Analytic Philosophy Vol. 58 No. 4 December 2017 pp. 418–439

LINGUISTIC CONVENTION AND THE ARCHITECTURE OF INTERPRETATION[1]

CRAIGE ROBERTS
The Ohio State University

Ernie Lepore and Matthew Stone (L&S), in *Imagination and Convention* (Oxford University Press 2015) offer critical analyses and comparisons of three major research programs in pragmatics—those of Larry Horn, Steve Levinson, and Dan Sperber & Deirdre Wilson. Lepore and Stone use these accounts, and the general Gricean approach they represent, as a foil for their own theoretical vision, which they call DIRECT INTENTIONALISM. Two themes run throughout their work:

- "[M]eaning is a matter of conventions, and listeners normally recover the meanings of utterances by recognizing the conventions involved, not by reasoning about the speaker in any deeper sense" (p. 199). Accordingly, the set of "possibilities a particular language allows for organizing discourse into patterns of inquiry, argument, and negotiation… goes beyond syntax and semantics as usually conceived, but… is part of the speakers' linguistic grammar nonetheless" (p. 92). They argue that a very large number of the types of implications that Grice and his followers have taken to be conversational implicatures instead arise largely as a function of conventional content; these implications are thus not non-detachable, they argue, and so not conversational at all.
- There remain aspects of utterance meaning which *do* clearly go beyond what is given by "linguistic grammar," even so broadly conceived; these include metaphorical meanings, sarcasm, irony, humor, and hints. But still, they argue, in recognizing these aspects of the meaning of a particular utterance, Grice's Cooperative Principle (CP) and maxims play no role at all. Instead, these types of meanings draw on the imaginations of both speaker (in generation) and addressee (in interpretation), involving a wide range of types of eclectic inferences. These enrichments of meaning are thus creative and improvisatory rather than regular and conversational in Grice's sense.

[1] Many thanks to B. Chandrasekaran, for extensive discussion and comments on all aspects of this essay, and to Shari Speer, for discussion of some of the relevant literature in psycholinguistics.

Analytic Philosophy Vol. 58 No. 4

On the basis of these claims L&S claim **that there is no place in pragmatic theory for Gricean conversational implicature** as understood in the theories they discuss, and ultimately that the cooperative principle and maxims play no important part in pragmatic theory.

One of their central complaints about the Gricean program derives from Grice's requirement that any conversational implicatures which are part of the meaning of a given utterance must be calculable on the basis of what is said, background knowledge, and the cooperative principle and maxims.

> By saying *p*, the speaker conversationally implicates that *q* just in case she is presumed to be following the maxims or at least the Cooperative Principle; and the supposition that she is aware or thinks that *q* is required in order to maintain that she is following these maxims; and she thinks (and would expect her audience to think that she thinks) that it is within the competence of her audience to work out or grasp intuitively, that the supposition that *q* is required. (Grice 1967)

Grice consistently insists that this calculation must take place on the basis of 'what is said'. Since 'what is said' certainly is tightly constrained by the conventional content of the utterance—roughly the 'literal content' of the expression uttered—this suggests that Grice took the determination of conventional content—the rule-governed part of meaning determination—to be in some sense prior to the determination of conversational implicature.

In the way it is most commonly understood in pragmatics, this priority is reflected in how meanings are calculated: First comes word recognition, then parsing and compositional semantics (perhaps determining syntactic structure and sense in parallel, rule-by-rule); and finally, having thereby derived the compositional content of the expression uttered, the result serves as the basis for reasoning to derive any intended implicatures. Implicatures based on Gricean reasoning are thus icing on the cake, enriching and modifying the truth conditional content conventionally retrieved. This common conception was reflected in the sausage machine model of parsing in interpretation (Frazier and Fodor 1978) that was influential in psycholinguistics in the 1980s: parsing takes place in two stages, via informationally encapsulated modules which make no appeal to semantic, pragmatic or discourse processes, first assigning lexical and phrasal nodes to the results of a word recognition module, then combining these to form a complete phrase marker. The result is packaged and then passed along to a semantic module for interpretation, and the result of that process is finally the basis for application of some pragmatic reasoning based on the proposition expressed ("what was said"), background knowledge, and general pragmatic principles like those of Grice.

The Gricean principles—Quality, Relevance, Quantity, and Manner—are all intended by Grice to be understood as reflexes of the Cooperative Principle, and hence presuppose that meaning conveyance itself is collaborative in that the interlocutors at least conspire to make their semantic intentions recognizable, and that the reasoning that derives implicatures involves the hearer's assumptions about what the speaker intended. The speaker, for her part, makes the utterance expecting this collaborative reasoning on the basis of the conventional content. This is very different from parsing as the latter is usually understood, which is based on purely grammatical principles, operating independently of interlocutors' intentions.

Lepore and Stone (p. 199) object to this general picture:

> We deny... that collaborative reasoning plays anything like the role in semantics and pragmatics that Grice argues for... [M]eaning is a matter of conventions, and listeners normally recover the meanings of utterances by recognizing the conventions involved, not by reasoning about the speaker in any deeper sense.

and (p. 83):

> Pragmatics can be, at most, a theory of disambiguation; pragmatic reasoning never contributes content to utterances.

Thus (p. 88):

> where alternative approaches have postulated pragmatic processes of enrichment, what's really going on is disambiguation: finding the right reading of the utterance, understood as a grammatically specified pairing of form and meaning.

I think L&S and I would agree on many points. The best pragmatic theory is (a) the most constrained theory that (b) makes the correct predictions about attested meanings with (c) the fewest assumptions —preferably all independently motivated, and (d) in the most perspicuous fashion. Though semantics owes us an explanation of the evident gap between what people say and what they mean, we want to avoid Wild West pragmatics, an unconstrained theory which lacks systematicity and predictive power. In this connection, it is clear that positing the calculation of implicatures *in the way illustrated by Grice and his successors* as part of the ordinary process of determining linguistic meaning is cognitively implausible and intractable: We produce and understand in real-time, quickly and efficiently. If all goes well, there is no evidence of time-consuming postcompositional

interpretive inferences, and there is a great deal of evidence that such inferences are processed in the course of compositional interpretation.[2] In addition, ideally we want to avoid open-ended, free pragmatic enrichment in the determination of what is meant, since this arguably over-generates wildly: at what point is the inferential process called to a halt? So I am sympathetic to L&S' strong stand in favor of a tightly constrained approach to interpretation, one which is both psychologically plausible and computationally tractable.[3]

In contrast to the approach they criticize, L&S propose a different account of how interpretation proceeds, the DIRECT INTENTIONALISM sketched in §IV of the book. On this view, setting aside "imaginative" interpretations like those involving metaphor and sarcasm, the purported implicatures that yield speech act determination, discourse

[2] Recent experimental studies on scalar implicatures (SI) converge on the conclusion that these are drawn in real-time during the course of processing. In cases where such implicatures play a role in directing attention to an intended referent in a visual array, fixation on the SI-indicated referent takes place prior to the utterance of the target NP and completion of compositional interpretation (Huang and Snedeker 2009, 2011). The calculation of an SI may be cognitively costly, slowing down processing by as much as 400 msec (Bott and Noveck 2004; Bott, Bailey, and Grodner 2012; Katsos et al. 2005; Huang and Snedeker 2009, 2011), though this is controversial (Sedivy 2003, Grodner et al. 2010; Breheny, Ferguson, and Katsos 2012). The reported delays are consistent with Newell's (1990:122) characterization of the "time scale of human action," according to which it takes approximately 100 ms to conduct a cognitive action like drawing an inference, and with the evidence (e.g., see Allopenna, Magnuson, and Tanenhaus 1998) that it takes about 150 msec to plan and make a saccade after the required inference is drawn (and hence shift gaze to the inferred target). Closely related experimental work by Atanassov, Schwarz, and Trueswell (2013) on NOT-*must* implicatures associated with use of *might* found the same delay in processing as that found by Huang and Snedeker (2009). Other studies argue that SIs are neither automatic nor default, but only take place in response to appropriate contextual factors (Katsos and Cummins 2010), like the Q(uestion)U(nder)D(iscussion) (Tian, Breheny, and Ferguson 2010). And ERP studies (Politzer-Ahles et al. 2013) on utterances whose meanings involve SIs argue that "inferential pragmatic aspects of meaning are processed using different mechanisms than lexical or combinatorial semantic aspects of meaning, that inferential pragmatic meaning can be realized rapidly, and that the computation of meaning involves continuous negotiation between different aspects of meaning," supporting an incremental theory of processing where semantics and pragmatics interact (see below).
Ito and Speer (2008) study the use of contrastive accent cues to permit hearers to anticipate upcoming referents in a visual array via implication of a contrast set; they argue that, like the SIs, this effect takes place early and rapidly, well in advance of confirming lexical information from a target NP; moreover, (Ito and Speer 2008, 2011) infelicitous use of contrastive accent results in slower detection of the correct target, a sort of prosody-driven garden-path effect. Again, such processing is "guided not only by the discourse context, but also by the task-relevant referential context of the visual field" (Ito and Speer 2011:86).
If other kinds of implicatures (e.g. relevance) involve similar kinds of inferential mechanisms to those observed in the SI and prosody studies, it is reasonable to assume that they also take place in real-time during semantic compositional processing, rather than postcompositionally.

[3] I think parallel considerations bear on production, but my focus here, as in L&S, will be on processing and interpretation.

coherence effects, anaphora resolution, and scalar implicatures all arise from the operation of grammatical rules. Interpretation is a kind of action which requires no regular inferential processes. Instead, L&S argue, it is a function of direct, Gibsonian affordances (Gibson 1979), "effects that agents can bring about just in virtue of the kind of being that they are and the kind of engagement they have with the world". In motor control, when we grasp an object, we do not *reason* about what we are about to do "despite the complex perceptual and motor activity involved" (p. 208), we just do it because that is the sort of being we are. Similarly, they claim, interpretation does not normally involve drawing inferences; it is directly a function of our linguistic competence. I take it that whether all this is so is an empirical question, but let us assume for the purposes of argument that that is the case.

But L&S argue for two other assumptions, and with these I will take issue: (a) that anaphora resolution (or domain restriction, or disambiguation more generally) takes place without access to non-linguistic information of the sort encoded in the Common Ground, and further (b) that this precludes any role for Gricean principles.

In fact, one can characterize L&S's alternative to the Wild West post-compositional model as the proposal that we build a bigger stockade—the grammar—and bring all the "horses"—the meanings that neo-Griceans take to be generated by reasoning—inside, in the sense of claiming that they are conventionally generated: L&S would put *into the grammar of a particular language* rules which most linguists would consider pragmatic: those for anaphora resolution in discourse, etc.

But there is a third model worth considering, also constraint-based, but not assuming the sequence of encapsulated processes that we find in the sausage machine. This is that the architecture of interpretation involves the coordination of input from multiple sources of information—not all of them part of the linguistic competence realizing the grammar for the language in question. On this model, all of these factors come to bear on interpretation in parallel.

The Third Model: Parallel Bottom-Up and Top-Down Processing

We know that in other kinds of complex cognitive systems, pragmatic reasoning based on background knowledge and expectations takes place in parallel with bottom-up processing of percepts. For example, according to contemporary theories of vision (especially the work building on Marr 1982; e.g. the summary in Shimojo, Paradoso, and Fujita 2001, and the overview and introduction in Smith and Kosslyn 2006), rapid processing of the distribution of image intensity on the retina produces percepts; this production itself is arguably accomplished via a set of parallel processes, as required to accomplish this

sophisticated task in real-time. This process is bottom-up, in the sense of being driven by the sensory stimulus, and it is largely cognitively impenetrable in Fodor's (1983) sense—neither accessible to conscious introspection nor affected by explicit reasoning. But there is also evidence of real-time constraints brought to bear on this bottom-up process by top-down selective attention and expectations, both based on contextually available information, goals, and conceptual sets (e.g. Allport 1989; Balcetis and Dale 2007; Corbetta and Shulman 2002). Recent work by Fecteau and Munoz (2006) present evidence that the two kinds of processes yield distinct neural signals (p. 387). See especially the work on *inattentional blindness* (Simons and Chabris 1999), wherein subjects regularly fail to consciously notice otherwise salient phenomena in a visual field if those are irrelevant to what the subjects are attending to. These constraints influence bottom-up pattern recognition, using information which is (at least in part) non-perceptually derived. That is not to say that this top-down influence itself is cognitively penetrable; there is no evidence that it involves conscious reasoning, for example; and according to Fecteau and Munoz, the two processes are at least partially concurrent in the pre-attentive phase, that which precedes conscious attention. Bottom-up processing, based on salience, seems to proceed in a fairly automatic, rapid way, entertaining whatever comes to attention as a function of visual salience, up to a point. There is evidence that unattended stimuli (both words and visual objects) do have priming properties, so that they are *perceived* at the first stage of processing of stimuli (Mack 2003; Mack and Rock 1998). But according to the contingent capture model of preattentive processing, a person's "current intentions and/or goals affect the speed and efficiency of preattentive processing", so that those stimuli that match what one is looking for "will be processed faster at the pre-attentive stage and will be more likely to be selected for attentive processing" (Folk and Remington 2006).

Hence, the top-down processes in vision clearly draw on background knowledge, goals, and rational expectations, unlike the sensory input itself. The way in which the two kinds of processes work together lends far greater speed and accuracy to visual processing; the expectations help to constrain, from the outset, the set of reasonable "parses" of the purely perceptual information that have to be entertained. But the bottom-up production of percepts is the stronger constraint on out-put—it is these that trigger processing after all. In experiments involving visual perception in monkeys, "salience and relevance yield distinct neural signals—salience is reflected in the initial registration of the target, and relevance is reflected in the elevated activity following the predictive cue" (Fecteau and Munoz 2006:387).

Given this common feature of human cognitive architecture (see recent work on attentional selection across cognitive domains in Schneider, Einhäuser, and Horstmann 2013), where parallel bottom-up and top-down processes converge for greater speed and accuracy, we

might expect to find similar bottom-up and top-down parallel processes in the real-time course of processing linguistic input. And in fact there is now ample experimental evidence that linguistic interpretation does involve this type of parallel architecture. To illustrate this, I will focus on just one of the types of phenomena L&S consider in some depth, anaphora resolution.

Anaphora Resolution in Real-Time

We now know that a hearer's expectations based on a task under discussion very strongly constrain the referential domain that is understood to be salient in the conversation, and thereby help to very quickly resolve anaphora as intended by the speaker. The psycholinguistic work with eye-trackers in visual world paradigms is especially convincing in this regard. Chambers et al. (2002) found that 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 et al. (2008) found that 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 disambiguated his expressions with respect to the competitor blocks predicted whether the addressee fixated these competitors as she interpreted the same expressions.

Chambers, Tanenhaus, and Magnuson (2004), and Tanenhaus, Chambers, and Hanna (2004) found that relevance to a task constrains the referential domain for experimental subjects, as measured by eye gaze. And in production studies of child-directed speech, Rohde and

Frank (2011) found that "Speakers use reduced referring expressions such as pronouns when topical entities are easily retrievable and listeners show signs of engaging in joint attention to entities that have become part of the common ground."

Thus, subjects whose attention has been directed by the task at hand to a single candidate antecedent typically completely ignore competitors, which argues that the pragmatic direction of focus in the best case serves to narrow the "referential domain" so effectively that no after-the-fact disambiguation (resolution) on the basis of purely syntactic factors is called for at all. Crucially, this narrowed joint attention arises as a function of (a) practical features of the task at hand, and (b) relevance as given by common ground (as represented in the visual array, as well as in knowledge about the type of task at hand). But these features are banned from L&S' Conversational Record; they explicit deny that anything like the Common Ground plays a role in that essential background for interpretation, and thus preclude any role for it in meaning derivation.

Consider L&S' discussion of how anaphora is resolved in a "rule-governed" fashion across discourse. They acknowledge that unspecified "pragmatic principles" do play a crucial role in interpretation, but only in a very limited way:

> As when we disambiguate speech acts, we may need pragmatic principles to recognize the preferred resolution of anaphora and presupposition. But accounts of pragmatic inference err when they attribute the specifics of interpretation to the action of the principles themselves. The content of interpretation must be licensed by the conventional, rule-governed dynamics of discourse anaphora. (p. 93)

So what are these rules which govern the "dynamics of discourse anaphora" and thereby "license" anaphora resolution? As elsewhere, L&S are thin on details about the particular rules they assume; but they repeatedly mention two kinds of factors governing discourse anaphora: principles of Centering Theory, and rules about "discourse coherence".

Centering Theory (Grosz, Joshi and Weinstein 1983, 1995; see Walker, Joshi, and Prince (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 Centering does not play the independent role in anaphora resolution that L&S seem to assume, and that to the extent that such principles are applicable, they are instead subordinate to a requirement of relevance to task. Gordon, Grosz, and 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), considering a variety of realizations of Centering Theory, 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 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 strong preference for subject antecedents. And while Tetreault and 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 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 and 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 be 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-Schmit 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 and 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 have not already been established by coherence assumptions, or by previous intention recognition processes**" [my emphasis, 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.

The other type of conventional rule-governed element of discourse to which L&S appeal for anaphora resolution is coherence relations, and in particular the rhetorical relations which play a prominent role in discourse coherence (Asher and Lascarides 2003; Kehler 2002; etc.). These semantic relations between adjacent utterances, along with other features of discourse coherence, do clearly play a role in anaphora resolution. In a suite of experiments (Kertz, Kehler, and Elman 2006; Rohde, Kehler, and Elman 2006, 2007; Kehler et al. 2008; Rohde and Kehler 2008a,b), Kehler 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, and Elman (2006):

(1) Samuel threatened Justin with a knife, and
  a . . .Erin blindfolded him (with a scarf) [Parallel]
  b . . .Erin stopped him (with pepper spray) [Result]
  c . . .he blindfolded Erin (with a scarf) [Parallel]
  d . . .he alerted security (with a shout) [Result]

When the follow-up stands in a Parallel rhetorical relation with the first conjunct, we find the expected parallel thematic roles, as in (1a)

and (1c). But when the (just as easily processed) Result relation is more plausible, as in (1b) and (1d), subjects prefer to resolve the object pronoun *him* to the non-parallel prior subject *Samuel*:

> 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. (Kehler 2009:8)

But Kehler (2009) takes the analysis one step further. He points out that in their experimental materials Rohde, Kehler, and Elman (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 Question-Under-Discussion (QUD)-based discourse structure. Hence, the relation of the target utterance to the QUD, reflecting the speaker's adopted strategy, is the central factor in predicting anaphora resolution:

> . . .at any point during comprehension the hearer will have expectations about how the discourse will be continued with respect to coherence, and. . .the difficulty in interpreting the linguistic material to follow will be conditioned in part on those expectations. These expectations will then evolve based on subsequent linguistic input. (Kehler 2009)

Again, we find that expectations, here based on prior discourse structure, play a crucial role in anaphora resolution. Roberts (1996/2012) argues that the QUD reflects the interlocutors' immediate joint goal in discourse—to address that question—and that, in turn, this goal is subordinate to any overarching goals and intentions, such as those associated with joint tasks. From that perspective, the findings of Kehler and associates are consistent with those of Tanenhaus and associates about the expectation-driven resolution of anaphora on the basis of common ground and task, and with the results of Allen, Poesio, Tetreault et al. showing that Centering principles, insofar as they are useful, are subordinate to task structure.

There is another type of conventional element which plays a strong role in anaphora resolution which L&S fail to mention in this regard (though it is discussed with respect to information structure, see Chapter 8). This is prosodic focus. Consider the well-known example-types (Lakoff 1971):

(2) Julie said Alice was a socialist, and then she INSULTED her.
(3) Julie said Alice was a socialist, and then SHE insulted HER.

One gets different truth conditions for the second sentences of these two string-identical utterances, with *she* in (2) coreferential with *Julie*, *her* with *Alice*, the opposite resolution in (3). We can explain this difference straightforwardly on the basis of the different prosodic prominences in the two—with the pronouns unaccented in (2), accented in (3)—and an independently motivated semantics and pragmatics of prosodic focus.

Rooth (1992) argues that the prosodic focus associated with an utterance conventionally presupposes a certain set of alternatives. Roughly, we abstract on the focused element(s) in a given constituent 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. In the second conjunct of (2), we derive the set {*she Red her*, *R* a two-place relation}, while in (3) we get {*x insulted y*: *x* and *y* individuals}. 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, and 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:

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

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

Hence we take the alternative sets predicted by Rooth for the second conjuncts of (2) and (3) to be the questions presupposed. Moreover, since both conjuncts in a conjunction must address the same question, it is natural to resolve *she* in (2) to Julie, *her* to Alice, yielding the question 'what did Julie do to Alice?'. But in (3) the presupposed question is 'who insulted who?'; then this implies that calling someone a socialist 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 (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.[4] Since the QUD reflects the discourse task at hand, relevance to task again is argued to be a central factor in anaphora resolution. Roberts (1996/2015, 2004) and Büring (2003) offer evidence that what is needed is not just a single QUD (i.e., an alternative semantics), but a more complex, pragmatically determined strategy of inquiry.

So once again, though I agree with L&S that both the understood structure of discourse and conventional features like prosodic focus play central roles in anaphora resolution, we cannot resolve anaphora as intended without taking into account the common ground, the QUD, and pragmatically inferred rhetorical relations. Though there are well-known inventories of rhetorical relations, I know of no successful attempt to determine *on purely conventional, syntactic or other structural grounds* which rhetorical relation is intended between any two given adjacent utterances.

Taken as a body, the psycholinguistic work on anaphora resolution, the work on Centering and rhetorical relations, and the role of focus in constraining felicitous resolution argues that a *central mechanism* in anaphora resolution is the restriction of the referential domain based on the assumption of relevance to the QUD and/or the joint task at hand, relevance reflected in the strategy of inquiry (rhetorical relations) for addressing that question or task.

Thus, the "rules" L&S would appeal to in order to explain anaphora resolution do not have the usual properties of the grammatical rules of a given language. Grammatical rules are generally understood to be categorical and context-free. But insofar as Centering principles are useful, they probably represent the statistical likelihood of a particular pattern of cross-sentential resolution of anaphora, useful as a fall-back when relevance to task has not already appropriately restricted the referential domain. And rhetorical relations between utterances can only be grasped against the backdrop of the QUD and/or any domain tasks in which the interlocutors are engaged.

As noted, L&S explicitly deny that the CG plays any role on their Conversational Record, and thus preclude any role for it in meaning derivation. They do not discuss the role of the QUD. But Roberts (1996/2012) takes the QUD, and associated evident goals and intentions of the interlocutors, to play a central role in interpretation: an utterance is relevant to the QUD just in case it directly or contextually addresses that question, concretely contributing toward its resolution by eliminating at least one possible answer. Every utterance, in order to be felicitous, must be relevant to the immediate QUD, a constraint that is (I have argued) conventionally encoded via an anaphoric

[4] Note that in cases like (3), 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 is required is that the CQ itself must be relevant to the prior QUD. See Roberts (1996/2012), Büring (2003) for further discussion.

presupposition triggered by the utterance's prosodic focus. This requirement of relevance to the QUD, in turn, constrains the resolution of a number of types of context-sensitivity in the conventional content of an utterance—anaphora, presupposition, implicature, topic, etc.[5] In this way, Gricean Relation emerges as a generalization about the interpretive system in which the QUD plays a central, ongoing role. Again, this is something that L&S explicitly deny: They argue that Gricean principles have no role to play in interpretation.

Gricean Maxims and the Priority of What Is Said

The QUD represents the interlocutors' immediate discourse goal, what the cooperative, *competent* interlocutor attends to *in order to grasp what the speaker means* (aims at), as well as what the competent speaker can take to be the addressee's current focus for the purposes of successful production. Thus, the QUD regularly and generally constrains interpretation by establishing expectations about what subsequent utterances will be about.[6] And such expectations, as we have seen in the brief review of the relevant literature from computational linguistics and psycholinguistics above, play a persistent, multi-faceted role in anaphora resolution. Thus, top-down expectations do play a role in anaphora resolution.

Taking the non-linguistic CG and QUD into account offers another way to understand the Gricean *priority* of what is said over what is contextually implicated. Joint attention focused on a particular task—including that of addressing the QUD—constrains the referential field in real time, to those entities relevant for the task. The same kind of pragmatic constraints can be argued to work for other anaphoric processes, including (a) the Reference Time resolution (Partee 1984) that L&S argue to be crucial in deriving from utterances like *John and Mary had a baby and got married* the classical implicature that the event in the second conjunction of temporally follows the first; and (b) the truth-conditionally crucial domain restriction of quantificational and modal operators (Roberts 1989, 1995; von Fintel 1994; Stone 1997, etc.). Importantly, all of those phenomena display analogues of "donkey sentence" presupposition resolution, in which content in the first part of an utterance may crucially contribute to the local resolution of presuppositions triggered in a later part; this argues that the contextual update in question is incremental over the course of interpretation. And (c) it seems reasonable to assume that expectation-driven top-down constraints based on task and QUD play a role in

[5] See the bibliography of related work at http://www.ling.ohio-state.edu/~croberts/QUDbib/.

[6] See Schoubye and Stokke (to appear) for related discussion about the relationship between the Gricean 'what is said' and the QUD.

disambiguation more generally, both lexical and structural. But even on this general parallel processing view, the results of bottom-up compositional processing are logically prior to the top-down influence, since presumably, as in visual processing, the contextual factors merely feed and constrain the conventionally triggered compositional interpretation. This is the sense in which the compositional process is prior, just as the percepts are prior in processing visual information. Rapid bottom-up processing of the speech signal is the content that is interpreted, but always as constrained in real-time by the incrementally updated top-down expectations related to task.

This is quite a different model of interpretation from the sausage model, according to which we expect the system to regularly come up with multiple possible interpretations for the entire utterance, with pragmatics playing a role in *disambiguation* after the completion of semantic processing. Which of these—the parallel vs. the sequential model—is better seems to be an empirical matter. But at least the experimental psycholinguistic literature seems to weigh more and more in favor of a model in which at least some pragmatic contributions act as constraints in parallel with the compositional derivation of conventional content.

The parallel processing model affords a more generous understanding of Grice, in which his maxims, and in particular Relation, can be understood to be functional constraints on the adequacy of the system, rather than premises for postcompositional inference. The assumption of relevance to the QUD thus *realizes* Relation as a constraint on interpretation. In Grice's characterization of conversational implicature, quoted above, it would suffice that "it is within the competence of [the speaker's] audience to… grasp intuitively" that the implicature is required in order to make the utterance felicitous, thus not requiring that implicatures be derived via explicit reasoning. That is, Grice's requirement that implicatures be calcul**able** does not mean that these calculations are done at run-time. Calculability is a statement about affordances, not real-time inference-making.

This third model of the relationship between conventional content and context, involving top-down Gricean constraints and input into bottom-up compositional interpretation, strikes me as more explanatory of the empirical evidence than either of the other two models.

There is another type of argument that L&S repeatedly bring to bear in support of the view that Gricean maxims play no role in interpretation. Grice claimed that implicatures are non-detachable, so that if one expression gives rise to an implicature, distinct but synonymous expressions should too. That is to say, conversational implicatures are not dependent on a particular lexical item or expression. But L&S argue that this is *not* the case with the types of examples they consider. Whatever in the utterance triggers these implications is specific to the conventional character of what is said,

and would not necessarily arise from a distinct but synonymous utterance. In this, L&S often support their argument for the grammar-based character of the relevant "rules" with claims about differences between languages. For example, they claim that indirect speech acts may arise from a particular locution in one language (*Can you pass the salt?* for 'please pass the salt') but not from its translation equivalent in another. That seems to be correct: I'm told (B. Chandrasekaran, p.c.) that a Tamil speaker would not request the salt in that way. So there's evidence that we tend to precompile the illocutionary content of certain uses of particular locutions in particular languages.

But L&S want to make this claim quite generally for the factors that play into interpretation. They claim, for example (pp. 91–92, my emphasis):

> there is a preference to resolve a subject pronoun in one sentence to the subject of the preceding sentence... It might turn out to be the only reasonable choice, given the kinds of information and goals that people typically have in conversation. It might turn out to be a side-effect of the processing mechanisms people have for attending to entities as they track contributions to conversation. Or **it might just turn out to be [a] rule of English, one that doesn't even carry over to other languages**.

Again, this is an empirical matter. I haven't seen sufficient evidence about the way discourse anaphora works across languages to argue one way or the other. But I certainly have not seen any evidence to support it. And I find it quite a stretch to imagine that the way in which the QUD constrains anaphora resolution differs from language to language. Which arguments tend to be more prominent in a given language may certainly differ from which are prominent in another, perhaps partly as a function of their differing syntactic structures, e.g. in subject-initial vs. verb-initial languages. But these are superficial matters.

Even restricting ourselves to consideration of conventional clues to resolution, Roberts (1998) argues, on the basis of a survey of all the languages for which evidence was available at that time, that the way in which prosody reflects the QUD, and hence constrains anaphora resolution, is a language universal. But this is just what we expect when the factor in question is essentially pragmatic, arising from the interaction between specifically linguistic competence and other cognitive competencies. That does not mean that the way in which such a factor influences interpretation is not conventionally triggered. But that the underlying architecture that brings such considerations regularly to bear on interpretation is at the interface between properly linguistic processing and other cognitive capacities.

Conclusion: What Is a Gricean Theory?

In conclusion, as noted above L&S grant the need for some pragmatic principles to play a role in interpretation; but they fail to specify what those principles might be. One very general principle that is extremely important for effective agency is that one should strive for a coherent mental state. What it is to be coherent may be complex, but it certainly has at least these two features:

- **Consistency:** One's beliefs ought to be consistent with the facts, hence consistent as a body, since they are the foundation of correct action. Moreover, one's goals must be consistent with one's beliefs, in the sense that one can only rationally hold a goal which one believes one can accomplish.
- **Cohesiveness:** The intentions one adopts at a given time should be functionally consistent, so that the achievement of the associated goals is mutually compatible, and they should be organized to permit focused action. With respect to the latter, some goals are subgoals of others, the subgoals optimally organized into plans to achieve the supergoal they serve; and one's goals are prioritized, so that one can determine which goal to aim for at a given time. Priorities over goals constrain action in a rational agent, since we cannot attempt to achieve all our goals at once and in any case only those goals whose achievement *at the same time* is possible. One reason for the tendency to focus on one goal at a time is because the plan to achieve a goal may be complex, with choices to make at various junctures which must be informed by the available information, conditional options which must be verified, and likelihoods to be assessed. Attention to a single goal improves ready access to relevant encyclopedic content in one's beliefs, to feed the decision process as one proceeds through a complex plan to achieve that goal.

These general pragmatic principles for rational agency pertain as well to the coordinated actions of interlocutors in discourse, including the choices a speaker makes in production and the interpretation an addressee assigns to what is said in a given context. For example, the common ground is a common reservoir of information. Insofar as it purportedly reflects the beliefs of the interlocutors, it should be consistent, as well, and there is now a very large body of experimental work to support the importance of maintenance of the common ground in interlocutor interaction. (Besides the work cited above, see the work of Herb Clark, e.g. Clark 1996.) Similarly, the interlocutors' mutually evident goals and intentions, including their immediate communicative goals in the discourse interaction as represented by the QUD, are assumed to be cohesive and organized, in order to more effectively bring about a focus on the relevant content in the common ground, especially that in immediately preceding discourse,

in order to facilitate interpretation. As Grosz (1977) pointed out and much subsequent work in psycholinguistics substantiates (see Terken and Hirschberg 1994, as well as the work of Kehler, cited above), it is not recency that is the most important factor in anaphora resolution, but coherence and relevance to task, including the QUD. Thus, relevance to the QUD is a general requirement on felicity growing out of the need to maintain cohesiveness in a discourse interaction.

This is consistent with the third view of the architecture of interpretation, outlined above, in which a functional counterpart of Gricean relevance plays a central role as a top-down constraint on interpretation, serving to realize the cohesiveness of the discourse and thereby increase processing efficiency and effectiveness. I see neither any need nor any plausible way to try to turn this general constraint on felicity into a language-specific grammatical rule.

Thus, I think the critique of Grice in L&S misconstrues the role of Gricean principles in discourse. In contrast, in the third model suggested above, instead of acting like a rule or a theorem that comes into play in the course of interpretation after the conventional content—what is said—has been calculated, a maxim like Relevance (Grice's Relation) emerges as a feature of the competence of any human language user. It is not reflected in the grammar for a particular language, but instead is an over-arching constraint on interpretation that is reflected in the interaction between different cognitive capacities (or modules, if you will), thus a constraint built into the architecture of interpretation. And in this role, it constrains the interaction between context and the resolution of context- sensitive elements in the constituent uttered.

Put another way: I have characterized Gricean relevance as a functional constraint. I would maintain that any system which constrains interpretation in this sense is a Gricean system. L&S do not preclude a parallel architecture.[6] To the extent that they would admit something like the role sketched here for relevance, then their theory is Gricean in my sense. If not, then they owe us an explanation of the empirical results, summarized above, which offer robust cross-methodological evidence for the role of relevance to the Question Under Discussion and/or task at hand in rapid, real-time anaphora resolution. I strongly suspect similar issues arise in the experimental study of the other linguistic phenomena they consider.

References

Allen, James F., Lenhart K. Schubert, George Ferguson, Peter Heeman, Chung Hee Hwang, Tsuneaki Kato, Marc Light, Nathaniel G. Martin, Bradford W. Miller, Massimo

[7] This is also true, so far as I can see, of the theories of neo-Griceans like Horn and Levinson, but not for that of Sperber and Wilson, though arguing the latter would take me beyond the purview of the current comments.

Poesio, and David R. Traum. "The TRAINS Project: A case study in building a conversational planning agent", *Journal of Experimental and Theoretical AI* 7 (1995): 7–48.
Allopenna, P. D., J. S. Magnuson, and M. K. Tanenhaus. "Tracking the Time Course of Spoken Word Recognition: Evidence for Continuous Mapping Models," *Journal of Memory and Language* 38 (1998): 419–439.
Allport, A. Visual Attention. In *Visual Attention*. Ed. M. E. Posner. Cambridge, MA: MIT Press, 1989. 631–682.
Asher, Nicholas, and Alex Lascarides. *Logics of Conversation*. Cambridge, England: Cambridge University Press, 2003.
Atanassov, Dimka, Florian Schwarz, and John C. Trueswell. *On the Processing of "Might"*. *U.Penn Working Papers in Linguistics* 19.1:1–10, 2013.
Balcetis, Emily, and Rick Dale. "Conceptual Set as a Top-Down Constraint on Visual Object Identification," *Perception* 36 (2007): 5581–5595.
Bott, L., and I. Noveck. "Some Utterances are Underinformative: The Onset and Time Course of Scalar Inferences," *Journal of Memory and Language* 51 (2004): 437–457.
Bott, Lewis, Todd M. Bailey, and Daniel Grodner. "Distinguishing Speed from Accuracy in Scalar Implicatures," *Journal of Memory and Language* 66 (2012): 123–142. http://00t0holtgrav.iweb.bsu.edu/Lab/BottBaileyGrodner12.pdf.
Breheny, Richard, Heather J. Ferguson, and Napoleon Katsos. "Investigating the Timecourse of Accessing Conversational Implicatures During Incremental Sentence Interpretation," *Language and Cognitive Processes* 28(4) (2012): 443–467.
Brown-Schmidt, Susan et al. (2005) [anaphora resolution, cited in Poesio and Rieser Brown-Schmidt, Susan and Tanenhaus, Michael K. "Real-Time Investigation of Referential Domains in Unscripted Conversation: A Targeted Language Game Approach," *Cognitive Science* 32(4) (2008):643–684.]
Büring, Daniel. "On D-Trees, Beans, and B-Accents," *Linguistics and Philosophy* 26(5) (2003): 511–545.
Chambers, Craig, Michael K. Tanenhaus, Kathleen M. Eberhard, Hana Filip, and Greg N. Carlson. "Circumscribing Referential Domains in Real-Time Sentence Comprehension," *Journal of Memory and Language* 47 (2002): 30–49.
Chambers, C. G., M. K. Tanenhaus, and J. S. Magnuson. "Action-Based Affordances and Syntactic Ambiguity Resolution," *Journal of Experimental Psychology: Learning, Memory & Cognition* 30 (2004): 687–696.
Clark, Herbert H. *Using Language*. Cambridge, England: Cambridge University Press, 1996.
Corbetta, Maurizio, and Gordon L. Shulman. "Control of Goal-Directed and Stimulus-Driven Attention in the Brain," *Nature Reviews Neuroscience* 3 (2002): 201–215. https://doi.org/10.1038/nrn755.
Fecteau, Jillian H., and Douglas Munoz. "Salience, Relevance, and Firing: A Priority Map for Target Selection," *Trends in Cognitive Sciences (Review)* 10(8) (2006): 382–390.
von Fintel, Kai. *Restrictions on Quantifier Domains*. Ph.D. Dissertation, University of Massachusetts, Amherst, 1994.
Fodor, Jerry A. *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press, 1983.
Folk, C. L., and R. Remington. "Top-Down Modulation of Preattentive Processing: Testing the Recovery Account of Contingent Capture," *Visual Cognition* 14 (2006): 445–465.
Frazier, Lyn, and Janet D. Fodor. "The Sausage Machine: A New Two-Stage Parsing Model," *Cognition* 6 (1978): 291–325.
Gibson, James J. *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin, 1979.
Gordon, P. C., B. J. Grosz, and L. A. Gillion. "Pronouns, Names and the Centering of Attention in Discourse," *Cognitive Science* 17 (1993): 311–348.
Grice, H.P. *Logic and Conversation, William James Lectures, Harvard*. Published in Donald Davidson and Gilbert Harman (eds.), 1967, *The Logic of Grammar*. Reprinted in Grice (1989) Studies in the Way of Words, Cambridge, MA: Harvard University Press.

Grodner, D. et al. ""Some", and Possibly All, Scalar Inferences are not Delayed: Evidence for Immediate Pragmatic Enrichment," *Cognition* 116 (2010): 42–55.
Grosz, Barbara J. *The Representation and Use of Focus in Dialogue Understanding*, Technical Report No. 151, Artificial Intelligence Center, SRI International, Menlo Park, CA, 1977.
Grosz, Barbara J., Aravind K. Joshi, and Scott Weinstein. Providing a unified account of definite noun phrases in discourse. In *Proceedings, 21st Annual Meeting of the Association of Computational Linguistics*, (1983) 44–50.
Grosz, B. J., A. K. Joshi, and S. Weinstein. "Centering: A Framework for Modeling the Local Coherence of Discourse," *Computational Linguistics* 21(2):202–225, 1995. (The paper originally appeared as an unpublished manuscript in 1986.).
Huang, Yi Ting, and Jesse Snedeker. "Online interpretation of scalar quantifiers: Insight into the semantics-pragmatics interface", *Cognitive Psychology* 58 (2009):376–415.
Huang, Yi Ting, and Jesse Snedeker. Logic and conversation revisited: Evidence for a division between semantic and pragmatic content in real-time language comprehension, *Language and Cognitive Processes* 26 (8) (2011): 1161–1172.
Ito, Kiwako, and Shari R. Speer. "Anticipatory Effects of Intonation: Eye Movements During Instructed Visual Speech," *Journal of Memory and Language* 58 (2008): 541–573.
Ito, Kiwako, Shari R. Speer. Semantically-Independent But Contextually- Dependent Interpretation of Contrastive Accent. In *Prosodic Categories: Production, Perception and Comprehension, Studies in Natural Language and Linguistic Theory*, 82, eds. S. Frota et al. 2011. https://doi.org/10.1007/978-94-007-0137-3_4, _ Springer ScienceþBusiness Media B.V.
Katsos, Napoleon, and Chris Cummins. *Pragmatics: From Theory to Experiment and Back Again*. Language and Linguistics Compass Teaching and Learning Guide, 3/1:1–7, 2010. Available at http://crcummins.com/publications.html
Katsos, Napoleon, Richard Breheny, and John Williams. The interaction of structural and contextual constraints during the on-line generation of scalar inferences. *Proceedings of the 27th Annual Conference of the Cognitive Science Society*. Eds. B. Bara, L. Barsalou and M. Bucciarelli. 2005: Mahwah, NJ: Erlbaum. 1108–13.
Kehler, Andrew. *Coherence, Reference, and the Theory of Grammar*. Stanford, CA: CSLI Publications, 2002.
Kehler, Andrew. *Ellipsis and Anaphora in a QUD Model of Discourse*. Talk presented at the University of Michigan Workshop in Philosophy and Linguistics, 2009.
Kehler, Andrew et al. "Coherence and Coreference Revisited," *Journal of Semantics (Special Issue on Processing Meaning)* 25(1) (2008): 1–44.
Kertz, Laura, Andrew Kehler, and Jeffrey L. Elman. *Grammatical and Coherence- Based Factors in Pronoun Interpretation*. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, Vancouver, Canada, 2006.
Lakoff, George. The Role of Deduction in Grammar. In *Studies in Linguistic Semantics*, eds. C. Fillmore and T. Langendoen. New York, NY: Holt, Rinehart and Winston, 1971.
Mack, A. "Inattentional Blindness: Looking Without Seeing," *Current Directions in Psychological Science* 12(5) (2003): 180–184.
Mack, A., and I. Rock. *Inattentional Blindness*. Cambridge, MA: MIT Press, 1998.
Marr, David *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York, NY: Freeman, 1982.
Newell, Allen. *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press, 1990.
Poesio, Massimo, and Hannes Rieser. "An Incremental Model of Anaphora and Reference Resolution Based on Resource Situations," *Dialogue and Discourse* 2(1) (2011): 235–277. https://doi.org/10.5087/dad.2011.110.
Poesio, Massimo et al. (2004) *Centering: A Parametric Theory and Its Instantiations*. NLE Technical Note TN-02- 01, CS Technical Report CSM-369. Available at: http://cswww.essex.ac.uk/staff/poesio/publications/CL04TN.pdf. Appears in abridged form in *Computational Linguistics* 30(3):309-363, 2004.

Politzer-Ahles, Stephen et al. "Distinct Neural Correlates for Pragmatic and Semantic Meaning Processing: An Event-Related Potential Investigation of Scalar Implicature Processing Using Picture-Sentence Verification," *Brain Research* 1490 (2013): 134–152. Available at: http://users.ox.ac.uk/~cpgl0080/papers.html.

Roberts, Craige. "Modal Subordination and Pronominal Anaphora in Discourse," *Linguistics and Philosophy* 12(6) (1989): 683–721. Reprinted in Javier Gutierrez-Rexach (ed.) *Semantics: Critical concepts in linguistics*, Routledge, 2003.

Roberts, Craige. Domain Selection in Dynamic Semantics. In *Quantification in Natural Languages*, eds. Emmon Bach, Eloise Jelinek, Angelika Kratzer, Barbara H. Partee. Dordrecht: Kluwer, 1995.

Roberts, Craige. "Context in Dynamic Interpretation. In Laurence Horn and Gregory Ward (eds.) *Handbook of Contemporary Pragmatic Theory*", Blackwell (2004): 197–220.

Roberts, Craige. Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. In *Ohio State University Working Papers in Linguistics* Volume 49, eds. Jae Hak Yoon and Andreas Kathol, 1996/2012. Reprinted in the 1998 version with a new Afterword in *Semantics and Pragmatics*, Volume 5, 2012.

Rohde, Hannah, and Michael C. Frank. *Markers of Discourse Structure in Child- Directed Speech*. To appear in *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, Boston, MA, 2011.

Rohde, Hannah, and Andrew Kehler. *The Bidirectional Influence Between Coherence Establishment and Pronoun Interpretation*. Poster presented at the 21st Annual CUNY Conference on Human Sentence Processing, March, 2008a.

Rohde, Hannah, and Andrew Kehler. *Coherence-Driven Expectations in Discourse and Dialog*. Poster presented at the 21st Annual CUNY Conference on Human Sentence Processing, March, 2008b.

Rohde, Hannah, Andrew Kehler, and Jeffrey L. Elman. *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, 2006.

Rohde, Hannah, Andrew Kehler, and Jeffrey L. Elman. *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, 2007.

Rooth, Mats E. "A Theory of Focus Interpretation," *Natural Language Semantics* 1 (1992): 75–116.

Schneider, Werner X., Wolfgang Einhäuser, and Gernot Horstmann. *Attentional Selection in Visual Perception, Memory and Action: A Quest for Cross-Domain Integration*. Introduction to a special issue on Attentional selection in visual perception, memory and action, in *Philosophical Transactions* B 368(1628), 2013. https://doi.org/10.1098/rstb.2013.0053. http://rstb.royalsocietypublishing.org/content/368/1628/20130053

Sedivy, Julie. "Pragmatic versus Form-based Accounts of Referential Contrast: Evidence for Effects of Informativity Expectations", *Journal of Psycholinguistic Research* 32 (2003): 3–23.

Schoubye, Anders J., and Andreas, Stokke. (to appear) What is said? Forthcoming in *Noûs*.

Shimojo, Shinsuke, Michael Paradoso, and Ichiro Fujita. "What Visual Perception Tells Us About Mind and Brain," *Proceedings of the National Academy of Sciences of the United States of America* 98(22), 2001.

Simons, Mandy, David, Baver, Craige, Roberts, and Judith Tonhauser. "The Best Question: Explaining the Projection Behavior of Factives", *Discourse Processes* 54 (2017):187–206. Available open-access at: http://www.tandfonline.com/doi/pdf/10.1080/0163853X.2016.1150660?needAccess=true

Simons, Daniel J., and Christopher F. Chabris. "Gorillas in Our Midst: Sustained Inattentional Blindness for Dynamic Events," *Perception* 28 (1999): 1059–1074.

Smith, Edward E., and Stephen M. Kosslyn. *Cognitive Psychology: Mind and brain*. See Chapter 2: Perception. Upper Saddle River, NJ: Pearson, 2006.

Stone, Matthew. *The Anaphoric Parallel Between Modality and Tense*. IRCS TR 97-06. Technical Report: University of Pennsylvania, 1997.

Tanenhaus, Michael K., Craig G. Chambers, and Joy E. Hanna. Referential Domains in Spoken Language Comprehension: Using Eye Movements to Bridge the Product and

Action Traditions. In *The Interface of Language, Vision, and Action: Eye Movements and the Visual World*. Eds. J. M. Henderson, F. Ferreira. New York, NY: Psychology Press, 2004. 279–318.
Terken, Jacques, and Julia Hirschberg. "Deaccentuation of Words Representing "Given" Information: Effects of Persistence of Grammatical Function and Surface Position," *Language and Speech* 37(2) (1994): 125–145.
Tetreault, Joel R. Decomposing Discourse. In *Anaphora Processing*, eds. Antonio Branco, Tony McEnery, Ruslan Mitkov. Amsterdam, The Netherlands: John Benjamins, 2005.
Tetreault, Joel R., and James F. Allen. *Dialogue Structure and Pronoun Resolution*. In *Proceedings of the Discourse Anaphora and Anaphora Resolution Colloquium (CDAARC'04)*, Azores, September 23–24, 2004.
Tian, Y., R. Breheny, and H. J. Ferguson. "We Simulate Negated Information: A Dynamic Pragmatic Account," *Quarterly Journal of Experimental Psychology* 63(12) (2010): 2305–2312.
Walker, Marilyn A., Aravind K. Joshi, and Ellen F. Prince. Introduction. In *Centering Theory in Discourse*, eds. M. Walker, A. Joshi and E. Prince. New York, NY: Oxford University Press (359–400), 1998.