've just considered seem to only bear on anaphora resolution in indirect or secondary ways, if at all. And when they do, as with prosodic focus, this influence is mediated by the way that focus reflects the QUD. Coherence relations do play an important role, as Kehler et al. show; but again, this is because of the important part they play in determining the intentional structure of the discourse.

I take all this to argue for the view of salience and its role in anaphora resolution outlined above: one in which salience is a function of the Right Frontier of the intentional structure of the discourse segment in which an anaphoric trigger occurs, reflecting RELEVANCE to the most immediate goal or question under discussion.

Why would we expect this to be the case? What does RELEVANCE in this sense have to do with salience?

A great deal of work on human perception, and visual perception in particular, argues that in processing percepts we bring to bear expectations about what's relevant to task, using these to filter out at the pre-conscious level possible interpretations of those percepts that are incompatible with expectation. This is reflected in the famous phenomenon of **inattentional blindness** (Simons & Chabris 1999). To see how this works, watch their video, available here, before reading on: <http://www.youtube.com/watch?v=vJG698U2Mvo>.

This short video was shown to subjects, who were directed to count the number of times a team of players pass a basketball. The passing is fast and difficult to follow, with a different team passing a different ball in their midst, so most subjects paid close attention to their task. As a consequence, a very high number failed to notice that in the middle of the video a person dressed in a full black gorilla suit walks on-camera, stops in the middle to pound his chest while facing the camera, then turns and slowly walks off, while the players ignore the gorilla and continue passing the ball. Even though there's evidence that the presence of the gorilla is registered preconsciously, it fails to come to awareness.¹⁴ This finding has subsequently been replicated repeatedly, across a variety of types of task. Thus, attention to task arguably functions to limit attention to those entities that the subject expects to be relevant to that task.

¹³ Note that in cases like (9), it is *not* necessarily the case that the focally presupposed question—the Current Question (CQ) of Simons et al. (2017)—was evident prior to utterance. Instead, all that's required is that the CQ itself must be relevant to the prior, perhaps explicit QUD. See Roberts (1996/2012), Büring (1999) for further discussion.

¹⁴ A follow-up study by Memmert (2006) used eye-tracking to study experimental subjects while they viewed the same video. They found that even people who report not noticing the gorilla do look at it, as evidenced by eye-movements and fixations: "Observers who did not notice the unexpected object in the basketball game test by Simons and Chabris (1999) spent on average as much time (about one second) looking at the unexpected object as those subjects who did perceive it." They conclude: "Motivational factors therefore control the direction of attention and influence the information process before the organism consciously perceives the specific input." So arguably it isn't that subjects don't *see* the gorilla, but they somehow don't take note of it.

I hypothesize that relevance to the immediate goals and intentions in discourse plays a similar role in limiting the potential referents for anaphora resolution, and this is the reason for the Right Frontier constraint.

There's a commitment involved in having an intention to achieve a goal (Bratman 1987): when an agent has such a commitment, distraction is dysfunctional. Accordingly, when we're trying to achieve a goal, commitment entails that we focus on those entities, relations, information and activities that pertain to that achievement: those that directly or indirectly facilitate or impede that achievement. These are the RELEVANT entities.

Joint goals lead to joint commitment, hence (ideally) to joint focus on their achievement. The resulting joint focus on the relevant entities, relations and activities is partly a function of Common Ground: Even if I know of some entity that would facilitate achievement of our goals, it cannot be assumed that it's in the joint focus of my interlocutors until I have reason to think that they know of it as well, i.e. until it's been introduced to the Common Ground. Thus, there's a presupposed requirement of weak familiarity of all RELEVANT entities.

For discussion of a range of psycholinguistic work supporting this general perspective on the relationship between intention and attention in interpretation, see Appendix C.

This general story about intention and attention suggests a specific instance in the case of anaphora resolution:

Attentional Masking Hypothesis: The joint goals and plans of interlocutors serve to make maximally salient those entities that play a role in the most immediate goal or question along the Right Frontier of the intentional structure of the discourse: Focusing attention on those entities—typically a very small set—masks from “view” any less salient entities that are weakly familiar to the interlocutors. I.e., as in inattentive blindness, only the most salient entities, so defined, come readily to attention as possible antecedents.

This suggests paired strategies for resolution and generation of anaphoric relations. Keep in mind that these are intended to bear not only on the resolution of pronominal anaphora, but on that involving demonstrative and definite descriptions as well:

Resolution strategy for interpreting anaphoric triggers: The search for an anaphoric antecedent, among those entities weakly familiar to the interlocutors in a discourse, proceeds as follows:

1. Look first to the most salient entities—those discourse referents that play a role in an immediately conjoined move or in the most immediate goal or question along the Right Frontier of the intentional structure of the discourse.
2. Should you fail to find a uniquely suitable antecedent in step (1), look successively to those that are still RELEVANT but decreasingly salient—entities pertaining to increasingly distant goals or questions along the Right Frontier.
3. Then consider entities recently discussed but irrelevant to the dominating goals.
4. And finally consider all weakly familiar entities.

5. The antecedent for an anaphoric trigger is the first discourse referent you encounter, in steps (1) – (4), which is informationally unique in satisfying the NP’s descriptive content among the discourse referents ranked at its level of salience.

Should this strategy fail to reveal an antecedent that is unique in satisfying the NP’s descriptive content at some level of salience, then the anaphora is infelicitous. Interpretation may still be saved by plausibility and informed guessing, but felicity is at least reduced, and the employment of these means is usually noticed.

Generation strategy governing choice of descriptive content for an anaphoric trigger: To guarantee the intended anaphora resolution for a definite NP, a speaker should choose one whose descriptive content is just sufficiently rich to uniquely identify the intended discourse referent among all those which are at least as salient, with salience as defined in terms of the intentional structure of the discourse.

These strategies and the underlying notion of salience account for a trade-off in felicity between richness of descriptive content and required degree of salience:

- (10) I was shopping in Kroger’s this afternoon when a young guy asked me for help choosing vegetables. He had his older buddy with him, and they both seemed clueless.
 - a) He asked which greens were kale.
 - b) The man asked which greens were kale.
 - c) The young man asked which greens were kale.
 - d) The young man I met at Kroger’s asked which greens were kale.
 - e) The man I met last week in Weiland’s had a similar problem with rutabagas.

In (10), we have the beginning of a story. The first sentence sets up a scenario and begins the Narration: the required Topic of Narration is established by *I was shopping in Kroger’s this afternoon*, which sets up the Topic: ‘what happened while I was shopping this afternoon at Kroger’s’. The *when*-clause is related to that Topic by Narration, making the (discourse referent for the) young guy salient in the topical situation. The second sentence gives some Background for the Topic, making the older buddy salient as well. So both the young guy and his older buddy are arguably salient at the point where one of the target continuations (a) – (e) occurs; due to the telic predicate (and the way that Reference Times are moved along in discourse—see Partee 1984), the chosen target sentence will be understood to be conjoined with the first utterance in a Narration relation.

Neither *he* (a) nor *the man* (b) is sufficiently rich descriptively to distinguish between the two salient agents, so neither is felicitous. One can force resolution to the younger man on the grounds of plausibility, either on the grounds of parallelism to the subject of the *when*-clause or because he’s the one who asked the speaker a question; but neither resolution reflects an optimal use of these triggers. The descriptive content of (c) is preferable here, both adequate to differentiate the two salient men and superior to that in (d), which is richer than required and so rather odd: The first utterance restricts attention to the scenario in Kroger’s, so there is no need to mention that location again. But (e) is fine, despite the fact that the Weiland’s man is irrelevant in (10)—the descriptive content is sufficiently rich to pick him out uniquely anyway (assuming

he's the only such man the two interlocutors have discussed), so the anaphora can be resolved. If we continue with (1c), then in a subsequent utterance *he* would take the younger guy as antecedent (subject to plausibility), because that would be the sole salient entity in its immediately preceding conjunct under Narration, (1c).

These strategies and how they interact with the descriptive content of anaphoric triggers illustrate the claim that the notion of anaphoric salience defined here is relative. The less descriptive content a trigger has, the higher the degree of salience typically required for its felicitous use.

Micha Elsner (p.c.) pointed out to me that there is a class of apparent counter-examples to the resolution strategy, represented by examples like (11):

(11) Ralph saw a man enter the convenience store. The man was carrying a paper bag.

Ralph presumably being a man, there are two men in the scenario described. But the subject of the second sentence, *the man* is clearly intended to refer to the man Ralph saw entering the store, and not to Ralph. There is no sense of infelicity here, despite the fact that the descriptive content applies to both salient entities.

Note that *he* could be used here to pick out either Ralph or the other man; at the outset, this would lead to ambiguity, though the content of the utterance could help to disambiguate. E.g. either of the follow-ups in (12) and (13) strike me as relatively felicitous, if perhaps forcing some accommodation. The pronoun in either is disambiguated by plausibility considerations:

(12) He thought the bag looked suspicious.

(13) He pulled a gun from the bag and told the clerk to give him the money from the till.

Use of *Ralph* in (12), *the man* in (13) would, I think, be slightly better. But, again, why is *the man* ok in (11)?

I think the answer lies in the fact that to disambiguate to Ralph we can just use his proper name. Proper names are special in that they have no descriptive content. And in most contexts of use, they directly pick out their bearers, with little risk of ambiguity or lack of clarity. Hence, there's only one salient referent for whom the best non-pronominal identification in this discourse context is with the description *the man*: the guy Ralph saw. That suggests a possible revision of the resolution strategy:

Resolution strategy for interpreting anaphoric triggers (revised): The search for an anaphoric antecedent, among those entities weakly familiar to the interlocutors in a discourse, proceeds as follows:

1. Look first to the most salient entities—those discourse referents that play a role in an immediately conjoined move or in the most immediate goal or question along the Right Frontier of the intentional structure of the discourse.
2. Should you fail to find a uniquely suitable antecedent in step (1), look successively to those that are still RELEVANT but decreasingly salient—entities pertaining to increasingly distant goals or questions along the Right Frontier.

3. Then consider entities recently discussed but irrelevant to the dominating goals.
4. And finally consider all weakly familiar entities.
5. The antecedent for an anaphoric trigger is the first discourse referent you encounter, in steps (1) – (4), which is informationally unique in satisfying the NP’s descriptive content among the discourse referents ranked at its level of salience, *ignoring those entities which satisfy that descriptive content but could be unambiguously referred to in some other way consistent with the interlocutors’ joint information.* (e.g., with a proper name).

IV. Open questions and conclusions

Recall that Poesio et al. (2004) in their work with anaphora resolution algorithms found that “Topic Continuity”—whereby supposedly there is a preference for same-Topic from one utterance to the next, is not terribly robust. Since Topic Continuity for subject-initial languages like English is a way of encoding a preference for the subjects of adjacent sentences to be coreferential, this is an argument that there is no overall strong preference for subject antecedents. Recall example (4), where when the follow-up stands in a Parallel rhetorical relation with the first conjunct ((a) or (c)), we find the expected parallel thematic roles. But when the (just as easily processed) Result relation is more plausible ((b) or (d)), subjects strongly prefer to resolve the object pronoun *him* to the non-parallel prior subject *Samuel*, with the results summarized by Kehler (2009:8) above.

Moreover, Transfer of Possession verbs like *give* or *serve to* are more likely to give rise to an interpretation of a subsequent utterance wherein it is the Goal of the transfer (the object) which is more prominent in the subsequent sentence, with a corresponding Goal bias in pronoun interpretation. Rohde et al. (2006) and Kehler et al. (2008) gave experimental subjects examples involving transfer-of-possession context sentences followed by ambiguous pronoun prompts, as in (14):

(14) John handed a book to Bob. He_____.

Here, the subject *John* has a Source thematic role and the object of the preposition *Bob* fills the Goal role. Following Stevenson et al. (1994), Rohde et al. (2006) hypothesized that “entity prominence shifts to the Goal – i.e., the recipient of the object of transfer – at the time the end state is reached. Rohde et al. thus predicted that coherence relations that focus on the end state of the previous eventuality (Occasion, Result) would display a strong Goal bias for pronoun interpretation, whereas relations that focus on other components of event structure (Explanation, Elaboration) would display a strong Source bias (see also Arnold, 2001). This is precisely what they found” (Kehler & Rohde 2017:11).

In a follow-up, Rohde et al. (2007) found that when subjects were asked to respond to a *Why?* question, leading to expectation of an Explanation relation to the response, there was a strong tendency to use a pronominal subject referring to the Agent of the transfer of possession verb, whereas a *What happened next?* question was more likely to elicit subject reference to the Goal.

And in a processing study in Kehler & Rohde (2017) they used materials like these:

- (15) [Source-referring pronoun] Jessica served chili to Emily. **She** explained to Emily...
 a. [WHY] ... in the kitchen that morning that everyone needs to try chili once.
 b. [WHAT-NEXT] ... in the kitchen that night that the secret to chili is real jalapenos.
- (16) [Goal-referring pronoun] Jessica served chili to Emily. **She** explained to Jessica...
 a. [WHY] ... in the kitchen that morning that she can only eat soft foods.
 b. [WHAT-NEXT] ... in the kitchen that night that the chili was a bit too spicy.

They found that when subjects expected in advance a *why* question, they gave evidence of processing difficulties after the highlighted goal-referring pronoun in (16a), but not to the source-referring pronoun in (15a). But with the expectation of a *what-next* question, there was evidence of processing difficulties after the source-referring pronoun in (15b), but not to the goal-referring pronoun in (16b). Together, these three studies argue that **it isn't subject-hood alone that raises the likelihood of resolution of an ambiguous pronoun, but the combination of thematic role and predicate type**, in conjunction with expectation of coherence relation.

But what about evidence that arguments are more likely to be pronominalized when they have a subject antecedent, even when the resulting pronoun is potentially ambiguous? Here's a hunch:

- (17) **Subject pronominalization:** Subjects statistically tend to be correlated with topics (which seems to be the case across languages). Then since topics are more prominent in discourse, i.e. tend to be what the discourse is about (Roberts 2011), subjects tend to be more prominent than non-subjects. But this default holds only when *all other things are equal*. When, as in some of the examples considered above, we do not expect that the subject will continue to be the topic, this tendency is over-ridden.

And here is an interesting question I have not addressed in the preceding: Is it the case that discourse referents introduced by utterance of an NP—hence those that are strongly familiar in Roberts' (2003) terms—are more salient than those that are merely weakly familiar? Consider the following:

- (18) [Context: both interlocutors A and B know Marcia, A's good friend who's an accountant and a lesbian:]
 A: Marcia got married!
 Alternative next utterances:
 A': She's a doctor!
 B: Who is she?
 B': Who's the lucky woman?

In (18), as in Elsner's (11), presumably there are two lucky women (a near-synonym for *bride*), but because we could refer to Marcia by name, one might take *she* in A', or B, or *the lucky woman* in B', to unambiguously refer to Marcia's new wife, especially since the interlocutors know Marcia's an accountant (making Marcia implausible as antecedent for *she* in A') and both know Marcia (so there's no need to ask who Marcia is in B or B'). I have my own intuitions about what's felicitous here, leading me to think that the merely weakly familiar 'Marcia's wife'

is salient enough for ready resolution of all these anaphoric relations. But I don't know of any experimental work which bears on how addressees generally would understand such sequences or judge their relative felicity, and which hence might support or refute my (possibly biased) intuitions.

Compare the “anaphoric island” *a bicyclist* (Postal 1969) in (19):

- (19) John, a bicyclist, always turns on his headlight.
a) It's red.
b) It's a Schwinn.

In (19), John's bicycle is no less weakly familiar than is Marcia's wife in (18). However, in (19a), I think *it* can only refer to the headlight, suggesting that the explicitly mentioned, strongly familiar headlight is more salient than the merely weakly familiar bicycle. I can retrieve the bicycle as antecedent for *it* in (19b)—and in fact that's the only plausible antecedent (since Schwinn is a famous US maker of bicycles, not headlights), but perhaps somewhat awkward. Again, I don't know of experimental research that bears on the question of felicity in such examples.

Summarizing my main claims:

- Our goals, plans and intentions are hierarchically organized, by relative importance and subservience. Ideally, this structure constitutes a strategy for achieving (at least) our most important goals, and, resulting constraints on the order of intending to achieve them.
- **Intention guides attention:** when we are committed to a goal, we attend not only to the goal (e.g. an action to carry out, or the resolution of a question), but also to the entities relevant to its articulation and achievement.
- Goals that are strategically most immediate are those we attend to first, and hence we attend first to the entities relevant to those immediate goals.
- **Salience is constrained by attention** (attentional blindness): The closer one attends to an entity (as a function of one's intentions), the more salient it becomes.

Thus just as there are degrees of attention, there are corresponding degrees of salience.

My underlying assumption is that in a felicitous discourse, one which gives interlocutors the wherewithal to retrieve each others' intended meanings, **the interlocutors know what they're talking** about—the theme, or QUD or topic of conversation. Moreover, they know which entities that topic is about. All theories of anaphora resolution assume that we track such entities, the questions posed, and the information shared about them. What I propose here is that the hierarchical intentional structure which gives coherence to discourse—the strategy of inquiry reflected in various question-answer and rhetorical relations—is central in the determination of salience as that bears on anaphora resolution.

But do keep in mind: Salience isn't so much about something taken into consideration as a constraint after-the-fact on the result of compositional interpretation—for anaphora resolution, disambiguation, domain restriction, etc. Rather, it's a reflection of how **RELEVANCE to the interlocutors' goals prospectively affords the hearer certain understandings as readily**

available and prominent. We don't so much choose between different possibilities as *recognize* the one being promoted—for anaphora resolution, the most obviously contextually salient entity of the appropriate type. For fuller discussion of this general perspective on the role of context in interpretation, see Roberts (2017).

I have no illusions that any of this constitutes the last word on this complex subject. But I do think that considering the intentional structure of an interaction will turn out to consistently play an important role both in the determination of salience and, hence, in anaphora resolution. I hope that this concrete proposal will help guide further experimental and corpus investigation into the problem.

Appendix A: Asher & Lascarides' (2003) discourse relations¹⁵

Highlighting in yellow indicates a role for Intentional Structure, in the sense of Roberts (2013) (not according to A&L). As above, a SARG is a Speech Act Related Goal.

Relation over arguments α , β	Veridical*	Indicative	Interrogative	Directive	add'l notes
Content-level relations					
d ↓, Topic 4.7, 146	✓	subordinating			required for Background, Continuation, Narration
Alternation (dynamic \vee) 4.8.5, 169					
Background 4.8.3, 165 7.6.4, 331	✓	coordinating	subordinating (answer to β is background to α) (12)		temporal overlap b/n α and β like Narration, requires Topic related to it by FBP
Consequence (dynamic \Rightarrow) 4.8.5, 169					
Continuation 4.7, 146	✓	coordinating			like Narration, requires Topic, but lacks spatio-temp'l consequences
Def(easible)-Consequence 4.8.5, 169	?			coordinating (19)	e.g., for bridging to prototypical participants; see (7) below
Elaboration 4.8.1, 159 7.6.4, 331	~✓ (modified Sat.Schema)	subordinating	subordinating (answer to β elaborates on α) (13)	subordinating (18)	temporal-part relation between β and α
Explanation 4.8.1, 159 7.6.4, 331	✓	subordinating	subordinating (why?)		temporal precedence b/n events dual of Result
F(ocus)B(ackground)P(air) 4.8.3, 165	✓	subordinating			only used for semantics of Background
Narration 4.8.2, 162 7.6.4, 331	✓	coordinating	subordinating (and then what?)(14)	coordinating (17)	spatio-temporal consequence (end of α = beginning of β) same-topic
Result 4.8, 155 7.6.4, 331	✓	subordinating [causal]	subordinating [causal]	subordinating α normally results in β being true (20)	
Q(uestion)A(nswer)P(air) 7.6.1, 313	✓		subordinating true direct answer		QUD subordination
Text Structuring relations one of these is required for ellipses per Asher (1993); they may coexist with other relations					
Contrast 4.8.4, 168	✓	subordinating (21)			α , β must have similar semantic structures with contrasting theme. holds with

¹⁵ from their Appendix D, pp.459ff

					Background in (21)
Parallel 4.8.4, 168	✓	subordinating (22)			also requires similar semantic structures, common theme holds with Background in (22)
Cognitive-level relations semantics specified partly in terms of intentions, beliefs of interlocutors					
Acknowledgement 8.4, 361		Subordinating			S(β) has accepted or achieved S(α)'s SARG(s)
I(ndirect)QAPS 7.6.1, 313		Subordinating	IQAP _{r(equest)} Subordinating		α must be a Q, β contextually entailing a direct answer to α
N(ot)E(nough) I(nformation) 7.6.1, 313		Subordinating			α must be a Q, β implying that S(β) can't answer
Plan-Correction 7.6.2, 320		Subordinating			β entails that S(β) won't accept/can't help S(α) achieve her SARG(s)
Plan-Elab 7.6.2, 320	✓	Subordinating (31)	Q-Elab Subordinating (33) Often a type of Elaboration _q	R-Elab Subordinating (34)	β provides info/asks Q/directs to elaborate a plan for achieving the SARG(s) of S(α)
P(artial)QAP 7.6.1, 313		Subordinating	Subordinating		α a Q, β non-monotonically entailing a partial answer
Divergent relations					
Correction 8.3, 345		Subordinating (35), (36)			
Counterevidence 8.2, 343		Subordinating (37)			Like Correction , but only defeasible.
Dis(R) 8.3.1, 350		??			R(α, β), part of discourse context, is now in dispute
Metatalk relations					
Consequence 7.6.5, 333		(38)			If α is true, then S(β) has the SARG of β .
Explanation* 7.6.5, 333	✓	Subordinating (39)	Explanation* _q (40)		β (or its answer) explains why S(α) has α 's SARG
Result 7.6.5, 333		(41)	?	?	e _{α} caused S(α) to utter β

*satisfies the Satisfaction Schema

Their illustrative examples for some of these relations, as noted in the table:

(7) If John scuba dives, he'll bring his regulator.

(12) A: Max arrived at the party at 8pm last night.
B: Who was there at the time?

- (13) A: Kluwer are accepting manuscripts at the moment.
B: What kind of manuscripts?
- (14) A: John arrived at the party at 8pm last night.
B: And then what happened?
- (17) Go into John's office and get a red file folder.
(18) Go to John's office and take a red file folder with you.
- (19) Smoke a packet of cigarettes a day and you will die before you're 50.
(20) Turn left at the roundabout and you will see traffic lights.
- (21) a. John loves sport.
b. But he hates football. [Contrast]
- (22) a. John loves sport.
b. Bill loves sport too. [Parallel]
- (31) A: I want to catch the 10.20 train.
B: It's leaving from platform 1.
- (33) A: Can we meet next weekend?
B: How about next Saturday?
- (34) A: I want to catch the 10.20 train to London.
B: Go to platform 1.
- (35) A: John distributed the copies.
B: No, it was Sue who distributed the copies.
- (36) A: John went to jail. He was caught embezzling funds from the pension plan.
B: No! John was caught embezzling funds, but he went to jail because he was convicted of tax evasion.
- (37) A: John doesn't have a girlfriend.
B: He's been going to New York a lot lately.
- (39) Close the window. I'm cold.
- (40) A: It's getting late.
B: Aren't you enjoying yourself?
- (41) It's getting late. Can we leave now?

Appendix B: Scoreboard of a discourse game, version of Roberts (2022)

A context of utterance is a body of information captured on a scoreboard in the sense of Lewis (1979), as developed in Roberts (1996/2012, 2015), given here with **new detail about G**:

The **scoreboard K for a language game** at time t is a tuple, $\langle I, M, <, CG, QUD, G \rangle$, where:

I is the set of interlocutors at t

M is the set of illocutionary moves made by interlocutors up to t , with distinguished subsets:

$A \subseteq M$, the set of assertions

$Q \subseteq M$, the set of questions

$D \subseteq M$, the set of directions

$Acc \subseteq M$, the set of accepted moves

$<$ is a total order on M , the order of utterance

CG , the common ground, is the set of propositions treated as if true by all $i \in I$ at t .

For all $a \in A \cap Acc$, $a \in CG$.

CG reflects all information about the current state of play in the scoreboard K itself.

$QUD \subseteq Q \cap Acc$, the ordered set of questions under discussion at t , is such that for all $m \in M$ at t :

a. for all $q \in Q \cap Acc$, $q \in QUD(m)$ iff CG fails to entail an answer to q and q has not been determined to be practically unanswerable.

b. QUD is (totally) ordered by $<$.

c. for all $q, q' \in QUD$, if $q < q'$, then the complete answer to q' contextually entails a partial answer to q .

d. for all $q \in QUD$ there is a $g \in G_{com}$ (see below) such that g is the goal of answering Q ,

G is a set of sets of goals in effect at t , such that

for all $i \in I$, there is a (possibly empty) G_i which is the set of i 's publicly evident goals, including those i is publicly committed at t to trying to realize; and

$G = \{ G_i \mid i \in I \}$.

For all $d \in D \cap Acc$, d indexed to interlocutor i , there is a goal g of realizing d such that $g \in G_i$ iff the applicable conditions for i 's realization of d may yet arise and it has not been determined that the realization of d by i is impracticable.

Moreover, for all $i \in I$:

a. for all $g \in G_i$, g is a conditional goal, representing the intention to achieve the goal should certain conditions arise in the actual world at some $t' > t$.

b. i 's priorities are reflected in additional structure(s) over G_i : Some goals sub-serve others, some goals are hierarchically organized into plans, and the way that the agent i prioritizes her goals is reflected in a partial order.

and we can define:

$G_{com} = \{ g \mid \forall i \in I: g \in G_i \}$, the set of the interlocutors' common goals and plans at t .

$G_Q = \{g \in G_{com} \mid \text{there is some } Q \in QUD \text{ and } g \text{ is the goal of answering } Q\}.$

For all $i \in I$, if i is a sincere, competent and cooperative interlocutor in D , we can use G_Q to characterize two kinds of publicly evident goals and plans to which i is committed (at time t):

Discourse Goals of $i = G_Q$

Domain Goals of $i = G_i \setminus G_Q$

$G_{com} \setminus G_Q$: the set of common Domain Goals of all the interlocutors

The *strategy of inquiry* which aims at answering q , $Strat(q)$:

For any question $q \in Q \cap Acc$, $Strat(q)$ is the ordered pair $\langle q, S \rangle$, where S is the set such that:

If there are no $q' \in Q$ such that $QUD(g') = \langle \dots q \rangle$, then $S = \emptyset$.

Otherwise, for all $q' \in Q$, $QUD(q') = \langle \dots S \rangle$ iff $Strat(q') \in S$.

Appendix C: Psycholinguistic work pertaining to the role of intentions in directing attention¹⁶

Eye-tracking studies on anaphora resolution in task-oriented dialogue:

The intentions adopted by a cooperative subject in an experimental task reflect the intentional structure of the experimental interaction, and hence what a subject takes to be at issue (RELEVANT) in interpreting a particular utterance that directs him to perform a given task. I.e., RELEVANCE with respect to a task one intends to perform plays the same role in attentional salience as RELEVANCE to the QUD.

- Chambers, Tanenhaus, Eberhard, Filip & Carlson (2002): subjects dynamically restructure their attentional field as sentence comprehension proceeds, in accordance not only with the visual array, but with task-relevant pragmatic information about the intended referents made available in the utterance itself: “[C]andidate referents are evaluated in terms of their relevance to the immediate task and...this information is used in tandem with linguistic information to incrementally define referential domains,” so that otherwise potential competitors in the visual field are not attended to by subjects when they are pragmatically irrelevant.
- Brown-Schmidt & Tanenhaus (2008): even in unscripted conversation:
...we observed typical lexical competitor effects for expressions uttered by the experimenter outside the context of the conversation...[but] decreased competition from lexical competitors when interpreting expressions within the conversation because of conversationally constrained referential domains. . .[The experimental evidence argued that] two factors—proximity and relevance to the task—did significantly predict whether speakers would modify their expressions with respect to the entire sub-area, suggesting that these factors played a role in the speaker’s decision as to what was in the referential domain... [and that] the addressee interpreted expressions with respect to similarly constrained referential domains. The same factors that predicted whether the speaker

¹⁶ See Roberts (2017) for a broader discussion of how an intentional structure of the sort assumed here bears on the architecture of discussion.

disambiguated his expressions with respect to the competitor blocks predicted whether the addressee fixated these competitors as she interpreted the same expressions.

- Chambers, Tanenhaus & Magnuson (2004), and Tanenhaus, Chambers & Hanna (2004): relevance to a task constrains the referential domain for experimental subjects, as measured by eye gaze.
- Hanna & Brennan (2007): tracking a confederate's gaze can be used not only to recognize, but even to anticipate referential intention.

The role of coherence in children's processing of pronouns:

- Spenader et al (2009): a study of pronoun interpretation in children aged 4;5 – 6;6. Earlier work by Elbourne (2005) had suggested that children at that age have a “pronoun interpretation problem”, but work on production in such children (Bloom et al. 1994, de Villiers et al. 2006) had shown they have almost perfect usage. Spenader et al. argued that the materials for testing pronoun comprehension in Elbourne's experiments were “inherently awkward, either failing to establish any topic or directly contradicting pragmatic rules for topics in their pronominalization”. They conducted experiments where the experimental materials were controlled for coherence and the relevant pronouns “were clearly established as the topic of the target sentence”, with the result that the so-called pronoun interpretation problem disappeared completely, so that even young children “are highly proficient at using pragmatic clues in interpretation.”¹⁷

The role of the QUD in discourse integration:

- Frazier & Clifton (2012): experimental evidence for a central role of the QUD in discourse coherence, favoring interpretations which resolved the QUD over those that did not. This is consonant with the proposals about ellipsis and the QUD of Kehler (2009).

The QUD, Focus and the role of prosody in processing

If we understand prosodic phrasing and prominence to reflect the relevant discourse alternatives (the QUD), as argued since Rooth (1992), then the following studies support Frazier & Clifton's conclusions:

- Speer & Blodgett (2006): critical overview of work on prosody in processing. Many studies support the contention that both phrasing and prominence play a role in syntactic disambiguation and reference resolution.
- Schafer (1997); Blodgett (2004): evidence for a prosody-first model of the role of prosodic phrasing in syntactic disambiguation: the phonological processing builds on an abstract prosodic representation, which serves as input to the syntactic and semantic processors. “Processing is incremental at this level (as at all levels), so it's constantly updated and available to influence processing at other levels.” (Speer & Ito 2006:529). Intonation phrase boundaries trigger wrap-up of any outstanding processing, including interpretation and (for Blodgett 2004) syntactic parsing. Contra the view that first-pass parsing is entirely driven by syntactic factors.

¹⁷ Spenader et al. use the Centering Theory of Grosz et al. (1995) to control for topicality. As discussed above, I don't think that this theory will ultimately suffice to yield a general account of topicality and salience, but that does not undercut the value of their experimental findings: children clearly use coherence and topicality in pronoun resolution.

- Bock and Mazzella 1983; Birch and Clifton 1995, 2002; Cutler 1976; Cutler and Foss 1977; Terken and Nooteboom 1987; Davidson 2001; Ito 2002: “have shown robust effects of intonation [prominence] on discourse comprehension in adults with tasks, such as phoneme detection, discourse verification, and speeded utterance acceptability judgments.” (Speer & Blodgett 2006)

Hypothesis: The early role of prosody in production reflects something even more general than attentional masking: the general orientation of processing towards what is RELEVANT to the QUD.

Corpus studies of anaphora resolution:

- Tetreault & Allen (2004) concluded that some semantic information (about events and situation types, object types, and other content that could be automatically retrieved) significantly improved pronoun resolution algorithm; but Tetreault (2005), considering the intrasentential centering theories of Grosz and Sidner (1986) and Kamayama (1998) concluded that “Our results show that incorporating basic clausal structure into a leading pronoun resolution does not improve performance.”

Difficulties:

- There is no ready way of segmenting the discourse automatically to reflect the text’s QUD structure.
- Grosz & Sidner (1986) did not conceive of the intentional structure of discourse in terms of a structure of questions for discussion, so this particular development of their proposal has not, to my knowledge, been investigated in corpus studies or in the development of algorithms for discourse segmentation.

References:

- Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*. 106(4):748-765.
- AnderBois, Scott, Adrian Brasoveanu, & Robert Henderson (2015). At-issue proposals and appositive impositions in discourse. *Journal of Semantics* 32.1:93-138. doi: 10.1093/jos/_t014.
- Arnold, Jennifer (2001) The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes* 21(2):137-162.
- Asher, Nicholas & Lascarides, Alex (1998a) Questions in dialogue. *Linguistics and Philosophy* 21:237-309.
- Asher, Nicholas & Alex Lascarides (2003) *Logics of Conversation*. Cambridge University Press.
- Beaver, David & Brady Clark (2008) *Sense and Sensitivity*. Blackwell.
- Birch, S. L., and C. Clifton, Jr. (1995) Focus, accent, and argument structure: effects on language comprehension. *Language and Speech* 38:365–91.
- Birch, S.L. & C. Clifton, Jr. (2002) Effects of varying focus and accenting of adjuncts on the comprehension of utterances. *Journal of Memory and Language* 47:571–88.
- Blodgett, Allison (2004) *The interaction of prosodic phrasing, verb bias, and plausibility during spoken sentence comprehension*. Ph.D. dissertation, The Ohio State University.

- Bloom, P., A. Barss, J. Nicol & L. Conway (1994) Children's knowledge of binding and coreference: Evidence from spontaneous speech. *Language* 70, 53–71.
- Bock, J. K. & J. R. Mazzella. (1983) Intonational marking of given and new information: Some consequences for comprehension. *Memory & Cognition* 11.64–76.
- Bratman, Michael E. (1987) *Intentions, Plans, and Practical Reason*. Harvard University Press, Cambridge, MA.
- Brown-Schmidt, S., E. Campana & M. K. Tanenhaus (2005) Real-time reference resolution by naïve participants. In J. C. Trueswell & M. K. Tanenhaus (eds.) *Approaches to Studying World-Situated Language Use*. Cambridge, MA: MIT Press, 2005, 153–171.
- Brown-Schmidt, Susan & Tanenhaus, Michael K. (2008). Real-time investigation of referential domains in unscripted conversation: a targeted language game approach. *Cognitive Science* 32(4):643-684.
- Büring, Daniel (2003) On D-Trees, Beans, and B-Accents. *Linguistics and Philosophy* 26(5):511-545.
- Büring, Daniel (2005) *Binding Theory*. Cambridge Textbooks in Linguistics. Cambridge University Press.
- Chambers, Craig G. & Michael K. Tanenhaus (2004) Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory and Cognition* 30(3):687-696.
- Chambers, C.G., Tanenhaus, M.K & Magnuson, J.S. (2004). Action-based affordances and syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 30, 687-696.
- Chambers, C.G., Tanenhaus, M.K, Eberhard, K.M., Filip, H. & Carlson, G.N. (2002). Circumscribing referential domains in real-time sentence comprehension. *Journal of Memory and Language*, 47, 30-49.
- Clarke, A. D., M. Elsner & H. Rohde (2015) Giving good directions: order of mention reflects visual salience. *Frontiers in Psychology, Special issue on Models of Reference* 6(1793), 1-10.
- Clifton & Frazier (2010) Main Assertion Principle *** [need reference]
- Cohen, Philip R., Jerry Morgan & Martha E. Pollack (eds.) *Intentions in Communication*. MIT Press, Cambridge, MA.
- Cutler, A. 1976. Phoneme-monitoring reaction time as a function of preceding intonation contour. *Perception and Psychophysics* 20.55–60.
- Cutler, A., and D. J. Foss. (1977) On the role of sentence stress in sentence processing. *Language and Speech* 20.1–10.
- Davidson, D. (2001) *Association with focus in denials*. PhD dissertation. Michigan State University.
- Dekker, Paul J. (2012). *Dynamic semantics*. Studies in Linguistics and Philosophy 91. Springer Netherlands.
- de Villiers, J., J. Cahillane. & E. Altreuter (2006) What can production reveal about Principle B? In K. U. Deen, J. Nomura, B. Schulz & B. D. Schwartz (eds.) *The Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition–North America*, 89–100. University of Connecticut Occasional Papers in Linguistics 4.
- Elbourne, Paul D. (2005) *Situations and Individuals*. MIT Press.
- Féry, Caroline & Vieri Samek-Lodovici (2006) Focus projection and prosodic prominence in nested foci. *Language* 82(1):131-150.

- von Stechow, Kai-Uwe (1994) *Restrictions on quantifier domains*. Ph.D. dissertation, UMass, Amherst.
- Frazier, Lyn & Charles Clifton, Jr. (2005) The syntax-discourse divide: Processing ellipsis. *Syntax* 8(2):121-174.
- Frazier, Lyn & Charles Clifton, Jr. (2006) Ellipsis and discourse coherence. *Linguistics and Philosophy* 29(3):315-46.
- Frazier, Lyn & Charles Clifton, Jr. (2010) Imperfect ellipsis. *Syntax* 13(4):279-297.
- Frazier, Lyn & Charles Clifton, Jr. (in progress) Identifying what's under discussion: Introducing and narrowing alternatives during discourse processing. Ms. UMass/Amherst.
- Fukumura, K. & P. G. van Gompel (2010) Choosing anaphoric expressions: Do people take into account likelihood of reference? *Journal of Memory and Language* 62:52-66.
- Ginzburg, Jonathan (1994) An Update Semantics for Dialogue. *Proceedings of the Tilburg International Workshop on Computational Semantics*.
- Ginzburg, Jonathan (2012) *The Interactive Stance: Meaning for Conversation*. Oxford University Press.
- Gleitman, Lila R., David January, Rebecca Nappa, & John C. Trueswell (2007) *Journal of Memory and Language* 57:544-569.
- Gordon, P. C., B.J. Grosz, & L.A. Gilliom (1993) Pronouns, names, and the centering of attention in discourse. *Cognitive Science* 17:311-347.
- Gordon, P. C., & K. Scarce (1995) Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory and Cognition* 23:313-323.
- Groenendijk, Jeroen & Martin Stokhof (1991) Dynamic predicate logic. *Linguistics and Philosophy* 14.1:39-100. doi:10.1007/BF00628304.
- Grosz, Barbara J. (1977) *The Representation and Use of Focus in Dialogue Understanding*, Technical Report No. 151, Artificial Intelligence Center, SRI International, Menlo Park, CA.
- Grosz, Barbara, & Candice Sidner (1986) Attention, intentions, and the structure of discourse. *Computational Linguistics* 12:175-204.
- Grosz, Barbara, Scott Weinstein & Aravind Joshi (1995) Centering: A framework for modelling the local coherence of discourse. *Computational Linguistics* 21:203-225.
- Halliday, Michael (1967) Notes on transitivity and theme in English: Parts 1 and 2. *Journal of Linguistics* 3, 37-81, 199-244.
- Hanna, Joy E. & Susan E. Brennan (2007) Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *Journal of Memory and Language* 57:596-615.
- Heim, Irene (1982) *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. dissertation, University of Massachusetts, Amherst.
- Heim, Irene (1983) On the projection problem for presuppositions. WCCFL 2, Stanford University, 114-126.
- Hinrichs, Erhard (1985) *A Compositional Semantics for Aktionsarten and NP Reference in English*. Ph.D. dissertation, The Ohio State University.
- Johnson, E. & P. W. Jusczyk (2001) Word segmentation by 8 month olds: when speech cues count more than statistics. *Journal of Memory & Language* 44:548-67.
- Kamayama, Megumi (1998) Intrasentential centering: A case study. In M. A. Walker, A.K. Joshi & E.F. Prince (eds.) *Centering Theory in Discourse*. Oxford University Press.

- Kamp, Hans (1981) A Theory of Truth and Semantic Representation. In Jeroen Groenendijk, Theo M. V. Janssen and Martin Stokhof (eds.) *Formal Methods in the Study of Language, Vol. I*, Mathematische Centrum, Amsterdam. Reprinted in Groenendijk, Janssen and Stokhof (eds.), 1984, *Truth, Interpretation and Information*, Foris, Dordrecht.
- Kamp, Hans & Uwe Reyle (1993) *From Discourse to Logic: Introduction to the modeltheoretic semantics of natural language*. Springer, Netherlands.
- Kaufmann, Magdalena & Stefan Kaufmann (2012) Epistemic particles and performativity. *Proceedings of SALT 22*:208-225.
- Kehler, Andrew (2002) *Coherence, Reference, and the Theory of Grammar*. CSLI Publications, Stanford, CA.
- Kehler, Andrew (2009) Ellipsis and anaphora in a QUD model of discourse. Talk presented at the University of Michigan Workshop in Philosophy and Linguistics.
- Kehler, Andrew, Laura Kertz, Hannah Rohde & Jeffrey L. Elman (2008) Coherence and coreference revisited. *Journal of Semantics (Special Issue on Processing Meaning)* 25(1):1-44.
- Kehler, Andrew & Hannah Rohde (2017) Evaluating an Expectation-Driven QUD Model of Discourse Interpretation. *Discourse Processes* 54(3), 219-238.
- Kertz, Laura, Andrew Kehler & 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.
- van Kuppevelt, Jan (1995) Discourse structure, topicality and questioning. *Journal of Linguistics* 31(1):109-47.
- van Kuppevelt, Jan (1996) Inferring from Topics: Implicatures as topic-dependent inferences. *Linguistics and Philosophy* 19(4):393-443.
- Lewis, David (1979) Scorekeeping in a language game. In R. Bäuerle, U. Egli, and A. von Stechow (eds.) *Semantics from a Different Point of View*. Springer, Berlin.
- Martin, Scott (2011) Weak Familiarity and Anaphoric Accessibility in Dynamic Semantics. Proceedings of the 16th Conference on Formal Grammar, Ljubljana, Slovenia, 2011. To appear in *Lecture Notes in Computer Science*.
- Memmert, Daniel (2006) The effects of eye movements, age, and expertise on inattention blindness. *Consciousness and Cognition* 15:620-627.
- Nouwen, Rick (2003) Complement anaphora and interpretation. *Journal of Semantics* 20.1:73-113.
- Poesio, Massimo & Barbara DiEugenio (2001) Discourse structure and anaphoric accessibility. In Ivana Kruijff-Korbayová & Mark Steedman (eds.) *Proceedings of the ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*.
- Poesio, Massimo & Hannes Rieser (2011) An incremental model of anaphora and reference resolution based on resource situations. *Dialogue & Discourse* 2.1:235-277. doi: 10.5087/dad.2011.110.
- Polanyi, Livia (1985) A theory of discourse structure and discourse coherence. In P.D. Kroeber, W.H. Eilfort & K.L. Peterson (eds.) *Papers from the General Session at the 21st Regional Meeting of the Chicago Linguistics Society*.
- Portner, Paul (2007) Imperatives and Modals. *Natural Language Semantics* 15(4), 351-383.
- Postal, Paul (1969) Anaphoric islands. *Proceedings of the Chicago Linguistic Society* 5:205-239.
- Prince, Ellen (1981) On the interfacing of indefinite-*this* NPs. In Joshi, Webber & Sag (eds.) *Elements of Discourse Understanding*, Cambridge U. Press, 231-250.

- Reinhart, Tanya (1983) Coreference and Bound Anaphora: A Restatement of the Anaphora Questions. *Linguistics and Philosophy* 6:47–88.
- Roberts, Craige (1989) Modal Subordination and Pronominal Anaphora in Discourse. *Linguistics and Philosophy* 12:683-721.
- Roberts, Craige (1995) Domain Selection in Dynamic Semantics. In Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee (eds.) *Quantification in Natural Languages*, Kluwer, Dordrecht, 661-700.
- Roberts, Craige (1996/2012) Information Structure: Toward an integrated theory of formal pragmatics. In Jae Hak Yoon and Andreas Kathol (eds.) *OSUWPL Volume 49: Papers in Semantics*, 1996. The Ohio State University Department of Linguistics. In press, with a new afterword, in *Semantics and Pragmatics*.
- Roberts, Craige (1996b) Anaphora in intensional contexts. In Shalom Lappin (ed.) *Handbook of Semantics*. Blackwell.
- Roberts, Craige (1998) Focus, the Flow of Information, and Universal Grammar. In Peter Culicover and Louise McNally (eds.) *The Limits of Syntax*, Academic Press, pp.109-160
- Roberts, Craige (1998b) The Place of Centering in a General Theory of Anaphora Resolution. In Marilyn A. Walker, Aravind K. Joshi & Ellen F. Prince (eds.) *Centering Theory in Discourse*, Oxford University Press, pp.359-400.
- Roberts, Craige (2002) Demonstratives as definites. In Kees van Deemter and Roger Kibble (eds.) *Information Sharing: Reference and Presupposition in Language Generation and Interpretation*, CSLI Press, pp.89-196.
- Roberts, Craige (2003) Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26:287-350.
- Roberts, Craige (2004a) Discourse context in dynamic interpretation. In Laurence Horn and Gregory Ward (eds.) *Handbook of Contemporary Pragmatic Theory*, Blackwell, pp.197-220.
- Roberts, Craige (2004b) Pronouns as definites. In Anne Bezuidenhout and Marga Reimer (eds.) *Descriptions and Beyond*, Oxford University Press, pp.503-543.
- Roberts, Craige (2010) Retrievalability and incomplete descriptions. Ms., The Ohio State University. <http://www.ling.ohio-state.edu/~croberts/Retrievalability.pdf>.
- Roberts, Craige (2011) Solving for interpretation. Talk at the Workshop on Meaning and Understanding at the Centre for Advanced Study, Oslo. [This unpublished talk, cited by Westera, includes a broad review of the work on anaphora resolution, some of that summarized here in Appendix C.]
- Roberts, Craige (2012a) Information Structure: Toward an integrated theory of formal pragmatics. *Semantics and Pragmatics* 5.6:1-69. (Invited as a classic unpublished paper.) <http://dx.doi.org/10.3765/sp.5.6>. Revision of Roberts (1996).
- Roberts, Craige (2012b) Topics. In press in Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.) *Semantics: An International Handbook of Natural Language Meaning*. Mouton de Gruyter.
- Roberts, Craige (2013) Information Structure: Afterword, with bibliography of related work. *Semantics and Pragmatics* 5.7:1-19. <http://dx.doi.org/10.3765/sp.5.7>.
- Roberts, Craige (2015) Accommodation in a language game. In Barry Loewer & Jonathan Schaffer (eds.) *A Companion to David Lewis*. Wiley, Hoboken, NJ, pp.345-366.
- Roberts, Craige (2017) Linguistic convention and the architecture of interpretation. *Analytic Philosophy* 58.4:418-439.

- Roberts, Craige (2018) Speech acts in discourse context. In Daniel Fogal, Daniel Harris & Matt Moss (eds.) *New Work on Speech Acts*. Oxford University Press.
- Roberts, Craige (2019) Contextual influences on reference. In B. Abbott & J. Gundel (eds.) *Oxford Handbook on Reference*. Oxford University Press.
- Roberts, Craige (2021) Modal Subordination: “It would eat you first!” In Daniel Gutzmann, Lisa Matthewson, Cécile Meier, Hotze Rullmann & Thomas Ede Zimmermann (eds.) *The Wiley Blackwell Companion to Semantics*. Wiley. DOI: 10.1002/9781118788516.sem102
- Roberts, Craige (2022) Imperatives in a dynamic pragmatics. Ms., OSU and NYU, re-submitted for review.
- Roberts, Craige, Mandy Simons, David Beaver & Judith Tonhauser (2009) Presupposition, Conventional Implicature, and Beyond: A unified account of projection. In *Proceedings of the Workshop on Presupposition, ESSLLI '09, Bordeaux, July, 2009*.
- Rohde, Hannah (2008) *Coherence-driven effects in sentence and discourse processing*. Ph.D. dissertation, University of California at San Diego.
- Rohde, Hannah & Michael C. Frank (2011) Markers of Discourse Structure in Child-Directed Speech. To appear in the proceedings of the 33rd annual conference of the cognitive science society, Boston, MA.
- Rohde, Hannah & 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, 2008.
- Rohde, Hannah & Andrew Kehler (2008b) Coherence-driven expectations in discourse and dialog. Poster presented at the 21st Annual CUNY Conference on Human Sentence Processing, March, 2008.
- Rohde, Hannah, & Andrew Kehler (2014) Grammatical and information-structural influences on pronoun production. To appear in *Language, Cognition and Neuroscience*.
- Rohde, Hannah, Andrew Kehler & 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*, Vancouver, Canada, 697-702.
- Rohde, Hannah, Andrew Kehler & 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*, Memphis, TN, 617-622.
- Rooth, Mats E. (1992) A theory of Focus Interpretation. *Natural Language Semantics* 1:75-16.
- Schafer, A. J. (1997) *Prosodic parsing: The role of prosody in sentence comprehension*. Ph.D. dissertation, UMass/Amherst.
- Schwarzschild, Roger (1999) GIVENness, AvoidF and other Constraints on the placement of accent. *Natural Language Semantics* 7.2:141-177.
- Schlenker, Philippe (2005) Non-Redundancy: Towards a semantic reinterpretation of Binding Theory. *Natural Language Semantics* 13(1):1-92).
- Simons, Daniel J. & Christopher F. Chabris (1999) Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception* 28:1059-1074.
- Simons, Mandy, David Beaver, Craige Roberts & Judith Tonhauser (2017) The Best Question: Explaining the projection behavior of factives. *Discourse Processes* 54.3:187-206, a special issue on the Question In Discussion. Available open-access at: <http://www.tandfonline.com/doi/pdf/10.1080/0163853X.2016.1150660?needAccess=true>
- Smyth, R. (1994) Grammatical determinants of ambiguous pronoun resolution. *Journal of Psycholinguistic Research* 23:197-229.

- Solan, L. 1980. Contrastive stress and children's interpretation of pronouns. *Journal of Speech and Hearing Research* 23.688–98.
- Speer, Shari & Allison Blodgett (2006) Chapter 13: Prosody. In Matthew Traxler & Morton Ann Gernsbach (eds.) *Handbook of Psycholinguistics*, 2nd edition. Elsevier.
- Speer, Shari & Kiwako Ito (2009) Prosody in first language acquisition. *Language and Linguistics Compass* 3(1):90–110.
- Spencer, Jennifer, Erik-Jn Smits & Petra Hendriks (2009) Coherent discourse solves the pronoun interpretation problem. *Journal of Child Language* 36.1:23-52.
- Sperber, D. and D. Wilson (1986) *Relevance: Communication and Cognition*. Harvard University Press, Cambridge, MA.
- Stevenson, Rosemary J., Rosalind A. Crawley & David Kleinman (1994) Thematic roles, focus, and the representation of events. *Language and Cognitive Processes* 9:519-548.
- Tanenhaus, Michael K., Craig G. Chambers. & Joy E. Hanna (2004). Referential domains in spoken language comprehension: Using eye movements to bridge the product and action traditions. In J.M. Henderson & F. Ferreira (Eds.), *The Interface of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press, pp. 279-318.
- Terken, Jacques, & Julia Hirschberg (1994) Deaccentuation of words representing “given” information: Effects of persistence of grammatical function and surface position. *Language and Speech* 37(2):125-145.
- Terken, J. & S. G. Nootboom (1987) Opposite effects of accentuation and deaccentuation on verification latencies for given and new information. *Language and Cognitive Processes* 2:145–63.
- Tetreault, Joel R. (2005) Decomposing Discourse. In Antonio Branco, Tony McEnery & Ruslan Mitkov (eds.) *Anaphora Processing*. John Benjamins.
- Tetreault, Joel R. & James F. Allen (2004) Dialogue Structure and Pronoun Resolution. In *Proceedings of the Discourse Anaphora and Anaphora Resolution Colloquium (CDAARC'04)*, Azores, September 23-24, 2004.
- de Villiers, J., J. Cahillane & E. Altreuter (2006) What can production reveal about Principle B? In K. U. Deen, J. Nomura, B. Schulz & B. D. Schwartz (eds.) *The Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition–North America*, 89–100. University of Connecticut Occasional Papers in Linguistics 4.
- Westera, Matthijs (2017) *Exhaustivity and Intonation: A unified theory*. Ph.D. dissertation, University of Amsterdam.
- Westera, Matthijs & Hanna Rohde (2019) Asking between the lines: Elicitation of evoked questions in text. Amsterdam Colloquium.