Craige Roberts
NASSLLI 2012, Austin

QUESTIONS IN DISCOURSE
Lecture 4: RELEVANCE, Salience and Anaphora Resolution

Last time we talked about Matter and Anti-Matter, sketching a view of projection as driven by Relevance to the QUD, partly via prosodic congruence.

In Tonhauser et al. (2012) we offer a range of diagnostics to differentiate several classes of meanings with respect to:

- whether they tend to project
- whether they impose a contextual felicity constraint on the context of utterance (basically, requiring prior familiarity or entailment—cf. the definedness conditions of Heim 1982)
- whether they have a local effect, i.e., contribute to the proffered content of the constituent in which they occur

In David Beaver's slides from yesterday, slide 44 summarizes the properties of three classes of Anti-Matter, in terms of these three properties, and in slide 45 he exemplifies these, and further differentiates the backgrounded triggers into those which arise conventionally and those which do not.

Today we're going to focus on one of these classes, the anaphoric elements which impose felicity constraints on prior discourse. The anaphoric triggers are a typologically diverse group, including pronouns, definite descriptions, additives like *too*, ellipses, and domain restriction. They all impose on prior context the requirement of an antecedent. This antecedent must be retrieved—the anaphora resolved—in order to grasp the intended contribution of the anaphoric trigger to proffered content.

Recall that from the Gricean perspective on meaning, Retrievability is a strong constraint:

> **Retrievability**: In order for an utterance to be a rational, cooperative act in a discourse interaction *D*, it must be reasonable for the speaker to expect that the addressee can grasp the speaker's intended meaning in so-uttering in *D*.

In view of the gap between conventional content and conveyed meaning, this is a strong principle, since it requires that cooperative speakers expect that their *uniquely* intended meaning_{nn} can be recognized as such by an addressee. Thus, we posed the question:

> **How do addressees regularly, easily retrieve a speaker's intended meaning, given that it is underdetermined by the conventional content of what she says?**

I.e.: What might be the grounds for the assumption of Retrievability on the part of a speaker?

Today I want to speak to that question, and to sketch for you a QUD-based notion of **salience** and talk about how it bears on **anaphora resolution**. In the interest of time, I'll focus on the anaphora triggered by pronouns and so-called *incomplete* definite descriptions—those like *the cup* whose descriptive content underdetermines their intended denotation. But in other work I have argued that Relevance to the QUD plays a central role in domain restriction and in the felicity conditions on Sluicing and VP Ellipsis.

I. Anaphora Resolution and Salience

Consider these three inter-related theses about the process of interpretation:

(a) **Dynamic interpretation**: Those aspects of an utterance which have already been interpreted bear on the interpretation of those still being processed. Hence, context changes non-globally.
(b) **Simultaneous solution**: Pragmatic reasoning plays a crucial role in the course of this dynamic interpretation. This is evident in the fact that interpretation of any single element of an utterance which is underdetermined by its conventional content must be compatible with the resolution of all such elements in the utterance, as well as resulting in a plausible, contextually felicitous interpretation.
(c) **The centrality of Relevance**: Interpretation is driven and constrained by the interlocutors' publicly evident intentions and goals, as reflected in the requirement of Relevance to the QUD. The interlocutors' recognition of and cooperative commitment to those intentions is essential to their collaboration in conveying and Retrieving meaning$_{nn}$.

(1) A move *m* is **Relevant** to the question under discussion *q* iff *m* addresses *q*, directly or indirectly yielding a partial answer to *q* in CG. (Roberts 1996)
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, introduced with utterance of an imperative).

In this section, I will first sketch a functional explanation of (c) and suggest how we might approach a theory of Salience based on this hypothesis. If the latter proved to be useful, that in turn might lay the foundation of an account of how, in processing an utterance in context, interlocutors can efficiently seek a simultaneous solution (b) in the course of interpretation (a).

Here is an outline of the explanation:

(1) Human behavior generally is driven and constrained by intentions, those intentions for any given agent ranked according to the relative importance to the agent of realizing the intentions. (I will simply assume that (1) is true.)
(2) What we intend constrains what we attend to, our attentional field.
(3) Salience (like commitment to intentions, a gradient notion) is a function of attention. I.e., something is salient to someone to the extent that they are attending to it. Thus, we might

expect that some entities (those associated with more highly ranked intentions) would be more central in the attentional field than others.

(4) Salience facilitates the Retrieval of those aspects of the intended meaning of an utterance which are not explicitly given by the conventional content of the utterance.

From (1) – (4) it follows that it would generally be extremely useful at any point in discourse to have a publicly evident goal which the collaborative interlocutors jointly intend to achieve, and which is more highly ranked than any of their other joint intentions. Since we take the general function of discourse to be to contribute to the interlocutors' Common Ground, it would seem reasonable to take as a goal a circumscribed domain of inquiry. This is exactly what a question offers: a set of alternative possibilities from which one attempts to derive the one true answer, the realized possibility. Hence, on the reasonable assumption (Roberts 1995,2004, Ginzburg 1995,2012) that the default goal at any time in discourse is to address the QUD, a goal which all cooperative interlocutors intend to address insofar as possible, and that doing so is what constitutes Relevance (see Handout 1 from this course), then (c) follows from (1)-(4). That is, we *predict* that the QUD will play a role in a wide range of phenomena involving interpretive resolution: *Inter alia* we predict that discourse salience is a function of the QUD (at least when interlocutors are not distracted), which thereby constrains the Retrieval of the intended meaning of an utterance. Another way to put this is that the intentional structure of a discourse, especially the QUD, constrains the search for meaning$_{nn}$. We seek a most-obviously Relevant way of yielding a simultaneous solution to the interpretive puzzles in an utterance, without necessarily working through all potential solutions in parallel.

To put some more meat on these bones, consider the following outline of a theory of salience, based on the QUD-intentional structure framework. For concreteness, consider first the role of salience in anaphora resolution. Though there is still debate about the meaning of definite descriptions, there is an emerging consensus (even on the part of E-type theorists like Elbourne) that (a) they are anaphoric, presupposing an antecedent in prior discourse—along the lines proposed by Heim (1982), Kamp (1981) or van der Sandt (1992), and (b) their descriptive content is presuppositional, as well, in the sense that their presupposed antecedent is presumed by the speaker to bear the property denoted by that descriptive content.[1] These theories of discourse anaphora take the antecedents to be not NPs *per se*, but **discourse referents**—as that notion is spelled out in the Heim/Kamp/van der Sandt theories. In interpreting a definite, an addressee must determine exactly which antecedent the speaker intends, out of all those familiar to the interlocutor. The NP's descriptive content is a both a constraint on and a clue to the intended antecedent (which must also satisfy that content).

The notion of the *descriptive content* of a noun phrase is complex and pragmatically subtle. For present purposes, here is a first approximation:

(34) The **descriptive content** of a definite or demonstrative description is the denotation of the nominal complement of *the/this/that/these/those*. The descriptive content of a personal or demonstrative pronoun is the semantic content corresponding to its person, number, and gender, as morphologically marked.

[1] See Roberts (2003), Elbourne (2005), *inter alia*.

(35) A definite noun phrase with descriptive content *np* is **informationally unique** relative to a given context *c* with domain DR just in case there is exactly one $d \in$DR s.t. the context entails that *np* holds of *d*.

The prevalence of so-called *incomplete descriptions*, illustrated in (36), argues that some kind of domain restriction generally plays a role in the satisfaction of (35): [2]

(36) John has a cat and a dog. He walks the dog twice daily and lets the cat out at night.

In any reasonable common ground among ordinary speakers, there will be information about the existence of many cats and dogs. Hence, neither of the underlined definite descriptions in (36) has a descriptive content sufficiently rich by itself to satisfy informational uniqueness, picking out a unique discourse referent in that context which satisfies that content. The obviously felicitous use of these descriptions in this context argues that there is a regular way in which speakers reasonably assume that their addressees will, nonetheless, Retrieve the uniquely intended interpretation for these NPs. Here is a hypothesis about how we do that:

(37) **The attentional restriction of the domain for Retrievability**: Even though a given definite NP is not informationally unique relative to the entire Domain of familiar Discourse Referents in the DR of the context of utterance, if a speaker can reasonably assume in advance that the addressee's attention is restricted to some subset of DR, a subset in which the intended antecedent discourse referent *is* informationally unique relative to the NP's descriptive content, then this guarantees that that antecedent is readily Retrievable.[3]

There is, it seems, a trade-off between the richness of a definite NP's descriptive content and the required degree of salience of the intended antecedent in order for the NP's use to be felicitous—i.e., for the intended antecedent to be readily informationally unique in the attentionally restricted domain, and hence Retrievable. To briefly illustrate, consider (38):

(38) I was shopping in Kroger's this afternoon when a guy asked me for help choosing vegetables. He had his little boy with him, and the child was fussy.

a) He	wasn't sure which greens were kale.
b) The man	wasn't sure which greens were kale.
c) The man in Kroger's	wasn't sure which greens were kale.
d) The man I met today	wasn't sure which greens were kale.

[2] For example, one could take (37) to suggest how one retrieves the implicit domain variable in a definite, along the lines suggested in von Fintel (1994), or the descriptive content of a definite whose descriptive content has been deleted, as in the treatment of E-type pronouns in Elbourne (2003). That would not be my preference, but for reasons orthogonal to the present discussion.

[3] This presumes that the ideal case obtains, wherein the context is non-defective; i.e. all the interlocutors agree on the Score and have equal access to its content. Of course, this is not always the case. But this is how meaning conveyance works when it works, and cases involving defective contexts are to be explained in terms of the respects in which they depart from the ideal.

In the context given, in order to convey anaphoric reference to the man mentioned in the first utterance, a speaker would be most likely to use either (a) or (b). It may be that (38a) is sub-optimal from a purely processing point of view, since when *he* is first encountered the addressee might take its antecedent to be the discourse referent for either the man or the boy. But in this context, plausibility makes the man be the more likely antecedent.[4] (c) or (d) would be odd, because their descriptive content is richer than required to pick out the uniquely most salient man. This oddness might reflect an implicature based on Grice's Quantity 2 maxim: Don't say more than is required for the purposes of the conversation. Thereby we are enjoined to use the leanest descriptive content that will guarantee Retrievability. But if we want to refer to someone other than the salient man or his son, even if the intended referent is familiar from prior discussion, we would be forced to use richer descriptive content:

(38) e) <u>The man I met (last week) in Weiland's</u> had a similar problem with rutabagas.

With contrastive accent on the subject of (38e), the speaker clearly intends to refer to someone other than the man at Kroger's—perhaps someone she and her addressee had discussed on a previous occasion. This impression is reinforced by the predicate, which doesn't plausibly hold of the Kroger's guy (first introduced in the first sentence), because of the anaphoric *similar*. With the Weiland's guy as intended antecedent, use of (38b) would of course be infelicitous, the antecedent nonRetrievable as a function of the descriptive content and the relative salience of familiar discourse referents.

How is the attentional restriction of the domain affected? Recall the definition of Relevance in the previous section, characterized in terms of the propositional content of an utterance. We can extend that notion to talk about the Relevance of (discourse) entities to the QUD:

(39) **The set of Relevant discourse referents**: In a discourse with scoreboard *S*, discourse referent *d* (in DR) is Relevant to the QUD *q* just in case for some property *P*, the question of whether *d* has *P* is evidently Relevant to *q*.

Then:

(40) **Salience** is a partial order of the elements of DR (the set of Discourse Referents), determined by the degree to which those entities would be immediately in the attentional field of anyone cooperatively paying attention to that context.

(41) Factors in a **salience ranking** in discourse include the following, themselves ranked in descending order of importance:
1. High perceptual salience in the situation of utterance.
2. RELEVANCE to the evident current purposes of the interlocutors, especially the QUD (cf. Grosz & Sider 1986)

[4] To make the pronoun even more felicitous, change *a guy* in the first sentence to *this guy*. It has been noted (Prince 1981) that specific indefinite *this*, among other things signals the speaker's intention to make its referent be the Topic in subsequent discourse. As in (2) above, a pronominal subject is usually taken to refer to the Topic, if there is one. See Roberts (to appear) for discussion of what it is to be a Topic.

3. Coherence, reflected in felicitous rhetorical relations in a relevant strategy of inquiry, with consequent relations between thematic roles in the two utterances (Kehler 2009)
4. Relative recency (Terken & Hirschberg 1994)

With respect to the first of these elements, high perceptual salience, this is a distracter, potentially trumping RELEVANCE to the discourse goals of the interlocutors. I have in mind cases where someone (not necessarily a speaker) points to draw another's attention to some entity or situation (*This is PJ Harvey's new album*); cases like Heim's (1982) donkey walking into a classroom room (*It stinks!*); or a situation where there is an explosion a few blocks away as one sits in a café in Costa Rica in 1982 (*That was a bomb!*). The first, deictic case arguably involves linguistic meaning$_{nn}$: Deixis is coordinated by a speaker with the canonical use of a demonstrative in order to heighten perceptual salience. But such heightening also arises non-linguistically, especially if in response to an unusual or immediately threatening situation. We can characterize the events in the other cases as distracting from the interlocutors' Discourse Goals because they bear on over-arching Domain Goals—the attentiveness brought on by oddity in a particular kind of situation, alertness associated with survival, reproduction, etc.—which goals are typically ranked higher than the goal of participating cooperatively in a discourse by addressing the QUD. That these cases are ranked more highly than Relevance to the QUD, then, is expected under the intentional characterization of context suggested by (28).

When a discourse referent is Relevant under (39), then it is more highly ranked than any other familiar discourse referents (in DR). Hence, in (38), the Relevant man in Kroger's is salient, whereas the man from Weiland's is not.

With respect to the relative ranking of the third and fourth factors in (41), there is empirical evidence that mere recency of mention is not highly ranked where salience is concerned. Terken & Hirschberg (1994) provide clear evidence that parallelism is more important in determining likely antecedence than recency; and Smyth (1994) offers experimental evidence for the preference for antecedents with the same grammatical role as the pronominal. But, in turn, one might speculate that Relevance is an important factor promoting parallelism between antecedents and anaphoric elements, for two reasons.

First, there is cross-linguistic evidence that surface order is constrained by discourse factors, and Centering Theory has long assumed that there is a preference across utterances for entities under continued discussion to persist in the same grammatical and/or thematic roles, especially in subject position. Also cross-linguistically, Topicality (Roberts to appear) pertains to entities which are in some sense under discussion or (as with English specific indefinite *this*) about to be under discussion.[5] Hence, Topics are Relevant and usually definite. Moreover, Topical NPs are typically ordered before less Topical ones referred to in the same utterance—though this is not absolute across all utterances even in a relatively topic-oriented language. Given the persistence of surface order and the way that it reflects Topicality, it seems reasonable to speculate that the role of parallelism observed by Terken & Hirschberg (1994) may partly reflect Topic

[5] See also Büring (2003) on Contrastive Topics, though as Roberts (to appear) notes (and I think Büring would concur), not all Topics are Contrastive Topics.

continuation. Topicality may also play a role in explaining an apparent bias toward taking subjects as antecedents (Stevenson et al. 1994, Arnold 2001).

Perhaps even more interesting, Kehler (2009) calls into question the utility of grammatical role parallelism *per se* in predicting anaphora resolution. Kehler (2002) pointed out some confounds in Smyth's materials, raising doubts about his argument for parallelism as an independent factor in anaphora resolution. Kertz, Kehler & Elman (2006) and Kehler, Kertz, Rohde & Elman (2008) did experiments to control for these confounds, and their results argue that coherence, as reflected in felicitous rhetorical relations, is more successful than grammatical role parallelism in predicting the preferred resolution. E.g., if the understood rhetorical relation between the two utterances containing target and potential antecedent was Result, as in (42b,d), the intended resolution of the anaphora was statistically far more likely to involve a non-parallel relation between antecedent and anaphor (95% for subjects, 94% for objects) than if the rhetorical relation was Parallel, as in (42a,c), where, on the preferred resolution, antecedent and anaphoric trigger were more parallel 98% of the time for subjects, 90% for objects).

(42) Samuel threatened Justin with a knife, and
a. …Erin blindfolded him (with a scarf) [Parallel]
b. …Erin stopped him (with pepper spray) [Result]
c. …he blindfolded Erin (with a scarf) [Parallel]
d. …he alerted security (with a shout) [Result] (Kertz et al. 2006)

Kehler (2009) captures the relevant generalization about how contextually-conditioned expectations play a role in anaphora resolution:

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

So he concludes there really isn't a "grammatical role parallelism bias"; instead, that's an epiphenomenon of a certain kind of data.

Experimental results in subsequent work with Rohde (Rohde et al. 2006, Rohde et al. 2007; Rohde & Kehler 2008a, Rohde & Kehler 2008b) support the thesis that coherence relation is the central factor in predicting anaphora resolution. But as Kehler (2009) points out, in many of their experimental materials Rohde and her associates used different types of questions to bias to different coherence relations—e.g., *What happened next?* to bias to the relation Occasion, or *Why?* to bias to Explanation. Kehler (2009) then agrees with Roberts (2004) that we can understand different coherence relations as reflecting different strategies of inquiry in a QUD-based discourse structure. Hence, he argues that the relation of the target utterance to the QUD, reflecting the speaker's adopted strategy, is the central factor in predicting anaphora resolution. And he shows that this constraint is not specific to pronoun resolution, but can be seen in full NPs, as well, as reflected in prosody.

From all this, we might conclude that it is not parallelism, but coherence in a strategy of inquiry, which should be the third factor in (41). One might even subsume the observations about Topicality under the third factor, noting that Topicality is often a reflection of an Elaboration sequence. This approach deserves more careful investigation.

Finally, entities which are neither perceptually salient, obviously Relevant nor recently mentioned are not at all salient.

For the time being, take the factors in (41) to be ordered as given. Now we can characterize the way in which anaphora resolution proceeds as a function of salience, Relevance and plausibility, as follows:

(43) **Attentional Masking Hypothesis**: The search for an anaphoric antecedent among the accessible discourse referents proceeds as follows:[6] Look first to the most salient entities, then to all those that are less salient but still Relevant, and finally to all elements of DR, the domain reflecting all familiar entities in the Common Ground. The antecedent is the first discourse referent you encounter which is informationally unique among the discourse referents ranked at its level of salience in satisfying the NP's descriptive content (while being plausible in view of what is predicated of the NP).

(44) **Descriptive content condition**: To guarantee Retrievability in using a definite NP, a speaker should choose one whose descriptive content is just sufficiently rich to uniquely identify the intended discourse referent among all those which are at least as salient.

Hence, (43) and (44) entail that from the addressee's point of view, the **alternative possible antecedents for a definite NP** are those discourse referents which are at least as salient as the most salient discourse referent(s) satisfying the NP's descriptive content. In (38), assuming that the most salient discourse referent is that for the man in Kroger's, *he* in (38a) will correctly lead to Retrieving that discourse referent as antecedent. If we take the son to be (nearly) as Relevant as the father, then *the man* in (38b) would be preferable. But the descriptive content of the subjects in (38c) and (38d) is richer than necessary. Since the intended antecedent of the subject of (38e) is not salient, the richer descriptive content is motivated and successful.

Examples like (1), repeated here:

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

and the examples involving deixis in (3) and (3′) argue that what is important for salience is not just that something be in the immediate visual field of the addressee, perhaps as directed by

[6] Accessibility is a constraint on antecedence that depends upon the scopes of logical operators in discourse. See Heim (1982), Kamp (1981), and Chierchia & Rooth (1985). Following Roberts (2003), I would include in the potential set of antecedents for a definite all discourse referents which are accessible and *weakly familiar* (in her

deixis, but that s/he be attending to it, hence that it be Relevant to her immediate goals and associated intentions. In (1), there is nothing intrusive or unusual about the mug and other visually accessible entities, so the interlocutors' attention is arguably focused on the QUD, here about the definition. I would argue that this is the same kind of attentional salience as what licenses the use and interpretation of the pronoun in examples like (2):

(2) A: What's up with John$_i$?—I saw him talking with Mac$_k$ earlier.
B: He$_{i/\#k}$ found a dent in his fender.

Addressing A's question requires attending to John, with the follow-up assertion by A about Mac serving only as a sidenote—presumably only Relevant to suggest a motivation for the question (and hence perhaps clarify the kind of explanation of John's behavior being sought). Hence both A and B, on the assumption that they are collaboratively attending to the question, can assume that John is more salient than Mac: The QUD effectively restricts the interlocutors' attentional field and ranks the entities in it.

So long as the descriptive content of a definite NP, along with what is predicated of it, is sufficiently rich to uniquely determine one element in the interlocutors' QUD-limited attentional field, in accordance with Attentional Masking (43) and the Descriptive Content Condition (44), there is no sense that the NP's descriptive content is incomplete. So, short definite descriptions like *the dog* make perfect sense when there is no more than one dog under discussion, as in (36).

(36) John has a cat and a dog. He walks <u>the dog</u> twice daily and lets <u>the cat</u> out at night.

But, of course, the attentional field can change quite quickly in discourse, as in Lewis' (45), where *the cat* in the last clause is readily understood to have a different intended referent than *the cat* in the first sentence:

(45) The cat is in the carton. The cat will never meet our other cat, because our other cat lives in New Zealand. Our New Zealand cat lives with the Cresswells. And there he'll stay, because Miriam [Cresswell] would be sad if the cat went away. [David Lewis 1979]

and also, of course, in the examples that argue for dynamic update of the context in the course of interpretation of a single utterance, like the classic donkey sentences and the bridging versions in (8) and (21).

II. Experimental evidence for the role of the QUD in interpretation

<u>The QUD in the interpretation of ellipsis</u>:

- Frazier & Clifton (in progress): experimental evidence for a generalization of their Main Assertion Principle (Frazier & Clifton 2005, Clifton & Frazier 2010): "Antecedents [for ellipsis] which are part of the main assertion are preferred, especially across sentence boundaries". Their new principle takes the QUD to be central in the interpretation of an utterance:

General interpretation principle: The comprehension system favors interpretations of an utterance which permit the utterance to comment on the QUD.

The Main Assertion Principle follows from this, so that ellipsis resolution can be seen to depend on the QUD. Consonant with the proposals about ellipsis and the QUD of Kehler (2009) and Roberts (in preparation).

The QUD in scope disambiguation and in the calculation of scalar implicatures:

- Hulsey et al. (2004); Gualmini et al. (2008); Zondervan (2008); Zondervan et al. (2008); Gualmini & Schwarz (2009): use a truth value judgment task to investigate two *prima facie* unrelated phenomena, scope disambiguation and the calculation of scalar implicatures. They take their results to strongly support the hypothesis that the QUD plays a central role in each, in both children and adults:

  > [C]hildren are sensitive to the context when they are interpreting a sentence containing a scope ambiguity. The contextual property they take into account is the question that was raised in the context, usually referred to as the *Question Under Discussion* (QUD).….[Accordingly, they propose the Question Answer Requirement:] The Question Answer Requirement: The selected interpretation of an ambiguous sentence, whether true or false, is required to be a good answer to the Question Under Discussion. (A good answer is an interpretation that at least *entails* an answer to the QUD.) (Hulsey, Hacquard, Fox & Gualmini 2004)

The QUD constraining presupposition projection (as proposed by Simons et al. 2010).

- Amaral, Cummins & Katsos (2011), Smith & Hall (2011), and Xue & Onea (2011): provide experimental support for the QUD-based account of presupposition projection, as in examples (9) – (14) above.

The QUD, Focus and the role of prosody in processing

Speer & Blodgett (2006): recent critical overview of work on prosody in processing. Many studies support the contention that both phrasing and prominence play a role in syntactic disambiguation and reference resolution.

Prosodic phrasing in interpretation:

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

Prosodic prominence in interpretation:

- Bock and Mazzella 1983; Birch and Clifton 1995; Cutler 1976; Cutler and Foss 1977; Terken and Nooteboom 1987; Davidson 2001; Ito 2002: "have shown robust effects of intonation [prominence] on discourse comprehension in adults with tasks, such as phoneme detection,

discourse verification, and speeded utterance acceptability judgments." (Speer & Blodgett 2006)

- Roberts (1996,1998, in preparation): the prosodically cued focal structure of an utterance must be congruent with the contextually given QUD in the context of utterance.
- Bock & Mazzella 1983, Terken & Nooteboom 1987, many other psycholinguistic studies ignore the literature in pragmatics and semantics on Focus, instead inquiring into the role of prosodic prominence in marking "new" vs. "given" information or referential NPs. But there are at least three different notions of "givenness" in discourse. The notion which is arguably Relevant in Focus: What is thematic with respect to the QUD (thematically-given—part of the question but not the answer), vs. what is rhematic (rhematically-new—roughly, that part of the utterance intended to be the answer)? Moreover, arguably pragmatic Focus as reflected in prosodic prominence does constrain phrasing, in that under most theories of prosodic constituency and their relationship to Focus (e.g. Selkirk 1996), there can be no more than one pragmatic Focus per intermediate intonational phrase. Hence, the complement of prosodically reflected Focus, prosodic backgrounding is also phrasally constrained. Again, taking the Focus to be the rheme, and backgrounded content to serve as the theme with respect to the understood QUD, both prominence and phrasing play a role in reflecting the QUD addressed by an utterance (Roberts 1996,1998,2010b; Féry & Samek-Lodovici 2006; Beaver & Clark 2008).

Hypotheses about prosody in interpretation:

- Prosody plays a role in processing at the very outset.
- Prosody, including phrasing and, especially, prominence (or prosodic Focus), play a role in guiding attention in processing and interpretation.
- Attention-recognition guides intention recognition.
- Prosodic Focus is correlated with alternative sets, including the QUD (via the theme/rheme distinction) (Rooth 1985,1992).

From these we can conclude:

- One of the central roles of prosody in those languages in which it is used to mark Focus is to help track what is Relevant to the QUD on the scoreboard at the moment of utterance (Roberts 1996,1998,2010b).

Can the range of results reported in the study of prosody in the experimental psycholinguistic literature be illuminated and integrated under this hypothesis? Does the early role of prosody in production reflect something even more general than attentional masking? —the general orientation of processing towards what is Relevant to the QUD.

Corpus studies

<u>Studies of anaphora resolution</u>:

- Poesio & DiEugenio (2001): overview of work on anaphora resolution in the framework of Grosz & Sidner (1986), which is related to the framework presented in section 2. Mixed results at best.
- Tetreault & Allen (2004) concluded that some semantic information (about events and situation types, object types, and other content that could be automatically retrieved) significantly improved pronoun resolution algorithm; but Tetreault (2005) looked "at naive

versions of Grosz and Sidner's theory and Kameyama's intrasentential centering theories" and concluded that "Our results show that incorporating basic clausal structure into a leading pronoun resolution does not improve performance."

Difficulties:

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

III. Pragmatic principles in acquisition: the QUD and the LAD

Meaning Retrieval is a question of intention recognition.

Accordingly: Discourse is a collaborative task, a game with a constrained, mutually recognized structure, based on sets of mutually evident intentions.

. As such, it bears careful consideration as we attempt to understand not only how language is processed, but how it is acquired, indeed its very nature from a cognitive point of view.

Chomsky:

- linguistics is a branch of psychology.
- the language faculty in the human mind is modular, with the relatively autonomous sub-components interacting in such a way as to optimize both efficiency and effectiveness in linguistic processing.
- The development of this modular faculty is triggered and guided by a genetically given Language Acquisition Device (LAD), which leads very young infants to recognize and strive to reproduce the prosody, the phonemic units, and the phonotactics of their native language, this ability evolving in similar ways across languages and cultures.[7]

Assuming Chomsky is right about these general claims:

- From what we know about the human mind and brain, we would expect that both in their evolution and in their interaction on-line the different components of the language faculty (and its associated grammar) are each constrained by the others, and that the interfaces between these components also bear on the character and operation of the components individually.
- Evidence from the role of contextual information and practical reasoning in on-line processing and interpretation, of the sort reviewed above, argues that the language faculty interacts on-line with other, non-linguistic cognitive capacities.
- Then the human linguistic capacity did not develop in a cognitive/neurological vacuum: Almost certainly, given our understanding of the nature of the mind and of the evolution of a nervous system (of which the human mind is a by-product), the components of the linguistic faculty evolved together with these other, related aspects of our cognitive capacity, each

[7] This view would not necessarily commit one to any particular theory of either the LAD or the grammar that results. E.g., this is consistent with an exemplar-based theory.

functionally expressing a particular set of inter-connected roles in the optimal design of a human signaling system.

Chomsky (1979):

- The conditions and substrate for such a development could only have arisen in a mind/brain capable of exploring higher mathematics.

Likely true, but not sufficient:

- The language faculty is specifically designed to process an utterance in the context of a mind that is capable of recognizing the subtle semantic intentions of an agent, his meaning$_{nn}$ (Grice 1957).
- The interaction, between conventional content, as Retrieved by language-specific mechanisms, and some representation of the evident intentions of the speaker, is engineered to be mediated by a particular kind of organized body of information, of the sort modeled in (28). I.e., this kind of information which must be available to the interface between linguistic processing, on the one hand, and cognitive processes like practical reasoning (Wallace 2008), and information storage and retrieval, on the other.

<u>Experimental psycholinguistic evidence on the role of attention and intention-recognition in acquisition</u>**:**

- Bloom (2000): overview of experimental work arguing that intention recognition not only plays an on-going role in utterance processing, but is crucial in grasping meaning at the outset, in language acquisition.
- Baron-Cohen 1988,2009; Boucher 2003, Rapin & Dunn 2003, de Villiers, Stainton & Szatmari 2007: The problems with language acquisition displayed by autistic children arguably result in part from their difficulty in developing a theory of others' minds.
- Baron-Cohen (1991): identifies the infant's understanding of attention in others, usually well-developed by 9 months of age, as a "critical precursor" to the development of a theory of mind, leading to an understanding that deixis can be used to foster joint attention, and that this is related to directing interest, in turn a precursor to semantic reference.

<u>Correlated work on the acquisition of prosody</u>: Speer & Ito (2009)
experimental work on phrasing:

- Mehler et al. 1988; Jusczyk et al. 1993: newborns as early as 3 days old can discriminate between two spoken languages on the basis of their prosody.
- Jusczyk et al. 1993; Morgan and Saffran 1995; Morgan 1996; Johnson and Jusczyk 2001: 6-month-olds use various aspects of prosody to determine the location of words in the stream of running speech.
- Jusczyk et al. 1995; see also Hirsh-Pasek et al. 1987: By 6-months old, infants show sensitivity to whether speech is prosodically well-formed: Artificial pauses were introduced into recordings of naturally occurring speech, and infants preferred to listen to passages with the pauses inserted at prosodic boundaries over the same passages with pauses inserted in the middle of prosodic phrases.
- Snow (1994): evidence from longitudinal observation of children's spontaneous speech for the hypothesis that acquisition of the control of prosodic phrasing corresponds to the

acquisition of verb argument structures, the two appearing around the same time in development.

experimental work on the acquisition of prosodic prominence:

- Friederici et al. (2007): recordings of event-related potentials in 4-month-old German and French infants showed differences in electrophysiological brain responses to native vs. non-native stress patterns.
- Schmitz et al. (2006): German infants develop sensitivity to the location of the accent in prosodic phrases by 8 months of age.

A potential problem: See discussion in the paper of Hornby and Hass 1970; Hornby 1971; Wieman 1976; Macwhinney and Bates 1978; Culter and Swinney 1987; Wells et al. 2004.

Hypothesis: Prosody is central in both acquisition, and processing both (a) because phrasing plays an important role in guiding parsing, and (b) because of the vital role that prosodic prominence plays in indicating the attentional structure, or Focus of the utterance. Thereby, via the correlation between that attentional structure and the intentional structure of discourse context, through congruence to the QUD, prosody is key to Retrieving the speaker's intended interpretation.

Conclusion: We retrieve implicit elements of the intended meanings of utterances with a view to relevance to the QUD and the interlocutors' evident domain goals and associated intentions. Recognizing and tracking such intentions is, then, central to solving for interpretation.

References and bibliographic note:

See the paper version of this talk for most of the work cited:
http://www.ling.ohio-state.edu/~croberts/Solving_for_interpretation.Oslo.paper.pdf
plus:
Tonhauser, Judith, Mandy Simons, David Beaver & Craige Roberts (2012) Towards a taxonomy of projective content. Accepted with revisions, *Language*.

Please do keep me posted on any work you do, or work by others that bears on the questions we discussed in this class. It is my intention to try to keep up to date the on-line bibliography at: http://www.ling.ohio-state.edu/~croberts/QUDbib/.