

Craige Roberts
July 6, 2023; revised July 10, 2023

Anaphora in Formal Pragmatics
LSA Summer Institute at UMass, Summer 2023

Lectures II and III: Dynamic pragmatics

1. Pragmatics in classical formal semantics: The two-phase account
2. Dynamic semantics
 - Problems for dynamic semantics
 - 2.1 Semantically intrusive pragmatic phenomena
 - 2.2 CCPs and illocutionary force
 - 2.3 What's dynamic: context update, not semantic content
3. Pragmatic competence as part of our cognitive architecture
 - 3.1 Experimental evidence
 - 3.2 Parallel bottom-up and top-down processing
 - 3.3 A comparable top-down feature of linguistic competence
 - 3.4 Gricean maxims and the priority of what is said
4. Dynamic pragmatics
 - 4.1 Context of interpretation
 - 4.2 Dynamically updated context feeds contextualized LF
 - 4.3 Lexical **CHARACTER** and how it feeds interpretation
5. Deriving intrusive pragmatic phenomena
 - 5.1 Deriving donkey anaphora
 - 5.2 Dynamic update with auxiliary content
6. Conclusions
 - 6.1 Logical form and Discourse Representations
 - 6.2 Summary and prospects

1. Pragmatics in classical formal semantics: The two-phase account

In classical compositional semantics, pragmatics comes to bear on semantic interpretation at two different points:

Phase 1: Values for indexicals, free variables are available in the context of utterance and feed compositional interpretation.

Phase 2: Post-compositional calculation of conversational implicatures, other enrichments of what is said.

With respect to Phase 2 phenomena:

Grice requires 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 his conversational maxims—Quality, Relation, Quantity and Manner. He consistently insists that this calculation must take place on the basis of ‘what is said’.

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.

Figure 1 (see the figures in the separate file accompanying this handout) is a schematic representation of the components of semantics and pragmatics and their interaction in the classical two-phase view of pragmatic effects on interpretation. This general approach is that developed in Montague (1973), Kaplan (1977), and most recently adopted in Heim & Kratzer (1998).

On this approach:

- The conventional input to compositional semantics is given by the syntactic logical form (LF) of an utterance, along with the semantic values of its lexical components, as indicated by **red lines** in Figure 1.
- What is interpreted is a *contextualized logical form* (LF) with its lexical entries, annotated with contextual information for resolution of indexicals, variables.
- Context is a combination of Kaplan’s tuple (speaker, addressee, time, etc.)—given once per LF—and a Tarskian assignment of values for free variables, the assignment function varied in the usual way solely in the interpretation of quantification.
- The role of pragmatics is two-fold, indicated with **green lines** in the figure:
 - The context of utterance acts like a sky hook to magically pull values for variables from out of the ether.
 - Gricean maxims play a post-compositional role: generating implicatures on the basis of the output of the compositional rules.

To the extent possible, the classicist avoids the wastebasket of post-compositional pragmatic explanation.

- In Heim & Kratzer (1998) (following Heim 1983): Presuppositions are made explicit. These bear indirectly on the interpretation ultimately derived (the proffered content), via constraints on felicity in context.

But over the past forty years, a growing body of literature in linguistic semantics and philosophy of language has provided evidence that the classical two-phase view, with its static context given once and for all, is empirically inadequate. The most famous such evidence, of course, are the so-called donkey sentences of Geach (1962), like (1):

(1) If [_{s1} a farmer₁ owns a donkey₂ s₁], [_{s2} he₁ beats it₂ s₂].

Such examples argue that the interaction between semantic content and context is dynamic, the context changing over a single utterance in such a way that different sub-constituents are

interpreted with respect to different contexts. Such interpretations cannot be readily captured in the static semantics of the classical two-phase view of pragmatics.

Note on the e-type alternative: Heim & Kratzer (1998) suggest an e-type approach to donkey sentences, of the sort later elaborated by Elbourne (2005). But note that Elbourne himself concedes that the account he offers doesn't suffice to distinguish definite NPs from indefinites unless one assumes that the definites are anaphoric, proposing that one take a DRT approach to nominal anaphora. So the e-type approach to donkey anaphora simply pushes the problem of dynamic context update down the road. See Roberts (2004) for a detailed critical analysis of earlier e-type analyses.

2. Dynamic semantics

Starting with the work of Kamp (1981) and Heim (1982), many authors have attempted to capture this more sophisticated conception of context with a dynamic semantics (Groenendijk & Stokhof 1990, Muskens 1996, Anderbois et al. (2015), Martin 2016, Dotlačil & Roelofsen 2019, etc.). Figure 2 is a schematic representation of the semantics-pragmatics relationship in this type of semantics.

There are three essential features of dynamic semantics that differ from the classical picture.

- First, the notion of context modeled in this framework is richer than in the classical account: We now track not only the Common Ground (CG) and utterance-internal assignment function shifts, but also crucially use referential indices and assignment functions to keep track of what has been referred to across utterances. When an entity is referred to with an NP, the NP's referential index is entered into the conversational record as a discourse referent, an entity "familiar" to the interlocutors. Further, that information about the relevant entity is encoded as constraints on possible values for variables bearing that referential index. In subsequent discourse, the only assignment functions that are admissible are those such that the values they assign to that index have the properties previously predicated of it. Then if the same familiar discourse referent is encountered in subsequent discourse—identified via the same referential index—its value is already constrained to reflect the information in the CG about that entity.
- Second, in the interpretation of an utterance in dynamic semantics, the context of interpretation for sub-constituents is no longer given once and for all as the context of utterance. Rather, just as Tarskian assignment functions can be shifted utterance-internally, so now the entire context is updated dynamically, with the meanings of some sub-constituents within an utterance updating the context for interpretation of other constituents.

Thus, dynamic semantics significantly expands the role of the assignment functions we found in classical semantics; they are no longer arbitrarily chosen, but are used to track information used for the resolution of non-bound anaphora, and of presuppositions more generally, including donkey sentences like (1) and comparable temporal phenomena (2), which are challenging for the classical view of semantics/pragmatics because, though unbound, their utterance-internal resolution crucially bears on truth conditions.

(2) Whenever Mary wrote a letter₃, Sam answered it₃ two days later. [Partee 1984]

A given context makes available a restricted set of assignment functions, those such that the values they would assign to variables previously used in the discourse are in keeping with what has already been asserted about those values. Technically, to do this we need world/assignment pairs, since more than one assignment may verify the content of an utterance in a given world. As usual, the semantics uses the assignment functions and other contextual information in the CG to track information for the resolution of non-bound anaphora, and presuppositions generally, including the anaphoric presuppositions in donkey sentences (1) and comparable temporal phenomena (2). In this framework, any presuppositions triggered by a lexical item, including anaphoric presuppositions, constrain the felicity of the local context of interpretation for the trigger, which must satisfy those presuppositions.

- The third way in which dynamic semantics crucially differs from the static semantics is that to implement the utterance-internal dynamic context just described, the semantic content of an expression is itself taken to be dynamic: Instead of the classical static denotation—a property, proposition, or entity, etc.—a constituent denotes a Context Change Potential (CCP), a function taking the local context of interpretation to yield an updated context, as you see in Figure 2. Hence, e.g., the first conjunct of a conjunction or the *if*-clause in a conditional can contribute to the context of interpretation for the 2nd conjunct or the consequent—the context updated (at least provisionally, in the case of the *if*-clause) to reflect the information in the first conjunct or *if*-clause.

In Figure 2, as in other figures we'll consider, there are many ways of realizing the relations indicated by the arrows—see Heim's (1982) File Change Semantics, Kamp's DRS update functions (Kamp 1981, Kamp & Reyle 1993), Muskens' (1996) compositional dynamic semantics, Martin's (2016) compositional categorial dynamic semantics, and recent work on dynamic inquisitive semantics (Dotlačil & Roelofsen 2019, 2020; Roelofsen & Dotlačil 2022)—but, modulo the types required for dynamic inquisitive semantics, all assume something like CCPs and the same type of two-way flow of information between context and semantic content, wherein compositional semantics feeds the context of interpretation, which in turn feeds the compositional interpretation, etc.

Problems for dynamic semantics:

2.1 Semantically intrusive pragmatic phenomena

Dynamic semantics marked a vital step forward in understanding the relationship between semantics and pragmatics. It works well for many of the types of phenomena that were covered in classical static semantics by what we called Phase one pragmatics: indexicals (supposedly once and for all per utterance, but as we will see, these also can be shifted in local contexts under ROOFING), free anaphora (now extended to cases with ROOFED antecedents), other merely local

explicit satisfaction of presuppositions. I.e. anaphoric antecedents that are available *merely* locally, not globally.

But there are important respects in which this framework still fails to correctly capture the ways that context influences truth conditional content. The central problem is with the types of phenomena addressed in the classical pragmatics in Phase two, like conversational implicatures. These, too, have intrusive counterparts: cases where they arguably cannot be merely calculated *after* the complete compositional semantic interpretation has been derived.

Semantically intrusive pragmatic phenomena: Those in which dynamic context update cannot be argued to follow from the explicit content of sub-constituents of the utterance.

Heim (1982) knew that discourse anaphora wasn't always and only licensed by explicit antecedents. For example, for deictic uses of free pronouns, she appealed to accommodation. But she took examples like Partee's marbles sentences to argue for a supposed "formal link" requirement on felicitous anaphora (Heim 1990). As discussed in Lecture 1, that requirement is too strong in general, even in marbles sentences.

Cases involving semantically intrusive pragmatic phenomena:

Donkey bridging:

(3) If a farmer₁ owns a John Deere₂, he₁ uses the tractor₂ to plow.

The architecture of information flow in dynamic semantics makes no allowance for local context update except through the CCPs of explicit sub-constituents. Bridging requires inferences based on world knowledge and lexical associations.

Modal subordination: Pragmatically driven domain restriction:

Modal subordination involves tacit restriction of the scope of a modal operator (modal auxiliary, adverb of quantification, modal operator in the lexical semantics of an attitude verb, etc.) with irrealis content from a prior utterance. So (4) out of the blue is a simple prediction—the birds are gonna get hungry—but the same sentence in (5) has a conditional sense: only in case Edna forgets to fill the feeder.

(4) The birds will get hungry (this winter).

(5) If Edna forgets to fill the birdfeeder, she will feel very bad.
The birds will get hungry.

(Roberts 1989)

Interaction with pronominal anaphora:

(6) A wolf might break in. It would eat you first. (Fred Landman, p.c. to Roberts) [lecture 1]

(7) Harvey courts a girl at every convention.
She always comes to the banquet with him.

The girl is usually very pretty.

(Karttunen 1976)

Interaction with *too*:

- (8) If Mary comes, we'll have a quorum. If Susan comes too, we'll have a majority. (Heim 1992)

Interaction with sluicing:

- (9) Stefanie often eats out when she's travelling on business. If she doesn't know the city well, where is usually influenced by the latest Zagat ratings.

Mixed force (6), (7), (9) shows that modal subordination cannot be accounted for by somehow extending the scope of the operator that ROOFS the antecedent to take scope over the intrusive domain restriction. Nor can it be characterized as involving descriptive content somehow "borrowed" from prior utterances: This would fail to explain constraints on modal subordination across mixed attitudes, as illustrated in the difference between (10) and (11) (Roberts 1996)

- (10) Alice **fears** there's a squirrel in her kitchen cabinets.
She **hopes** to trap it alive and turn it loose outside.
- (11) Alice **denies** that there's a squirrel in her kitchen cabinets.
#She **hopes** to trap it alive and turn it loose outside.

With respect to (11), the accessibility relation for the second predicate, *hope*, is given by the beliefs of the subject, Alice. But, as noted in McCawley's (1981:337) discussion of these verbs, the first sentence entails that Alice does *not* believe in the existence of a squirrel in her kitchen cabinets. So instead of an entailment relation to license the anaphoric relation to the pronouns in the complement of *hope*, via weak familiarity, there is one which precludes it. This contrasts with (10), where one's fears entail something like near-belief, which appears to be sufficient grounds on which to base one's hopes. So clearly the weak familiarity of the squirrel in (10), plus the lexical semantics of *denies*, *fears* and *hopes* all play a role in licensing the anaphora in (10), precluding it in (11). The domain restriction in (10) is pragmatically inferred, intrusive.

Intrusive conversational implicatures:

Intrusive implicatures (Cohen 1971; Kadmon 2001; Chierchia 2004; Chierchia, Fox & Spector 2011; Simons 2010, 2011, 2013) also play a regular, presumably explicable role in truth conditional interpretation.

- (12) If the old king has died of a heart attack and a republic has been declared, then Tom will be quite content. [Cohen 1971]
- (13) Bill believes that some of his students are waiting for him. [Chierchia 2004]
standard accounts predict: 'it is not the case that Bill believes that all of his students are waiting for him' [true if Bill has no beliefs about the students]
attested implicature: 'Bill believes that not all of his students are waiting for him'

About (12) (her (4)) and (13) (her (5)), Simons (2010) says:

As argued by Cohen 1971, the implication of temporal ordering and causal connection. . . is part of the content of the antecedent of [(12)]. Chierchia 2004 argues that an utterance of [(13)] attributes to Bill the belief that some and not more than some of his students are waiting for him. . .

...

[With respect to (13):] As Chierchia points out, standard accounts predict that if this sentence gives rise to a scalar implicature, it is this:

It is not the case that Bill believes that all of his students are waiting for him.

This, though, could be true if in fact Bill had no beliefs at all as to whether all of his students were waiting for him. The observed implicature, that Bill believes that not all of his students are waiting for him, cannot be generated directly from the asserted content of [(13)], but only from the clause embedded under *believe*. But, as the content of this clause does not count as said, then it should not trigger any implicature calculation. This, then, is the calculation problem.

[The compositionality problem, then, is that] in examples like [(12)] and [(13)], the implicatures apparently generated by embedded clauses seem to fall under the scope of the embedding operators, and thus to contribute to the truth conditional content expressed: that is, to what is said.

Some authors, including Roberts (1989) and Kadmon (2001), treat local contexts in DRT more like a Stalnakerian Common Ground than purely the result of CCP update, introducing the results of accommodation and intrusive implicature. But from the point of view of the architecture of information flow captured by the theory itself, this is an ad hoc elaboration.

Intrusive scalar implicatures in conditionals:

(14) Kai ate the broccoli or he ate some of the peas. (Sauerland 2004)

(15) If I give an extension to some of my students, the others will be upset. (Simons 2011)

Local enrichment implicatures (Horn's R-implicatures; Levinson's I-implicatures):

(16) A: How will you get to SALT?

B: Either I'll rent a car or I'll fly.

'Either I'll rent a car *and drive that car to SALT* or I'll fly' (Simons 2011)

(17) If you need to get to SALT, you could rent a car or you could fly.

Local issue: how to get to SALT

Note that scalar implicatures are very sensitive to the question under discussion:

(18) [Context: A is a bartender. B sits down at the bar.]

A: Are you 21?

B: Yes.

A: How old are you?

B: 28

(Thomason, Hobbs & Moore 1996)

And see Welker's (1994) observation about the role of task-orientation in determining when implicatures arise:

- (19) A: I'm having a dinner party and I need exactly six more chairs.
B: John has four chairs.
- (20) A: I'm having a dinner party and I need exactly four more chairs.
B: John has four chairs. (Welker 1994)

The scalar implicature only predictably turns up in (19), not in (20). That is, it depends on the interlocutors' understood goal at hand. If A needs six chairs (19) and B knows that John has more than four chairs, he is uncooperative if he doesn't say so, implicating that so far as B knows John has no more than four. In (20), where all that matters is whether John has at least four chairs, the exhaustivity implication fails to arise.

The same factors turn out to constrain whether intrusive implicatures arise in contexts like that of Chierchia's (13):

- (21) A: I need six chairs.
B: Bill thinks that John has four chairs.
 'John has no more than four chairs'
- (22) A: I need four more chairs.
B: Bill thinks that John has four chairs.
 'John has at least four chairs'
 i.e. doesn't implicate that John has no more than four chairs.

And in the related conditionals:

- (23) If you need six chairs, I think John has four chairs.
(24) If you need four chairs, I think John has four chairs.

An EXH operator in LF won't predict *when* you get the scalar implicature. It only predicts that it's possible.

Semantically intrusive auxiliary content

There is another empirical challenge both for the classical view of the semantics/pragmatics interface and for dynamic semantics. This is the existence of another possible aspect of the semantic contribution of a constituent, besides its ordinary semantic content and any presuppositions it triggers. This is that aspect of meaning that contributes what Potts (2005) called *supplemental conventional implicatures*, contributed in English by nominal appositives, non-restrictive relative clauses (NRRs) and other constructions. Potts considered examples like (25), where the appositive *a very honest woman*, modifying *his wife*, is understood to reflect a belief of the speaker about Sam's wife, and *not* something that Sam himself believes, despite the fact that it occurs in the syntactic scope of *believe* with Sam as agent. So following (25a) with (25b) seems to be a contradiction, whereas (25c) is fine:

- Not locally entailed:

- (25) a) Sam believed that his wife, a very honest woman, was lying to him.
 b) #Sam's wife isn't an honest woman.
 c) ✓ Sam's wife wasn't lying to him.

Moreover, auxiliary content cannot bind or be bound by the semantic content of the clause in which it occurs, as we see in (26):

- (26) *No reporter₁ believes that Ames, often the subject of his₁ columns, is a spy. [Potts 2005]

On the basis of phenomena like these, Potts argued convincingly that auxiliary content does not directly contribute to the proffered semantic content of an utterance in which it occurs, instead acting as a sort of aside or parenthetical. He provides an account in a static semantics, where auxiliary content is always directly added to the globally available information about the speaker's "world view", except when shifted in direct quotation to the reported agent of the utterance quoted.

But as Amaral et al. (2007) pointed out, there are examples where auxiliary content has truth conditional effects which cannot be explained if we assume that context is static, the same across an entire utterance except as manipulated by quantifiers in the usual Tarskian manner. (27) is an example of donkey anaphora with auxiliary content; the antecedent of *it* in the main clause is *a bachelor's degree* occurring in a NRR which in turn is under the scope of *every professional man*:

- (27) [Every professional man I polled]₁ said that while [his₁ wife]₂, who had earned a bachelor's degree₃, had no work experience, he₁ thought she₂ could use it₃ to get a good job if she₂ needed one. [ARS 2007]

Thus, while auxiliary content does not interact compositionally with the ordinary semantic content of the utterance in which it occurs—as evidenced by the lack of local entailment and binding—it does come to have an essential effect on truth conditional interpretation. So if we want to capture the donkey anaphora in (27), we have to address this issue:

Auxiliary content updates context dynamically, intrusively, yet without contributing to the compositional semantics that yields proffered content.

All the intrusive phenomena considered above argue that a purely semantic account will not suffice to explain the phenomena that drive dynamic accounts of interpretation.

CCPs cannot capture semantically intrusive pragmatic phenomena.

2.2 CCPs and illocutionary force

Stalnaker (2014) argues that the assumption that the update of an indicative utterance (its CCP) is automatically a contribution to the Common Ground/Context Set of the input context is problematic from the point of view of Speech Act theory¹. Consider (28):

- (28) A and B are strategizing about how their gang is going to rob a bank:
- i. A: Suppose the police arrive while we're cleaning out the vault.
 - ii. B: We'll elude them by escaping over the roof.
 - iii. A: What if our short-circuiting software fails and the alarm goes off?
 - iv. B: Grab the cash in the drawers and run!
 - v. A: Suppose the guard gets untied.
 - vi. Should I shoot him?

In (28) the interlocutors are negotiating a plan. Besides directly planning how to achieve their goals, they consider various possible obstacles and contingencies and speculate about alternative ways of proceeding if/when they arise. So there are branching possibilities. In each branch, they consider 'what to do' *as if* acting out the plan. Earlier B might have said *then tie up the guard*, etc., just *as-if* they were in an actual situation in which that action was appropriate, modulo the anaphoric *then*. In this extended irrealis context, all proffered content is relativized to the hypothetical scenario being entertained, and as in modal subordination generally, the relativization has implications for the resolution of anaphora and other presuppositions. In (28i) A proposes consideration of one possible contingency; since this is a planning discourse, this raises the practical question of what to do in that circumstance. In (ii), B offers a plan to address that contingency; notably, the Reference Time for (ii) is clearly the immediate aftermath of the event described in (i), and *them* is resolved to the police. In (iii) A directly poses a question about how to plan for yet another possible contingency, and in (iv) B suggests what to do in that hypothetical circumstance, with the same relativization of Reference Time, and also restriction of the referential domain to entities in the relevant hypothetical situation, so that *the cash* and *the drawers* are understood to be those in the bank. In (v), A proposes consideration of yet another contingency, and then in (vi) asks whether she should adopt a particular provisional plan in that case, shooting the arbitrary guard under discussion, thereby resolving *him* and, once again, the Reference Time—she's asking whether she should shoot the guard (uniquely familiar in that situation) *at that point in time*. Note that the imperative in (iv) is a conditional suggestion: 'if the alarm goes off, grab the cash and run'. The modal *should* in (vi) takes the background scenario in the bank plus (v) as part of its modal base, and as its ordering source the understood goals and priorities both immediate (get the money) and longer-term (get away, avoid the worst potential legal consequences, stay alive), yielding a conditional interpretation of the question: 'should I shoot the guard if he gets untied'. And just so, the imperative in (iv) takes (iii) as one of the premises for its modal base, and the same general priorities for the ordering source. In all these cases, relevance to the question under discussion (as part of the practical strategy being developed) and the interlocutors' understood joint and individual goals in G play a direct role in restricting the applicable circumstances to yield the natural interpretation.

¹ See Portner 2018, Chapter 3 for an excellent overview of the literature; Roberts 2018 for arguments that illocutionary force is not part of the proffered content of an utterance, but is instead pragmatically inferred.

As (28) illustrates, taking an indicative as an assertion, an interrogative as a question, an imperative as a directive is arguably a pragmatic matter, not automatically part of compositionally determined meaning. So to build that into conventional content seems off the mark:

Grammatical mood does not determine speech act type.

2.3 What's dynamic: context update, not semantic content

Moreover, as Stalnaker (2014) pointed out, there is a sense in which dynamic semantics fails to clearly represent what is meant_{nm} by an utterance or its sub-constituents. An unembedded indicative clause not only serves (in the default case) to update the Common Ground, but also intuitively denotes something like a proposition, as we see when such a clause serves as complement to an attitude predicate, or the antecedent of a propositional anaphor like *that*. Yet in dynamic semantics its meaning is a function from contexts to updated contexts, the intuitive propositional content only retrievable indirectly via its effect on the CG.

And while it is useful, e.g. for donkey anaphora, to capture the fact that the use of a quantificational NP leads to the introduction of a new discourse referent and the update of the local context with the information in its nominal complement, the latter then serving as a new local context for its second argument, we might still want to insist that such a constituent *denotes* a generalized quantifier, with all the properties universally observed in that type captured as relations over two static argument types. I think understanding them in this way is important for a deeper explanation of donkey anaphora in both the quantificational and conditional cases. Chierchia (1995) argues that it is the conservativity (Barwise & Cooper 1981) of generalized quantifiers that explains why we would expect the content of the restriction of the operator to be available in the interpretation of its nuclear scope: $Q(R)(S) = Q(R)(R \cap S)$. If we take the adverb of quantification or modal underlying the semantics of a conditional to be a different type of generalized quantifier (deSwart 1991, Frank 1996), then the conditional donkey sentences are predicted, as well.

This is a conceptual problem for dynamic semantics, not empirical. But I think it underlies the empirical problem illustrated by the intrusive phenomena, including modal subordination, anaphoric bridging, and perspective shift under attitudes, as well as intrusive conversational implicature:

Though context update is *driven* by explicit semantic content, it seems that semantic content does not *determine* context update:

It is the way that context is updated that is dynamic, not semantic content *per se*.

And hence, in keeping with the pragmatic character of speech act determination:

Context update is essentially pragmatic, not semantic.

3. Pragmatic competence as part of our cognitive architecture

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's no evidence of time-consuming post-compositional interpretive inferences, and there's a great deal of evidence that such inferences are processed in the real-time course of compositional interpretation.

3.1 Experimental evidence for dynamic pragmatics

- 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 & Snedeker 2009,2011). The calculation of an SI may be cognitively costly, slowing down processing by as much as 400 msec (Bott & Noveck 2004, 2012; Katsos et al. 2005; Huang & Snedeker 2009,2011), though this is controversial (Sedivy 2003, Grodner et al. 2010, Breheny et al. 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 100ms to conduct a cognitive action like drawing an inference, and with the evidence (e.g., see Allopenna et al. 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 et al. (2013) on *NOT-must* implicatures associated with use of *might* found the same delay in processing as that found by Huang & Snedeker (2009). Other studies argue that SIs are neither automatic nor default, but only take place in response to appropriate contextual factors (Katsos & Cummins 2010), like the QUD (Tian et al. 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).

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 post-compositionally.

For recent relevant work on scalar and other conversational implicatures, see the QUD bibliography available on my website: <https://www.asc.ohio-state.edu/roberts.21/QUDbib/>. In particular, consider Degen & Tanenhaus (2019), Ronai & Xiang (2020), Skordos & Barner (2019), Smeets & Meroni (2020), Yang & Fiorentino (2018)

- Work on processing **auxiliary content** (Dillon, Clifton & Frazier 2014, Dillon et al. 2017) argues that “parentheticals” (non-restrictive relative clauses and appositive nominals) “are processed independently of their embedding utterance”, treated “differently from at-issue content in online processing”, e.g. contributing less processing difficulty than restrictive relative clauses when they intervene between a filler and a gap. I take this as further evidence that there are parallel independent components at work in processing.
- Ito & Speer (2008) study **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 & 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 & Speer 2011:86).
- Psycholinguistic work on **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’s 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 & Tanenhaus (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 et al. (2004), and Tanenhaus et al. (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 & 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 syntactico-semantic 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).

How can we make linguistic sense of such evidence of rapidly processed intrusive pragmatic phenomena?

3.2 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 et al. 2001, and the overview and introduction in Smith & 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, based on contextually available information, goals, and conceptual sets (e.g., Allport 1989, Corbetta & Shulman 2002, Balceris & Dale 2007). Fecteau & Munoz (2006) present evidence that the two kinds of processes yield distinct neural signals (p.387). See especially the work on *inattentional blindness* (Simons & Chabris 1999, Simons 2000),² 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 & 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

² See the video at: <https://www.youtube.com/watch?v=vJG698U2Mvo>

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 & Rock 1998, Mack 2003). But according to the contingent capture model of pre-attentive processing, a person's "current intentions and/or goals affect the speed and efficiency of pre-attentive 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 & 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, "saliency and relevance yield distinct neural signals – saliency is reflected in the initial registration of the target, and relevance is reflected in the elevated activity following the predictive cue." (Fecteau & Munoz 2006:387)

Given this common feature of human cognitive architecture (see recent work on attentional selection across cognitive domains in Schneider et al. 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. The psycholinguistic work reviewed in section 3.1 on scalar implicature, anaphora resolution, and the role of focus in constraining felicitous resolution can all be taken to argue that something like this is going on.

3.3 A comparable top-down feature of linguistic competence

In vision, the top-down selective attention and expectations are based on contextually available information, goals, and conceptual sets. Do we find something similar bearing on linguistic interpretation?

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. 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.³

³ See Schoubye & Stokke (2015) for related discussion about the relationship between the Gricean 'what is said' and the QUD.

Every utterance, in order to be felicitous, must be relevant to the immediate QUD, a constraining factor that is (I have argued) conventionally encoded via an anaphoric 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—presupposition, implicature, topic, etc. In this way, Gricean Relation emerges as a generalization about the interpretive system in which the QUD plays a central, on-going role.⁴ See the bibliography of related work at <http://www.ling.ohio-state.edu/~croberts/QUDbib/>.

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, as in other aspects of interpretation. Clifton & Frazier (2017) is a nice overview of a body of experimental work that together argues that the QUD plays an on-going role in processing, biasing and facilitating comprehension when the utterance is congruent with the QUD, a role that has effects in real time processing.

Thus, we hypothesize 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.⁵

3.4 Gricean maxims and the priority of what is said

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 that Lepore & Stone (2015) 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 Stechow 1994, Stone 1997, etc.). Importantly, all of those phenomena display analogues of “donkey sentence” presupposition resolution, in which under the scope of a ROOFING operator, 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 disambiguation more generally, both lexical and structural: only those lexical senses and syntactic parses that are RELEVANT to the QUD are entertained. But even on this general parallel processing view, the results of bottom-up compositional processing are logically prior to the top-

⁴ Again, see the QUD bibliography for extensive references: <https://www.asc.ohio-state.edu/roberts.21/QUDbib/>

⁵ Note that I am *not* claiming the QUD is the answer to all puzzles in pragmatics. Rather, it is striking that it is a pervasive feature of the contextual factors that bear on interpretation. See, again, the section of the QUD bibliography on psycholinguistic experimental literature.

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.

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 post-compositional 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 **calculable** doesn't mean that these calculations are done at run-time. Calculability is a statement about affordances, not real-time inference-making.

This ⁶model 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 classical two-phase model or the model proposed by Lepore & Stone (2015), wherein discourse processes are viewed as part of the grammar.

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 doesn't mean that the way in which such a factor influences interpretation isn't 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.

4. Dynamic pragmatics

Recall:

Utterance: <constituent, context>

where the constituent is a syntactically derived constituent under an analysis (e.g., LF).

This precludes an utterance being multi-sentential, since the maximal unit *syntactically* derived is a matrix sentence.

⁶ See Roberts (2017) for an extended critique of Lepore & Stone (2015).

For the arbitrary felicitous utterance, a theory of pragmatics should give an account of the gap between semantic content and attested meaning in context. I.e., it should predict its attested meaning_{nn}.

To do this, a pragmatic theory requires:

- 1) A compositional theory of syntax-semantics:
 - 1a) a theory of the meanings of the parts of the constituent uttered, thus including an account of **the semantic CHARACTER of lexical items**.
 - 1b) a rule-by-rule compositional semantics that composes the CHARACTERS of the parts in the way they're composed syntactically.
- 2) A theory of context: What is a context? What kinds of information does it contain and how (if at all) are they related? How does this information bear on interpretation?
Given that context changes across the derivation, we need to distinguish between and relate:
 - the **global context of utterance**: context available prior to utterance, plus the fact of its utterance (speaker, etc.)
 - the **local context of interpretation** for a given constituent: what feeds its local semantic interpretation
- 3) An account of the ways that context gets updated.
 - 3a) We have seen that the interaction with semantic content will have to be dynamic: updating the local context for some sub-constituents on the basis of the semantic content of others
 - 3b) Auxiliary content can come to bear anaphorically on co-ROOFED semantic content.
 - 3c) There is also evidence that other, non-linguistically-given information that becomes available non-globally may also be reflected in the local context of interpretation: locally relevant world knowledge and abductive inference for modal subordination, donkey bridging, intrusive implicature, etc.

To address the range of phenomena alluded to above, and others besides, I will offer an alternative picture of the interaction between semantics and pragmatics. Figure 3 sketches a schema for interpretation which crucially features a static semantics and a central role for a dynamic pragmatics (see Dever 2013, Stalnaker 2014, K. Lewis 2014, 2017, Portner 2018 for early discussions), the latter conceived of as a crucial non-grammatical component of the linguistic competence underlying interpretation. I will argue that the result has the virtues of dynamic semantics but without the drawbacks just discussed.

To get a feeling for the basic differences between these different models of interpretation, the black lines in Figures 1b, 2b and 3b trace the flow of contextual information in the static semantics, the dynamic semantics, and the dynamic pragmatics. In the static semantics of Heim & Kratzer (Figure 1b), though the notion of context may be richer than that in classical Montague Grammar by virtue of assuming a CG for non-anaphoric presupposition satisfaction, there is no context update. Pragmatic input is two-phase, as in Kaplan/Montague, with Gricean maxims and the like in the second phase yielding implicatures which enrich the compositional meaning to yield what is meant_{nn}. In dynamic semantics (Figure 2b), we have continuous update of the context of interpretation, some of this temporary, e.g. under the scope of an operator. But

the only formal input to this update is given by the compositional semantic content of the utterance.

But on the view of dynamic pragmatics represented in Figure 3b, context revision is fed in multiple ways. As in dynamic semantics, the output of (partial) static compositional semantic interpretation feeds context revision over sub-constituents of the utterance: The loop on the right from the compositional semantics through the dynamic pragmatic update rules feeds the local context that is brought to bear on the interpretation of sub-constituents of a contextualized logical form. But independently, auxiliary content (top left) directly feeds context update in the dynamic pragmatics, while lexically triggered presuppositions constrain local context in the usual way. Neither the auxiliary content nor presuppositions interacts with operators at Logical Form. In addition, pragmatic constraints on well-formed context (bottom right), such as Relevance to the QUD and the Right Frontier constraint on anaphora resolution (Polányi 1985, Asher & Lascarides 2003, Roberts 2016), continuously constrain context update to play a role on-line in disambiguation, anaphora resolution, implicature generation, and speech act recognition. The dynamic pragmatics so-fed determines a local context of interpretation for a given constituent, which feeds the contextualized logical form, in turn feeding compositional semantics. Thus, on this model there are multiple indirect contributions to truth conditional semantics, highlighted in green in Figure 3b; these contributions are not reflected in the syntactically determined LF, but only in the contextualized LF that is fed by the dynamic pragmatics.

The formal framework

All the elements of the architecture in Figure 3 are assumed to be developments of a general cognitive strategy for processing perceptual input, but specialized to contribute to linguistic interpretation. Such processes and their interactions are part of our linguistic competence.

Recall the psycholinguistic evidence reviewed earlier for similarities between the interaction of the kinds of information processing we see in Figure 3, and that of two aspects of visual processing. On this conception:

- Both visual and linguistic processing include both bottom-up (visual parsing/word recognition, syntactic parsing) and top-down (expectation/QUD-based) systems, operating in parallel and feeding each other in real time in the course of interpreting sensory input.
- Both systems are pre-cognitive, cognitively impenetrable, and rapid, features typical of Fodorean mental modules.

And crucially, the two systems operate in parallel, with limited input into the bottom-up system from the top-down, expectation-driven system. The experimental evidence from anaphora resolution especially supports the view of these systems as distinct sub-modules of our linguistic competence.

So conceived, pragmatics is not a jumble of post-hoc rules for rational cooperation. Our semantic competence did not evolve and does not operate in a vacuum. Instead, pragmatics is a specifically linguistic system for tracking particular types of information, the latter fed on-line to bottom-up semantic interpretation in very specific ways.

Assume that there is a rule-by-rule relationship between syntax and semantics (Bach 1979). Then syntactic parsing to determine Logical Form is the bedrock of the bottom-up compositional semantics that drives and tightly constrains interpretation, yielding the usual static results. Just as with percepts in vision, the LF imposes a hard constraint on interpretation, which cannot be readily over-ridden. But in parallel, a rich, specifically linguistic context evolves in a way constrained by expectations and intention-recognition. This context feeds the compositional semantics dynamically, yielding a contextualized Logical Form wherein different sub-constituents typically have distinct local contexts of interpretation. This evolving context is directly fed by the bottom-up semantic interpretation of sub-constituents of an utterance. But it is also affected by conventional elements of the utterance which do *not* directly feed Logical Form:

- Context update is constrained by presuppositions, including especially prosodically-given presuppositions (Rooth 1992), imposing felicity conditions on context (as in Heim 1983), including the felicitous alternatives that compose the QUD; and
- Context is directly enriched by the addition of auxiliary content (more below).

Moreover, context not only tracks the CG and discourse referents, but a much richer intentional structure of the interlocutors' interaction, guiding expectations about what's under discussion and what is being said (Roberts 2004, 2013, 2023). This richer context comes to bear top-down during compositional interpretation to constrain interpretations so that they are contextually relevant, coherent and consistent, and thereby the evolving context plays a central role in disambiguation, anaphora resolution, presupposition accommodation, intrusive implicature generation, and speech act recognition. As with intentions in visual processing, expectations based on intentions tracked in linguistic discourse lead to a kind of inattentional blindness to irrelevant or infelicitous interpretations, not only yielding a more relevant interpretation, but speeding processing.

All elements of the architecture in Figure 3 are thus language-specific developments of a general cognitive strategy for processing perceptual input: The result is faster, more accurate and efficient processing of complex sensory input, with selective inattention constraining the reasonable "parses" of the purely perceptual information that is the foundation of the process.

The context of interpretation represented in Figure 3 is thus a hypothesis about what kinds of information we track in the course of interpretation, and of how that information is brought to bear on interpretation.

4.1 Context of interpretation

I assume an evolving context of interpretation, along the lines of the InfoStructure scoreboard in Roberts (1996/2012), which one might represent as an elaborated version of the DRSEs in Kamp & Reyle (1993).

The **elements of a context of interpretation K** include:

- CS, the context set (Stalnaker 1979);
- DR, the set of familiar discourse referents (Kamp 1981; Heim 1982);
- QUD, the set of questions under discussion (Roberts 1996/2012);
- G, a set representing the interlocutors' evident goals, intentions and plans (Roberts 2004, 2013, 2023; cf. the To-Do lists of Portner 2004, 2007); and
- \mathbb{C}_D , the set of discourse centers currently under discussion (Roberts 2014, 2020, to appear).

The inclusion of the set \mathbb{C}_D is motivated by work on the way that doxastic perspectives come to bear on interpretation. It is used to track the doxastic perspectives available (in DRT terms: accessible) at a given point in discourse. This set is updated and downdated dynamically, the dynamics modeled after the way the Reference Times are updated in Partee (1984):

The set of **discourse centers in context D**, \mathbb{C}_D :

A discourse center \mathbb{C} : an ordered set $\langle d_i, e_j \rangle$, s.t. $d_i, e_j \in DR_D$ and d_i is a doxastic agent whose beliefs in $e_j \in E_D$ are under discussion in D at speech time in the actual world.

\mathbb{C}_D is an ordered set. It always includes as its last element the unordered set of discourse centers involving the interlocutors $\{ \mathbb{C}_{i,j}^*, \mathbb{C}_{k,j}^@, \mathbb{C}^{*+@} \}$, where:

- $\mathbb{C}_{i,j}^*$ is a distinguished center corresponding to the speaker d_i in the actual event of utterance $e^* = e_j$,
- $\mathbb{C}_{k,j}^@$ corresponds to the addressee d_k in $e^* = e_j$,
- $\mathbb{C}_{i+k,j}^{CS}$ ($= \mathbb{C}_{i+k,j}^{*+@}$): inclusive 'we', the joint (purported) doxastic point of view of speaker and addressee $d_i \oplus d_k$ in $e^* = e_j$, as reflected in CS_D .⁷

Other elements of \mathbb{C}_D are discourse centers introduced in conjunction with the interlocutors' direct consideration of alternative doxastic states, i.e. attitudes:

- In the scope of attitude predicate *pred*, $\mathbb{C}_{i,j}^{pred}$ (e.g., \mathbb{C}^{know} , \mathbb{C}^{hope} , \mathbb{C}^{claim} , etc.) is introduced as the first element of the local \mathbb{C}_D . \mathbb{C}^{pred} : the ordered pair of the agent d_i of the attitude in the arbitrary eventuality e_j in which it obtains.
- In Free Indirect Discourse, reflecting the doxastic perspective of some third party: $\mathbb{C}_{i,j}^{FID}$, the agent d_i whose perspective in a given (possibly fictional) eventuality e_j is being adopted globally in the narrative.
- Perspectival adverbials like *according to a* (at time t) or *from a's point of view* (at time t): a new element $\mathbb{C}_{i,j}^{POV} = \langle a, e \rangle$, d_i the agent a in the (possibly arbitrary) eventuality e_j , the point of view is available at $t = \text{time}(e_j)$.
- In an epistemic conditional, the *if*-clause enriches a presupposed modal base, MB, which introduces a doxastic perspective, $\mathbb{C}_{i,j}^{episMB} = \langle d_i, e_j^{episMB} \rangle$, d_i the anchor of the epistemic modality, e_j^{episMB} the arbitrary eventuality where $\bigwedge MB$ is true. The *if*-clause introduces a modified perspective: the same anchoring agent, a new arbitrary eventuality, where the *if*-clause is true as well as $\bigwedge MB$: $\mathbb{C}_{i,j}^{if} = \langle d_i, e_j^{episMB+if} \rangle$.

When the scope of the operator triggering introduction of some \mathbb{C} is closed, \mathbb{C} itself is removed from \mathbb{C}_D . There is no other way to introduce discourse centers to \mathbb{C}_D in English.

⁷ where \oplus is the join operator of Link (1983).

It is important to note that all the elements of this context apart from the DR are independently motivated by phenomena apart from the anaphora resolution that principally concerns us here. The CS, QUD and G are the loci for update with the three main types of illocutionary acts:

Assertion: (following Stalnaker 1979)

If an assertion of α is accepted by the interlocutors in a discourse context K , $|\alpha|^K$ is added to CG.

Direction:⁸ (cf. related proposals in Roberts 2004,2018, Portner 2007)

If a proposed direction $\mathbf{!}_iP$ is accepted by the addressee i in a discourse D , then G_i and i 's associated evident plans are revised to include the realization, under the applicable circumstances, of $\mathbf{!}_iP$.

G_i is revised to remove the realization of $\mathbf{!}_iP$ once it *or the larger goals it subserves* are no longer potentially applicable (e.g., it has been realized, or else it is determined that it cannot be practically realized).

Interrogation: (Roberts 1996)

If a question posed by $?\alpha$ is accepted by the interlocutors in a discourse context K , then $|\alpha|^K$, a set of propositions, is added to the QUD.

A question is removed from QUD iff either its answer is entailed by CG or it is determined to be unanswerable.

And the QUD plays a central role in generating implicatures and determining salience due to its role in determining what is RELEVANT (cf. Grice's Relation) in determining salience and generating implicatures.

RELEVANCE: Since the QUD reflects the interlocutors' goals at any point in a discourse, in order for an utterance to be rationally cooperative it must address the QUD.

An utterance m **addresses a question** q iff m either contextually entails a partial answer to q (m is an assertion) or is part of a strategy to answer q (m is a question) or suggests an action to the addressee which, if carried out, might help to resolve q (m is a suggestion).

CG/CS also plays an important role in presupposition satisfaction and projection, as well as in implicature generation. And G plays a role in joint plan development and recognition, including the inference of a RELEVANT Ordering Source for deontic modals.

The salient discourse centers at a given time in discourse serve to anchor a wide range perspective-sensitive expressions, including *inter alia* deictic motion verbs (Barlew 2017), and epistemic modals (Roberts in press), as well as indexicals of all types & shifting indexicals (Roberts 2014); and, as we will soon see, auxiliary content.

⁸ This characterization of Direction does not cover what Kaufmann (2012) calls Expressive uses of imperatives, like *Be well!* See Roberts (2015) for discussion.

4.2 Dynamically updated context feeds contextualized LF

The **context updates highlighted in blue** in Figure 3 are modeled after the way that the CCPs of particular constructions in Heim (1982, 1983) dynamically update context. As in Heim, these updates enable donkey anaphora and presupposition filtering. The updates are simplified here for ease of exposition, initially considering only the way in which CS_D is updated, with CS treated as in Heim (1982) as a set of world/assignment pairs:

$$CS_K = \{ \langle w, g \rangle : \text{all the familiar discourse referents are assigned values by } g \text{ in } w \text{ that are consistent with the information interlocutors share about them} \}$$

As in Kamp (1981) and Heim (1982):

- the **global context** of utterance K typically serves as the default local context of interpretation for simple clauses, the prejacent of a negated clause, the first conjunct of a conjunction, and the *if*-clause of a conditional.⁹
- When we add the content of an indicative clause S to context K , we intersect with K the proposition expressed by S in that context to yield a new context $K+S$.
- The update for negation yields the complement in the global context K of the result of interpreting the prejacent of negation in K , thus effectively removing any worlds from CS_K in which S is true while guaranteeing that any presuppositions triggered in S are satisfied in K .
- In both conjunction and the conditional, the second conjunct/consequent are interpreted with respect to a context that is the result of updating the global context K with the content of the first conjunct/antecedent. As in Heim/Kamp, the update with a conditional antecedent is merely provisional, so that the resulting updated context after the whole conditional has been interpreted *removes* worlds from K in which the antecedent is true but the consequent is false. But the update with the first conjunct of a conjunction is stable, persisting in the result of update with the whole conjunction.

In the next lecture we'll consider update with disjunctions.

Unlike in Heim (1982), what's dynamic in these update rules is not the semantic content of the constituents themselves, but the way the *context* differs for different sub-constituents. Notably, though the dynamics of context update yields the presupposition-satisfaction asymmetry in conjunction, whereby the presuppositions of the second conjunct can be satisfied by content in the first, but not vice versa, the semantics for *and* is classical and static, the resulting interpretation commutative, unlike in Heim (1982) or Groenendijk & Stokhof (1990). So:

$$[S_1]_K \wedge [S_2]_{K+S_1} \equiv [S_2]_{K+S_1} \wedge [S_1]_K$$

What's asymmetric is that the interpretation of the second conjunct takes the content of the first as part of its local context.

⁹ This is not the case for utterances in Free Indirect Discourse or involving modal subordination, but we'll ignore those for now.

4.3 Lexical CHARACTER and how it feeds interpretation

Unlike Kaplan's Character, lexical CHARACTER as given at the top left of Figure 3 is intended to reflect the fact that a lexical entry may have quite a rich body of associated semantico-pragmatic content:

- Only semantic content directly feeds compositional semantics, in logical form.
- Anaphoric presuppositions constrain contextual felicity, as in dynamic semantics. Like non-indexicals, indexicals crucially involve anaphoric presuppositions which must be locally satisfied—in English for *I* and *you*, by the actual speaker/addressee, but in shifting indexical languages possibly by other salient perspectival centers in discourse (Roberts 2014,2020).
- Auxiliary content directly feeds context update, as we'll see below. This is similar to the way that Potts (2005) models the contribution of auxiliary content to interpretation: His conventional implicatures contribute directly to information about the speaker's beliefs, without contributing to the proffered content of the utterance in which they are conventionally triggered. But on the present account, auxiliary content is instead anaphorically anchored to a salient, relevant discourse center, which needn't always be the speaker.

5. Deriving intrusive pragmatic phenomena

5.1 Deriving donkey anaphora:

Contextualized LFs reflect the ways that sub-constituents of an utterance may contribute to the local context for interpretation of other sub-constituents, as given by the **update rules**. We can see how this accounts for the possibility of donkey anaphora via the contextualized LF for (1) in (1'), with the updated context for the main clause highlighted in yellow:

(1') ${}^K[s_0 \text{ if } {}^K[s_1 \text{ a farmer}_1 \text{ owns a donkey}_2 \text{ s}_1] \text{ }^{K+S_1}[s_2 \text{ he}_1 \text{ uses it}_2 \text{ to plow } s_2]]_{s_0}$

Here is how context revision is calculated in the interpretation of (1'), with resulting **CS updates**:

1. $K+(1) = K + [if [S_1]^K \text{ then } [S_2]^{K+S_1}]^K = K \setminus (K+||S_1||^K \setminus K+||S_1||^K+||S_2||^{K+S_1})$
2. $K+S_1$: update DR_K to add d_1, d_2 ; update CS_K to yield $CS_{K+S_1} = CS_K \cap \{ \langle w', g \rangle \mid g(d_1) \text{ is a farmer in } w', g(d_2) \text{ is a donkey in } w', g(d_1) \text{ owns } g(d_2) \text{ in } w' \}$
3. $K+||S_1||^K+||S_2||^{K+S_1}$: update CS_{K+S_1} to yield $CS_{K+S_1} \cap \{ \langle w', g \rangle \mid g(d_1) \text{ uses } g(d_2) \text{ to plow in } w' \} = \{ \langle w', g \rangle \text{ in } w' : g(d_1) \text{ is a farmer \& } g(d_2) \text{ is a donkey \& } g(d_1) \text{ owns } g(d_2) \& g(d_1) \text{ uses } g(d_2) \text{ to plow} \}$
Substituting 2 and 3 into 1 yields 4:
4. $K+(1) = CS_K \setminus (\{ \langle w', g \rangle \mid g(d_1) \text{ is a farmer in } w', g(d_2) \text{ is a donkey in } w', g(d_1) \text{ owns } g(d_2) \text{ in } w' \} \setminus \{ \langle w', g \rangle \text{ in } w' : g(d_1) \text{ is a farmer \& } g(d_2) \text{ is a donkey \& } g(d_1) \text{ owns } g(d_2) \& g(d_1) \text{ uses } g(d_2) \text{ to plow} \})$

In step 2, we see that not only is the CS updated with the content of S_1 , but per usual in Heim/Kamp and other dynamic semantic accounts, we update the set DR_K with new dRefs for the indefinites *a farmer*₁ and *a donkey*₂, d_1 and d_2 .

If CS_K is the set of all $\langle w, g \rangle$ pairs (so that there is as yet no shared information in the discourse), the set of worlds that are first elements in $K+(1)$ is the proposition expressed, (1)'s proffered content: 'the set of worlds in which either no farmer owns a donkey, or there are farmers that own donkeys and they use the donkeys they own to plow'.¹⁰ Since S has no other context-sensitive elements, this is the proposition expressed in that context. If there is already shared information reflected in CS, so that the domain is globally restricted, this will just be *the set of worlds in the domain in which...* In any case, the result of that interpretation is added to the global context for the next utterance. But since *a farmer* and *a donkey* are ROOFED by the conditional in (1), the dRefs introduced in the course of its interpretation do not persist in the final interpretation, so do not license anaphora in subsequent discourse.

(3) is a donkey sentence involving anaphoric bridging:

(3) If ${}^K[s_1 \text{ a farmer}_1 \text{ owns a John Deere}_5 \text{ s}_1], \text{ }^{K+S_1}[s_2 \text{ he}_1 \text{ uses the tractor}_5 \text{ to plow } s_2]$.

In determining the local context for the main clause of (3), the need to retrieve an antecedent for the definite description *the tractor* drives an abductive inference based on world knowledge that *John Deere* is the brand-name of a famous tractor manufacturer, plus the assumption that the *if*-clause is relevant to the consequent: We infer that the arbitrary farmer owns a John Deere tractor, introducing a discourse referent d_5 with the information that d_5 is such a tractor. Updating K with S_1 plus this accommodated inference then makes directly available d_5 to serve as a weakly familiar antecedent for *the tractor* in S_2 .

¹⁰ This is, of course, the so-called universal (or "strong") reading of (1), which is what Heim's rules yield, along with the proportion problem. See my Brazil Lecture One for explanation. I do not mean to suggest that that is the correct interpretation of the sentence. See Roberts (2005) for discussion and citation of a great deal of relevant literature, King & Lewis (2021) for more recent work. Heim's rules are used here simply for illustration.

5.2 Dynamic update with auxiliary content

Auxiliary content, contributed by appositive NPs, nonrestrictive relative clauses (NRRs), and a variety of other triggers (Potts 2005) makes no direct contribution to the proffered content of the utterance in which it occurs, but nonetheless it may have a truth conditional impact in a dynamic fashion, straightforwardly afforded by this architecture.

Properties of auxiliary content:

The first two properties of auxiliary contents were mentioned above, but others are also worth noting.

- i. Auxiliary content doesn't directly contribute to the proffered content of the clause in which it occurs at LF, as we saw in (25). (Potts 2005)

- (25) a) Sam believed that his wife, a very honest woman, was lying to him.
b) #Sam's wife isn't an honest woman.
c) ✓Sam's wife wasn't lying to him.

- ii. Auxiliary content cannot be bound by/bind semantic content, as we saw in (26). (Potts 2005)

- (26) *No reporter₁ believes that Ames, often the subject of his₁ columns, is a spy. [Potts 2005]

On the basis of examples like (25) and (26), Potts concludes that auxiliary content does not take scope at LF; Schlenker (2013, 2021) offers apparent counter-evidence, but I think those examples can be understood to involve modal subordination rather than scope at LF.

- iii. Chierchia & McConnell-Ginet 2000) observed that although the auxiliary content of an NRR is projective, as displayed in the family of sentences test in (29), unlike presuppositions NRRs and appositives carry an anti-novelty implication, as we see in (30). Hence, auxiliary content is not presupposed:

- (29) a. Monty, who's from Kentucky, doesn't like corn grits.
b. Does Monty, who's from Kentucky, like corn grits?
c. If Monty, who's from Kentucky, likes corn grits, it isn't surprising.
d. It's false that Monte, who's from Kentucky, like corn grits.

(a) – (d) all commit the speaker to the truth of the proposition that Monte is from Kentucky: The auxiliary content projects.

- (30) Monte was born in Louisville, Kentucky.
...
#Monte, who's from Kentucky, loves corn grits.
#Monte, a Kentuckian, loves corn grits.

iv. Auxiliary content can't by itself serve to make an utterance RELEVANT to the QUD, hence is not part of what is proffered by an utterance in which it occurs. (Amaral et al. 2007)

(31) Chris: What are Harvey's bad habits?

Tom: ?#Harvey, a heavy smoker, is from Kentucky.

Together, properties (i) – (iv) support Potts' argument that the implications in question are neither semantic content nor presuppositions. But properties (v) and (vi) were not noted by Potts, and are crucial to the analysis I will propose:

v. Auxiliary content is perspectival: anchored to a salient doxastic perspective, which (*pace* Potts) is not always that of the speaker (Amaral et al. 2007), as confirmed experimentally by Harris & Potts (2009). So in (25) the appositive is anchored to the speaker, but in (32) the NRR *who has an awful lot of legos* is anchored to the arbitrary kid on the arbitrary play visit, and in (33) the appositives are both anchored to the agent of belief, Lois.

(32) [The speaker is a child psychologist studying play activity, with young subjects.]
Whenever I play with kids₂ in a poorly equipped daycare center₃, they₂ clearly hope that this stranger₄, who has an awful lot of legos₅, will leave some of them₅ behind when she₄ goes.

(33) a) Lois believes that Clark Kent, the milquetoast reporter, is attractive. (False)

b) Lois believes that Clark Kent, the superhero, is attractive. (True)

[Zsófia Zvolenszky, p.c., after Saul 1997]

vi. Auxiliary content may contribute indirectly to the truth conditional content of a clause not only via anaphoric dependencies, as in (27) (repeated below) and (32), but non-anaphorically, as in (33). In the latter, the appositives suggest the (contradictory) ways that the agent Lois views Clark Kent which lead her to conclude that he is or is not attractive.

(27) [Every professional man I polled]₁ said that while [his₁ wife]₂, who had earned a bachelor's degree₃, had no work experience, he₁ thought she₂ could use it₃ to get a good job if she₂ needed one. [ARS 2007]

The interpretive contribution of auxiliary content

Constituents that contribute auxiliary content do show up in LF, because they are part of the syntactic structure of the utterance in which they occur. Among other things, their interpretation depends on what they modify in that structure: The denotation of the NP modified by an appositive or NRR serves as argument of the auxiliary content contributed. But a constituent contributing auxiliary content is so-marked in some fashion. E.g., in English, appositives and NRRs are typically marked by pauses on either edge, a prosodic feature Potts (2005) called COMMA; Schlenker (2013) claims that French *lequel* is a relative pronoun that can only occur in NRRs. A constituent so-marked does not contribute to compositional semantics directly, but instead directly contributes to context-update, as follows:

In the absence of such rejection, auxiliary content in (32) updates local context, as follows:

(32') Contextualized LF of (13), © anchored to the arbitrary kids d_2 , updates highlighted:
 K [S_0 Whenever [S_1 I play with kids $_2$ in [a poorly-equipped-center] $_3$ S_1] $K+S_1$ [S_2 they $_2$ hope
 $K+S_1+©hope$ [S_3 that [NP_4 this stranger NP_4] $K+S_1+©hope$ [$COMMA_4$ who has [a lot of legos] $_5$ $COMMA_4$]
 $K+S_1+©hope+COMMA_4$ [VP_3 will leave some of them $_5$ when she $_4$ goes VP_3] S_3] S_2] S_0]

At the outset in the global context of interpretation K for (32), as usual in non-FID contexts the set of discourse centers includes only the speaker and addressee at utterance time. Take *whenever* to be a quantifier over eventualities. Its restriction S_1 introduces discourse referents e_1 for the arbitrary playing event and d_2 for the kids the speaker plays with in e_1 . Then, as usual in such constructions in dynamic semantics, the semantic content of adverbial S_1 is part of the context of interpretation for the nuclear scope S_2 , with d_2 serving as antecedent for *they* and the event time of e_1 as the reference time for the clause—the kids are the agents of the hoping, which occurs during the speaker’s arbitrary visit. Under the scope of *hope*, the discourse center $©_{2,1}^{hope} = \langle d_2, e_1 \rangle$ (the arbitrary hoping agent/kid in the play situation) is introduced to the set of discourse centers, yielding the updated context $K+S_1+©hope$ for the interpretation of the complement clause.

The complement includes the NRR consisting of $COMMA_4$ and the relative clause. The use of *this stranger* in the NRR implicates that it reflects the perspective of the arbitrary kid, so that the auxiliary content is anchored to $©^{hope}$, satisfying the anchoring presupposition of $COMMA_4$. The resulting auxiliary content doesn’t contribute directly to the semantic content of the complement. Rather the auxiliary proposition ‘this stranger has a lot of legos’, along with a discourse referent d_5 for the legos, is added to $DOX(\langle d_2, e_1 \rangle)$ —the set of beliefs attributed to the arbitrary kid at the time of the visit—enriching the local context of interpretation for the remaining nuclear scope, VP_3 .

Crucially, because it is part of the complement of *hope*, VP_3 describes the hopes of the kids. But hopes are founded on beliefs, as we saw in (10) and (11) above (the squirrel examples). So in contributing information about the belief state of the kids, the NRR contributes updated information about their beliefs, hence contributes non-trivially to the local context for interpretation of VP_3 . The auxiliary content then provides the locally available antecedent d_5 for *them*, the legos. Since the end of S_0 marks the limit of the scope of both *hope* and *whenever*, at that point the content of the material under their scopes is removed from the resulting context and the set of discourse centers is down-dated to remove $©^{hope}$. But the proposition calculated partly on the basis of that content is added to the global context of interpretation K : ‘whenever the speaker plays with kids in a poorly equipped center, the kids think that this stranger has a lot of legos and they hope that she will leave some of the legos when she leaves’.

So paraphrased, the appositive appears to take narrow scope relative to both *hope* and *whenever*. But in the derivation, the wide syntactic scope of *hope* only matters because it introduces the anchor $©^{hope}$. In turn, the agent of $©^{hope}$ d_2 is introduced by *kids* $_2$, which itself takes narrow scope relative to *whenever*. Since the anchor is thus *arbitrary*, the auxiliary content is ROOFED by the modal operator in the lexical content of *hope*, as well as that of *whenever*, and so does not contribute to the global context. In other words, this apparent narrow scope need not be

represented at LF, but is a reflection of merely local presupposition satisfaction for the anchor of the presupposed ©.

One more point about (32): The subject embedded under *hope*, *this stranger*, is clearly intended to refer *de dicto* to the speaker from the perspective of the kids. I have proposed elsewhere (Roberts 2014) an account of nominal content and belief attribution based on Aloní (2001), wherein nominal content is always understood to reflect a guise of the intended denotatum from the perspective of a salient discourse center. Aloní argues that in examples involving the classic puzzles about belief attribution, we capture the intended interpretation via a pragmatic perspective-shifting operator \wp : Her \wp takes a variable (that associated with the referential index on an NP) and yields a conceptual cover, a method of individuating entities which yields, for the index in question, what we might call the relevant guise of the *res* denoted by the NP. Roberts (2014) modifies \wp to require that it be anchored to a salient discourse center; in (32), the anchor is the salient ©^{hope}, the same arbitrary kid that anchors the auxiliary NRR. The result is that *this stranger* means something like ‘the individual who, from the perspective of the arbitrary kid during the visit, is the salient, present (because of the proximal presupposition of *this*) stranger in that circumstance’, thus yielding a *de dicto* description of the speaker in her role as visitor in that eventuality.

Then with no additional assumptions, we can explain the truth conditional effects of the appositives in (33):

- (33) a) Lois believes that Clark Kents, the milquetoast reporter, is attractive. (False)
b) Lois believes that Clark Kents, the superhero, is attractive. (True)
[Zsófia Zvolenszky, p.c., after Saul 1997]

We take the nominal appositives *the milquetoast reporter* and *the superhero* to be anchored to the agent of *believes*, Lois, and hence they contribute information to the context of utterance about Lois’ beliefs. Lois believes both that the denotation of the modified head, who the interlocutors know by the name Clark Kent, is a milquetoast reporter—hence, a wimp—and that he’s a superhero—hence, strong and brave. Since these are incompatible traits, we assume that these are guises which Lois attributes to the familiar Kent without realizing that both guises describe the same *res*. Her beliefs aren’t irrational; they’re just ill-informed (due to Kent’s own misrepresentations in the Superman story). So the appositives suggest guises of Kent with respect to which Lois holds her opinions about his attractiveness. They offer explicit content for the guise given by Aloní’s \wp in this case, which operates on the rigid designator *Clark Kent* to pragmatically yield the attested interpretations, in turn explaining the attested judgments of truth and falsity.

Again, this relativization of content to an available doxastic perspective is much more general than just the application described here to auxiliary content. I argue in Roberts (to appear) that epistemic modal operators are quite generally anchored to some doxastic perspective available in discourse, using discourse centers, as above. Hence, all the tools used in this analysis are independently motivated. The only innovation is the dynamic pragmatics itself and the treatment of auxiliary content as anchored to a discourse center and contributing to the local context of interpretation. Since Potts (2005) had already argued for the direct contribution of conventional

implicatures to context update, this way of modeling their contribution is just a way of using independently motivated dynamic pragmatics to capture Potts' idea, while permitting the merely local anchoring he had incorrectly taken to be impossible.

Finally, consider (34):

- (34) [Background: In (a) – (d) the speaker is gossiping about a prominent community member named Lauren and her son Spenser. The speaker doesn't know Lauren personally, but they have several mutual friends. The interlocutors know that Spenser has a boyfriend named George.]
- a) Spenser seems poised to propose to George.
 - b) Most of her friends think that Lauren, who would enthusiastically approve, would be misguided in that endorsement:
 - c) George is handsome and intelligent, but he isn't kind or loving.
 - d) I myself don't know George, but he certainly hasn't made many friends around here.

The meaning of (34b):

'most of Lauren's friends believe that if Spenser proposes, Lauren would enthusiastically approve of the proposal; and they think that if Spenser proposes and Lauren approves Lauren would be misguided in her endorsement of the proposal. They believe this because they believe that though George is handsome and intelligent, he is neither kind nor loving'

The highlighted elements in this paraphrase are not explicitly given in (b), but inferred as discussed below.

There are several notable features of this interpretation of (34b).

- i. The most likely understood anchor of the NRR *who would enthusiastically approve* is the arbitrary friend of Lauren; The speaker herself is understood not to know the principals well enough to know how Lauren would react.
- ii. The NRR predicate *approve* has an elided complement, which must be retrieved from context; it also contains auxiliary *would* which requires domain restriction. If we take *would* to be modally subordinate to the possible proposal from (a): 'if Spenser proposed to George', this supplies a plausible value for the ellipsis: 'approve of the proposal'. In turn this implicates that the speaker believes that most of Lauren's friends know of the possible pending proposal. Update with this context presumably takes place immediately on interpretation of the subject and NRR, so that it now locally entails all this information about the arbitrary friend of Lauren.
- iii. In the main clause, *would* is understood to be modally subordinate to the content of the NRR, and hence in turn to the possible proposal in (a): 'if Spenser proposes and Lauren enthusiastically approves'. Arguably this subordination does not require accommodation, given the update with the NRR. Then in the complement of the main verb *think*, the antecedent of *that endorsement* is 'Lauren's enthusiastic approval of the proposal' from the modally subordinate NRR, so that an anaphoric presupposition in the main clause is satisfied by the content of the NRR. Thus, taking main clause *would* to be modally

subordinate to the NRR both makes the NRR relevant to the main, proffered content, and also offers an antecedent for the descriptively abridged definite description.

- iv. Finally, (c) is understood as (i) an opinion of the arbitrary member of the group composed of those of Lauren's friends who disapprove of Spenser proposing to George, and (ii) rhetorically, an explanation for the disapproval: 'they say that he is handsome...'. On this interpretation, the speaker isn't committed to the truth of this assessment of George, permitting consistency with the first clause of (d).

Hence, this complex but readily comprehensible example involves (i) anchoring NRR content to an arbitrary agent under quantification, (ii) the modal subordination of the NRR to the irrealis content of the prior utterance (a), in turn licensing anaphora resolution in the NRR, (iii) modal subordination of the main clause *would* to the NRR, in turn licensing anaphora resolution in the main clause, and (iv) an FID-like extension of the arbitrary friend's perspective on the matter to the second sentence, so that the speaker isn't committed to its truth in the actual world. Hence, the global perspective in (34) is that of the speaker, but in (b), under the scope of *most of her friends* and in the FID-like continuation in (c), the reported perspective is that of the arbitrary friend of Lauren, both modally subordinate to the possible proposal and with all the resulting auxiliary content dynamically contributing to information about that perspective. This kind of example offers strong support for the claim that context update via both proffered semantic content and auxiliary content is dynamic, taking place rapidly in real time as the utterance is processed, and that this update crucially involves the update of salient perspectives.

6. Conclusions

6.1 Logical form and Discourse Representations

The remarks above were prepared for the annual lecture in honor of the Dutch logician Evert W. Beth. But in response, my colleague Frank Veltman asked: Where's the logic?

There are two prongs to my response:

First, like Grice (1967) I am concerned that we not develop overly-complicated logical forms for linguistic expressions when the attested complications arguably arise not from their semantic content, but from the ways in which they are used in discourse. No one who studies natural language semantics doubts the importance of context in the determination of what is meant. But not all would agree about what constitutes a context of interpretation, about what kinds of contextual information are needed to model how natural language is actually interpreted. I have argued here (and elsewhere) that in order to capture the complex ways in which context comes to bear on interpretation, it should contain more than the classical assignment functions, common ground, and indexical values; there is now a growing body of empirical (including experimental) evidence that interlocutors also track the QUD, the evident goals and intentions to which they are committed, and the salient doxastic perspectives under discussion. In addition, here I have added new empirical support for the thesis that the interaction between compositional interpretation and discourse context is dynamic. But I have argued that instead of changing the semantics of the expressions interpreted—their contributions to logical form—it is preferable to change our

understanding of the way that compositional interpretation draws on contextual information. In the human interpretation of sensory input generally, we find a general cognitive strategy of using parallel bottom-up parsing of input sensory data with top-down contextual input—driven by expectations based on intention-recognition—to yield rapid, accurate interpretation. I contend that only when we recognize how this more general cognitive strategy for interpretation is realized in the interpretation of linguistic expressions can the true logical form of the expressions themselves shine forth, elegant and simple.

But back to Frank’s question: We do still need to understand how in this interpretive process logical constraints are brought to bear on contextual factors in real time. And here, fortunately, I think we can address this question without re-inventing the wheel. Just as the update rules in Figure 3 replicate Heim’s (1983) rules for context update in dynamic semantics, so we can use Discourse Representation Theory as a representation of contexts, relying on the proof theory for DRSeS worked out by Kamp & Reyle (1993, 1996) to explain attested inferences and logical constraints.

Of course, the DRS construction rules in Kamp & Reyle (1993) are written as rules for directly translating syntactic phrase structure trees into DRS interpretations. But I see no reason why one could not instead take DRSeS to represent contexts of the sort proposed above, dynamically updated with the output of a static, rule-by-rule compositional semantics. Classical DRT already contains both discourse referents and conditions, and the latter are used by many authors (e.g. Roberts, 1989, Kadmon 2001) to represent the content of the interlocutors’ Common Ground. Roberts (1989) worked out an intensional DRT along those lines for modal subordination; other intensional versions have been proposed by Asher (1986) and Maier (2010, 2016). Adding discourse centers to an intensional DRS along those lines would be no more complex than Partee’s (1984) addition to DRT of Reference Times and rules for their update/downdate. To add the QUD to such a context would require elaboration of the classical DRTs, but the work on dynamic inquisitive semantics (Dotlačil & Roelofsen 2019) should prove useful there.

On this way of representing contexts, the static semantics would work off of Logical Forms that are contextualized to draw from DRS “contexts”—both global and local (sub-DRSeS under operators)—leading to updated contexts that feed other sub-constituents and subsequent utterances. And the addition of non-linguistically given content to a DRS can be appropriately constrained by pragmatic principles, including those for relevance (Roberts 1996), coherence (e.g., Asher & Lascarides, Roberts 2016), and, of course, consistency. Then for consistency, and for drawing pragmatic inferences, we can rely on the Kamp & Reyle (1993, 1996) proof theory.

Of course, all this would have to be worked out in detail. But I see no reason not to build on the work in DRT to answer Veltman’s question.

6.2 Summary and prospects

Here we reviewed several challenges to earlier views of the semantics/pragmatics interface:

- intrusive Phase 1 phenomena: donkey anaphora, shifting indexicals

- intrusive Phase 2 phenomena: donkey bridging, modal subordination, intrusive implicature
- intrusive auxiliary content: dynamic contextual-dependence (both phase 1 and 2) between non-proffered, auxiliary content and the output of the semantic composition of proffered content, as in (18).

These challenges are met by a view of the semantics-pragmatics interface which involves:

- (a) a rich notion of a context of interpretation, tracking the kinds of information shown to have a consistent contextual influence on interpretation;
- (b) rules for the evolution of context across an utterance, operating (comparably to the parallel architecture in vision competence) on simultaneous processes:
 - bottom-up dynamic input from the grammar: compositional semantic output for sub-constituents of the utterance in question; and
 - top-down constraints and affordances from linguistic mechanisms designed to yield in run time the effect of abductive application of rationally motivated principles and constraints, without actually involving inferential processes, e.g.:
 - inattentional blindness (Simons & Chabris 1999) resulting from focus on the QUD/task structure of discourse yields
 - a metric for relevance (Roberts 1996/2012), a constraint yielding implicature
 - salience (Roberts 2016, drawing on Polanyi 1985), inducing restriction of domains for reference, anaphora resolution and quantification
 - the determination of modal and imperative flavor (Roberts 1989, 2018, 2023, to appear; Moss 2015)
 - contrastive prosodic cues (themselves related to the QUD, hence universally serving as clues to task, Roberts 1998) give rise to referential expectations, scalar implicatures (Rooth 1992), etc.: a structural mechanism reinforced by Relevance to the QUD
 - constraints on the relationship between QUD and G, to yield imperative flavor and the deontic implications of directives (Portner 2004, Roberts 2018)
 - lexical priming and spreading lexical associations triggered by lexical content (Simons & Danks 2020), which contribute to word recognition, disambiguation, bridging (donkey bridging being an utterance-internal version), and enrichment implicatures

This approach allows for the simplification and/or explanation of features of interpretation elsewhere treated as syntactically unmotivated principles of the grammar. For example:

- We can derive Heim's Maximize Presupposition as a requirement to maximize cohesion across moves.
- Constraints on rhetorical relations can similarly be seen to characterize well-formed, maximally cohesive strategies of inquiry in the QUD (Roberts 2016).
- Some features commonly included in LF are arguably redundant and/or empirically inadequate e.g.: Hacquard's (2010) binding of epistemic modal auxiliaries (as argued by Roberts, to appear), Chierchia's (2004) use of a tacit operator EXH to generate intrusive scalar implicatures, Schlenker's (2013, 2021) scope of NRRs at LF, etc.

Thus, the account tightly constrains context update/downdate in the course of interpretation, avoiding the over-generation associated with free pragmatic enrichment (Sperber & Wilson 1986, Recanati 2004, 2010). Linguistic pragmatics is an interface between grammatically given content and higher cognitive function, specifically designed for that linguistic purpose (e.g., as evident in the reflection of the QUD in prosodic focus), but is not itself a part of the grammar (*pace* Lepore & Stone 2015, Stojnić 2021).

This evidence argues, *pace* K. Lewis (2017) *inter alia*, that linguistic competence consists not just of knowledge of grammar—phonology, syntax, compositional semantics—and its realization in a given language, but also crucially includes the capacity to wield and update a specifically linguistic context in the course of interpretation.

Like UG, formal pragmatics so-conceived is a theory of competence, not performance:

- There are many ways the mind might store the kinds of information that we seem to track in order to interpret quickly and efficiently. The scoreboard is just an abstract characterization of certain kinds of information that we arguably need ready access to and track in the course of interpretation.
- There are many ways that the posited pragmatic constraints might be realized in an actual implementation, including statistically or in a connectionist system. The constraints in pragmatic theory just tell us what such an implementation must *do* in order to realize this competence.

Finally, we have taken Bar Hillel's warning seriously: The pragmatics proposed here is not a wastebasket. Rather, we have been at pains to argue that the contents of a context of interpretation are organized in very specific ways that serve to constrain and facilitate rapid retrieval of the speaker's reasonably intended meaning. Insofar as this approach has empirical substance, as I have argued above, then we are indeed embarked on the explication of the role of context in interpretation.