

1 Entailment and Givenness

Theories of focus structure define GIVEN words and phrases as those whose surrounding discourse context provides an appropriate antecedent, known as a focus antecedent. Whether a phrase is GIVEN may affect its syntactic, phonological, and semantic behavior. In English, for instance, GIVEN phrases are not pronounced as prominently as non-GIVEN, or NEW phrases. Consider the pattern of acceptable pronunciations for the two questions below, where accented words are in SMALL CAPS:

- | | | | |
|-----|--------------------------------|-----|--------------------------------|
| (1) | Who is Joffrey's father? | (2) | Whose father is Jaime? |
| | a. JAIME is Joffrey's father. | | a. #JAIME is Joffrey's father. |
| | b. #Jamie is JOFFREY'S father. | | b. Jaime is JOFFREY'S father. |

When a question mentions Joffrey, utterances of *Joffrey* in the answer count as GIVEN and hence are not accented; when the question mentions Jaime, though, it is his name that must be unaccented. A name in the question acts as an antecedent for the same name in the answer, rendering this second use GIVEN.

Since at least Lakoff (1971)¹, authors have also argued for non-identical focus antecedents. For instance, it seems an antecedent may be quite different syntactically and phonologically, if it is identical semantically, as shown in (3):

- (3) Willie Mays lost his cool, but... (\approx Lakoff 1971, Ex. (27))
- | | |
|----|--|
| a. | ... the centerfielder of the Giants soon regained it. |
| b. | #... the centerfielder of the GIANTS soon regained it. |

Since (in 1971) Willie Mays played center field for the San Francisco Giants baseball team, the phrase *Willie Mays* can act as the antecedent for the co-referent phrase *the centerfielder of the Giants*. It seems that such superficial differences may be ignored; for the purposes of GIVENNESS, two phrases are identical if they have the same denotation (when evaluated at the same world and time index).

More puzzling than (3), though, are cases that lack even semantic identity. For instance, in (4a), the verb *insulted* is GIVEN (as indicated by its unaccented pronunciation), even though it has not been mentioned before. Lakoff points out that this example crucially relies on the fact that calling someone ugly is an insult; once the original action is changed to a non-insult, as in (4b), the unaccented pronunciation of *insulted* sounds odd. Thus, it seems as though some GIVEN items have antecedents that are not semantically identical.

- (4) (Lakoff 1971, Ex. (20))
- | | |
|----|---|
| a. | John told Mary that she was ugly and then SHE insulted HIM. |
| b. | #John told Mary that she was beautiful and then SHE insulted HIM. |

Rochemont (1986) picks up this thread with the following example:

¹Lakoff credits R. M. W. Dixon and Georgia Green (p.c.) for bringing such examples to his attention, and quotes some related examples from Green (1968).

- (5) (Rochemont 1986, Ex. (40))
 A: I saw some GORILLAS in the SUBWAY today.
 B: Oh really?
 (i) WE saw some gorillas/#GORILLAS in the ZOO today.
 (ii) WE saw some animals/#ANIMALS in the ZOO today.
 (iii) WE saw some #tigers/TIGERS in the ZOO today.

Options (i) - (iii) in (5) represent three different possible statements by speaker B. Rochemont points out that both the word *gorillas*, used in (i), and (surprisingly) the word *animals*, used in (ii), are preferably unaccented after a previous mention of gorillas, while words denoting other animals, such as tigers in (iii), are preferably accented. In our terminology, *animals* is GIVEN in this example, perhaps licensed by the non-identical antecedent *gorillas*.

Rochemont suggests that the relation between a GIVEN phrase and its antecedent is one of “informal entailment” rather than identity. Schwarzschild (1999) formalizes this notion of GIVENNESS as entailment after generalized existential closure.² We return to this formal definition in §5, but in the case of properties like *animals* and *gorillas*, it simply amounts to the subset relation: *gorillas* counts as an antecedent for *animals* because the set of gorillas is a subset of the set of animals, but *gorillas* may not antecede *tigers* since the two relevant sets do not even overlap. Words so related are also known as a hypernyms (supersets) and hyponyms (subsets).

In addition to the cases noted by Rochemont and Lakoff, further apparent examples of non-identical focus antecedents arise in examples of corrections:³

- (6) A: Jane saw some GORILLAS in the SUBWAY today.
 B: Actually, she saw SOME/ANOTHER kind of animal there, but it wasn't gorillas.
- (7) A: Where did you see (some) gorillas today?
 B: I saw SOME/ANOTHER kind of animal in the SUBWAY, but it wasn't gorillas.

In (6), speaker B challenges part of A's assertion. B crucially agrees with A that Jane saw some animals, but disputes that these animals were gorillas. The word reflecting this area of agreement (*animal*) is unaccented, indicating a GIVEN status. Again, this GIVENNESS seems to stem from the earlier utterance of *animal*'s hyponym *gorillas*. Similarly, in (7), B rejects part of the assumption of A's question, namely that B saw some gorillas, while agreeing that B did see some (other kind of) animal. As before, the word *animal* may remain unaccented in this context.

In this squib, I will challenge the shift away from a stricter, semantic identity relation between a GIVEN phrase and its antecedent. To this end, in §2, I present several counterexamples to this move: cases where identity is required

²Actually, Schwarzschild formalizes just the entailment part of “informal entailment”; Rochemont's system also includes contextual information.

³I am grateful to an anonymous reviewer for proposing examples similar to these.

between antecedents and GIVEN phrases. Since the solution presented here will involve the Question Under Discussion (QUD) model of discourse, §3 briefly summarizes this system. Next, §4 argues for a particular connection between the QUD model and GIVENness. There, I argue that cases like (4) and (5) are not actually counterexamples to an identity notion of GIVENness; essentially, these cases involve identical antecedents found in discourse structure rather than spoken aloud. Last, §5 presents the formal details of the proposal, and §6 discusses a few examples below the sentential level.

2 Counterexamples

The first indication that discourse structure plays an important role in determining the GIVENness of words with non-identical antecedents comes from the following two variants of (5) above:

- (8) A: I saw some GORILLAS in the SUBWAY today.
B: I saw some ANTELOPES in the PARK today.
C: I saw some animals in the ZOO today.
- (9) A: I saw some GORILLAS in the SUBWAY today.
B: I saw some SEQUOIAS in the PARK today.
C: #I saw some animals in the ZOO today.
(cf. ✓I saw something cool in the ZOO today.)

The discourse in (8) is simply an extension of Rochemont's example: a sentence mentioning another animal (*antelopes*) has been interpolated between Rochemont's original sentences, and the result is a felicitous discourse. However, the quite similar case in (9) sounds much worse; it is odd for C to leave the discourse-new word *animals* unaccented here, presumably because while gorillas are animals, sequoias are not. (The variant of (9) where C uses the unaccented phrase *something cool* instead of *animals* sounds fine.) The intervening phrase *sequoias* seems to interrupt the antecedence relation between *gorillas* and *animals*. Crucially for the argument being pursued here, a discourse like (9) improves greatly when it instead includes an identical, unaccented phrase:

- (10) A: I saw some GORILLAS in the SUBWAY today.
B: I saw some SEQUOIAS in the PARK today.
C: I saw some gorillas in the ZOO today.

So, here is a counterexample to Rochemont's proposal: a case where *gorillas* is GIVEN, as shown in (10), but *animals* is not, as shown in (9).

Next, in cases where the parallelism inherent in Rochemont's and Lakoff's examples is lacking, identity reemerges as a necessary condition for GIVENness. For instance, consider the following cases:

- (11) A: Nice to meet you. Tell me a little about yourself.
B: Well, GORILLAS are my favorite #KIND of animal.
(cf. ✓GORILLAS are my favorite kind of ANIMAL.)

- (12) A: What did the dog owner buy?
 B: She bought a PICTURE of a(n) dog/#animal.
 (cf. ✓She bought a picture of an ANIMAL.)
- (13) A: I bought these APPLES at the FARMER'S MARKET last week.
 B: Oops! That reminds me: I left a BAG of apples/#fruit/#produce/
 #food in the CAR. (cf. ✓I left a bag of FRUIT in the CAR.)

In sentence (11), the two relevant words are again *gorillas* and *animals*. In B's response, *animal* cannot be unaccented, despite the quite recent mention of its hyponym *gorilla*. This is unexpected: in (5), Rochemont's example above, a mention of the word *gorilla* seemed to be enough to license the GIVENness of the word *animal*. Here, though, the sentence must be pronounced with an accent on *animal*.⁴ Moving to (12), we find the word *dog* embedded deeply in the structure of a question about something else (namely, what someone bought). Subsequent mention of *dog* in B's reply is felicitously unaccented, but mention of its hypernym *animal* is must be accented. Finally, (13) is a case where the topic shifts abruptly. The relevant pair of words in (13) is *apples* and *fruit* (among other choices). Even across such a large shift of topic, the word *apples* in B's reply remains GIVEN (and hence may be unaccented) due to its recent mention. The word *fruit*, on the other hand, must be accented in the same position.

Tellingly, all the cases of unaccented hypernyms improve in slightly different discourse scenarios. (14) shows versions of the dialogues above where B does mention the hypernym earlier, in passing and without drastically altering the discourse context. In each case, the second mention of this hypernym may now be felicitously left unaccented:

- (14) A: Nice to meet you. Tell me a little about your interest in animals.
 B: Well, GORILLAS are my favorite KIND of animal.
- (15) A: What did the dog owner buy?
 B: Well, as an ANIMAL lover, she bought a PICTURE of an animal.
- (16) A: I bought these APPLES at the FARMER'S Market last week.
 B: Oops! Speaking of FRUIT, I left a BAG of fruit in the CAR.

Further examples are shown below:

- (17) (Here are some fun facts. . .)
 a. #The Burj KHALIFA is the TALLEST skyscraper.
 (cf. ✓The Burj KHALIFA is the tallest SKYSCRAPER.)
 b. #The TOMATO is ACTUALLY a fruit.
 (cf. ✓The TOMATO is actually a FRUIT.)

⁴It is also interesting to compare this case to (i):

- (i) Vanilla ice cream is my favorite FLAVOR of ice cream.

Unlike the word *gorilla*, which does not include the word *animal*, the phrase *vanilla ice cream* actually mentions its hypernym *ice cream*. Subsequent mentions of ice cream are therefore allowed to be GIVEN, unlike subsequent mentions of animals after gorillas are mentioned.

- (18) a. A: Where did you take your gorilla?
 B: I took it to a SHELTER for gorillas/#animals.
 b. A: What do you call a joke/#pun you make in the shower?
 B: A CLEAN joke.
- (19) a. I was eating a Boston CREAM pie while I drove THROUGH Boston/
 #Massachusetts/#the United States.
 b. I chose JOHN_i because I thought (he_i/#a man) would understand
 my SITUATION.
 c. The famous HOT DOG eating champion SELLS hot dogs/#food for
 a LIVING.
- (20) A: I saw some GORILLAS in the SUBWAY today.
 B: Hmm. That reminds me. There's a DOCUMENTARY on goril-
 las/#animals on PBS tonight.

Based on cases like these, I will argue in §4 that Rochemont's and Lakoff's examples (among others) are the exceptions rather than the rule. In these exceptional cases, the antecedent to a GIVEN phrase is found in an implicit discourse question instead of being explicitly spoken aloud. The reason that the cases in presented in this section do not fit Rochemont's pattern is that they project a different discourse structure from those in §1. First though, in §3, I will introduce the notion that discourse is structured by questions and answers.

3 Questions Under Discussion

Roberts (2012) (following Carlson 1982) argues that the basic unit of discourse is a question/answer pair, where answerhood is defined by (contextual) entailment:

- (21) A: Did you see any animals today?
 B: (i) (Yes,) I saw some animals today.
 (ii) (No,) I didn't see any animals today.
 (iii) (Yes,) I saw some gorillas today.
 (iv)#(No,) I didn't see any gorillas today.

Statements that directly address the question, like (i) and (ii), of course count as answers, but so do statements that only entail one of these responses. For instance, since seeing gorillas entails seeing animals, (iii) is a felicitous answer to A's question. Not every response where *gorillas* replaces *animals* is felicitous, though, as shown in (iv). This last response does not entail one of the direct answers, so it is not a felicitous response.

In Roberts's system, larger discourses are structured as a series of questions, some spoken aloud and others merely implied by the spoken portions of a discourse. All conversations contribute to answering what she calls the "Big Question," *What is the way things are?* (cf. ?). In practice, most conversations start with a subquestion⁵ of this larger question, as illustrated by Roberts's

⁵Roberts defines a subquestion as one whose complete answer (contextually) entails at

example:

- (22) Who ate what? (= Roberts's ex. 1)
- a. What did Hilary eat?
 - (i) Did Hilary eat bagels?
Yes
 - (ii) Did Hilary eat tofu?
Yes
 - b. What did Robin eat?
 - (i) Did Robin eat bagels?
Yes.
 - (ii) Did Robin eat tofu?
Yes.

Here, the first question under discussion – namely *Who ate what?* – is a subquestion of the Big Question. In this idealized example, each utterance (or “move” as Carlson and Roberts term them) either answers or poses a subquestion to the most recent unanswered QUD. For instance, move (22a) (*What did Hilary eat?*) poses a subquestion to the initial move, which happens to be the previous question asked. Move (22b) poses a subquestion to this initial question, too, even though several other question moves intervene. This is possible because the intervening questions have been answered, leaving *Who ate what?* as the most recent unanswered QUD. Following these strict rules leads to a much more felicitous discourse than the one in (23), where questions, subquestions, and answers are all mixed up:

- (23) Who ate what?
- a. What did Hilary eat?
Yes.
 - (i) Did Robin eat bagels?
...

Roberts points out, however, that certain discourses are felicitous despite bending the rules somewhat. For instance, (24) sounds fine, despite the fact that (24b) is not technically a subquestion of (24a)⁶:

- (24) (Roberts's ex. 13)
- a. What kinds of seafood will John eat?
 - b. Isn't John allergic to clams?

For this reason, among others, she argues that certain discourse moves may remain unspoken/implicit if they are sufficiently recoverable to all conversants. In (24), the implicit move is perhaps a question like *What reasons would John have for not eating clams?*, which is a technically correct subquestion of (24a)

least a partial answer to its superquestion.

⁶As Roberts explains, a complete answer of “No” to (24b) does not even partially answer (24a), going against the definition of questions and subquestions.

and a technically correct superquestion of (24b). In other words, the reason that (24) sounds better than (23) is that there is a recoverable implicit QUD which brings (24) into compliance with the proper rules of discourse.

Roberts (2012, p. 8) argues that, in general, listeners infer implicit QUDs using Plan Inferencing Rules (part of Planning Theory). However, she most explicitly discusses the role of prosody in guiding listeners to recover implicit QUDs. For instance, in an example due to Schwarzschild (1999, ex. 50, p. 165), prosody can disambiguate which (implicit) QUD a correction is answering, and therefore which statement is being corrected:

- (25) A: John borrowed the book that Max had purchased.
B: (i) No, MAX borrowed it.
(ii) No, Max BORROWED it.

In (25Bi), the implicit QUD is paraphraseable as *Who borrowed the book?*. Such a question requires prosodic prominence on its subject (which replaces the *wh*-word *who*). In (25Bii), this implicit QUD is paraphraseable as *What did Max do to the book?*, which requires prominence on the verb instead of the subject. The difference in prosody allows the listener to reconstruct the correct QUD. Therefore, since a correction and the clause being corrected must be able to answer the same QUD, (25Bi) is taken to correct the matrix sentence while (25Bii) is understood to correct the relative clause within it.

Büring (2003) adopts Roberts's notion of unspoken QUDs, but he argues for a limitation on the prosodic cues for such implicit moves: a focus-antecedent for a GIVEN word or phrase, Büring claims, cannot be introduced in an implicit discourse move. (Note that the unaccented words in (25) above all have explicit antecedents.) Büring (2003) discusses more complex, two-level implicit QUD structures, but his point holds in simpler contexts, as well. For instance, consider the potential prosody variants of the following short dialogue:

- (26) A: Oh, hi Bob! How are things going?
B: Not bad, Alice.
(i) My DAUGHTER got into HARVARD.
(ii) #My DAUGHTER got into Harvard.

Although the relationship between focus and prosody in complex, in general English phrases have accents on their rightmost (last-spoken) components. Therefore, in an out-of-the-blue context like (26), the most natural prosodic rendering of B's reply will place an accent on the rightmost word, *Harvard*. However, when the phrase *got into Harvard* is GIVEN, such as after the explicit question *Who got into Harvard?*, the prosodic prominence shifts away from this position, leaving the highest prominence on the subject (as in (26Bii)).

Now, if a listener like Alice were free to infer any implicit QUD that fit a spoken prosody, she could infer the question *Who got into Harvard?* as the implicit QUD that Bob had in mind for his reply. However, this prosody sounds quite odd in this general context, providing evidence for Büring's claim. We can capture this explicitly using an Optimality-style constraint:

- (27) *IMPLICITANTECEDENT
A GIVEN phrase may not have its only antecedent in an implicit QUD.

In the next section, we will see cases where this constraint seems to be violated.

4 Implicit QUDs and Givenness

Roberts (2012) actually discusses an example based on one from Lakoff (1971):

- (28) Mary called Sue a REPUBLICAN, and then SHE insulted HER.
(Roberts's ex. 52)

This example is parallel to the case in (4), but here *contextual* entailment is important: the context must be one where calling someone a Republican is tantamount to insulting them. Roberts's analysis of this case is that the verb *insulted*'s GIVENNESS arises not from the earlier VP of the form *called X a Republican* but rather from an implicit QUD that both clauses of (28) answer: *Who insulted whom?* As shown in (29), this question may be felicitously spoken aloud before an utterance of (28):

- (29) A: Who insulted whom?
B: Mary called Sue a REPUBLICAN, and then SHE insulted HER.

Mary called Sue a Republican is a felicitous answer to the question in (29A), and Roberts's system allows another clause, such as *then she insulted her*, to continue answering this same question. When such a question is spoken aloud, an identity-based constraint on GIVENNESS is sufficient because the antecedent for the GIVEN verb *insulted* in B's statement comes from A saying this verb first. Roberts's explanation of (28) (which lacks an explicit preceding question) can also be understood as maintaining an identity definition of GIVENNESS but (pace Büring 2003) allowing the relevant antecedent to be in an implicit QUD rather than being spoken aloud.

This analysis works for Rochemont's example in (5), too, as suggest by the version of (5) in (30) where a preceding (explicit) question has been added:

- (30) A: Where did people see animals today? I saw some GORILLAS in the
SUBWAY.
B: WE saw some animals in the ZOO.

Again, in this explicit version, the GIVEN noun *animals* finds an identical antecedent in the question. Assuming an implicit QUD identical to this question in Rochemont's original example (5) will again allow us to maintain an identity-based relationship between the GIVEN node and its (unspoken) antecedent.

Finally, such an implicit question could conceivably precede the new cases of apparent non-identical focus antecedents presented above in (6) and (7). These cases are repeated here with an explicit question:

- (31) What kind of animal did Jane see where?

- A: Jane saw some GORILLAS in the SUBWAY today.
 B: Actually, she saw SOME/ANOTHER kind of animal there, but it wasn't gorillas.
- (32) What kind of animal did you see where today?
 A: Where did you see (some) gorillas today?
 B: I saw SOME/ANOTHER kind of animal in the SUBWAY, but it wasn't gorillas.

One major problem remains, though: Roberts's analysis goes against the *IMPLICITANTECEDENT constraint assumed in Büring.⁷ If her analysis is to be maintained, it will be necessary to posit some exception to this constraint. This exception will have to crucially apply to the cases from §1, but not those from §2. The intuition that I will pursue below is as follows: while every two explicit discourse moves must share a common ancestor QUD in Roberts's system, discourse structures where such a common QUD is more specific are preferred over those where the common QUD is less specific. I propose the following constraint, ranked higher than *IMPLICITANTECEDENT, to capture this intuition:

- (33) SPECIFICCOMMONQUD
 Prefer structures where the lowest common ancestor QUD of two explicit discourse moves is more specific.

Since Roberts' system is perfectly hierarchical, with (more specific) sub-questions branching below (less specific) super-question ancestors, any two moves will share an ancestor, even if it is simply the root of the tree, which I take to be the Big Question here. For two assertion moves, this means that there will always be question in the QUD structure that both moves (at least partially) answer. For an assertion move and a question move, this means that there will be an ancestor QUD that the assertion answers and the question is entailed by (i.e., is a sub-question of). If there is more than one such question, there will always be a lowest one. The SPECIFICCOMMONQUD constraint privileges structures where such lowest common QUDs are more specific (defined via question-entailment) when compared to competitor QUD structures. Qualitatively, to the extent that QUDs can be thought of as conversational topics, this constraint will prefer structures where the implicit topic of a discourse is narrowly construed, based closely on the evidence provided by explicit discourse moves.

To illustrate this new constraint, consider four potential QUD structures for (28) above, with the lowest common ancestor QUD rendered in bold. Each line represents a discourse move, with implicit moves in angle brackets:

- (34) **⟨What is the way things are?⟩**
 John called Mary a Republican.
 She insulted him.

⁷Although officially published after Büring (2003), Roberts (2012) was written earlier.

- (35) ⟨What is the way things are?⟩
 ⟨**Who did what to whom?**⟩
 John called Mary a Republican.
 She insulted him.
- (36) ⟨What is the way things are?⟩
 ⟨**Who insulted whom?**⟩
 John called Mary a Republican.
 She insulted him.
- (37) ⟨What is the way things are?⟩
 ⟨**Who insulted whom?**⟩
 ⟨Did John call Mary a Republican?⟩
 John called Mary a Republican.
 ⟨Did Mary insult John?⟩
 She insulted him.

Among other things, the SPECIFICCOMMONQUD constraint prevents structures that simply take every assertion to answer the Big Question, as in (34). The structure in (35) is preferable to (34) because it features a more specific lowest common QUD for the two explicit discourse moves. But (36) is even better, because its lowest common QUD is even more specific. Notice that, unlike (34) and (35), (36) violates *IMPLICITANTECEDENT. This is allowed, though, since SPECIFICCOMMONQUD outranks *IMPLICITANTECEDENT. Finally, additional specific QUDs targeting just one discourse move, as in (37), do nothing to improve violations of SPECIFICCOMMONQUD, because the lowest common QUD is the same as in (36). Since (37) actually violates *IMPLICITANTECEDENT more than (36) does, (36) is the overall more preferred structure.

To summarize, then, the examples of non-identical antecedents seen so far (collected in (38)), all involve cases where *IMPLICITANTECEDENT is violated in order to reduce violations of SPECIFICCOMMONQUD:

- (38) a. John told Mary that she was ugly and then SHE insulted HIM.
 b. Mary called Sue a Republican, and then SHE insulted HER.
 c. A: I saw some GORILLAS in the SUBWAY.
 B: We saw some animals in the ZOO.
 d. A: Jane saw some GORILLAS in the SUBWAY today.
 B: Actually, she saw SOME/ANOTHER kind of animal there.
 e. A: Where did you see (some) gorillas today?
 B: I saw SOME/ANOTHER kind of animal today, in the SUBWAY.

The most specific common QUD for (38a-b) is (paraphraseable as) *Who insulted whom?*, which violates *IMPLICITANTECEDENT, but provides a nice specific common QUD. Similarly, listeners to (38c) infer a common QUD akin to *Where did who see some animals?*, which violates *IMPLICITANTECEDENT in order to be as specific as possible while still accepting both statements as answers. The most specific question encompassing both A's assertion and B's correction

thereof in (38d) is *What kind of animal did Jane see where?*, which provides an antecedent to B's use of *animal*. Finally, B's reply in (38e) doesn't directly answer A's question, revising it to the most specific ancestor to both A's and B's moves, something like *What kind of animal did you see where today?*

Unlike these cases, though, the examples in §2 are not amenable to structures satisfying SPECIFICCOMMONQUD in order to violate *IMPLICITANTECEDENT. Take (9), for instance, repeated here:

- (39) A: I saw some GORILLAS in the SUBWAY today.
 B: I saw some SEQUOIAS in the PARK today.
 C: #I saw some animals in the ZOO today.

The most specific common QUD between B's and C's statements will not involve animals and therefore will not provide an antecedent for the word *animals*. A's earlier statement does not change this situation: since QUD structure is hierarchical, the common ancestor for A's and C's statements must also accept B's statement as an answer.

Some other cases from §2 are repeated here:

- (40) a. #GORILLAS are my favorite KIND of animal.
 b. I was eating a Boston CREAM pie while I drove THROUGH Boston/
 #Massachusetts.
 c. A: What did the dog owner buy?
 B: #She bought a PICTURE of an animal.
 d. A: I bought these APPLES at the FARMER'S MARKET last week.
 B: Oops! #I left a BAG of fruit in the CAR.

In Roberts's system, discourse moves must ask or answer a question, and thus the smallest move in her system is a full clause. Therefore the single-clause sentence (40a) is a single discourse move in Roberts' system, which cannot trigger SPECIFICCOMMONQUD. (Although see §6 for discussion of subsentential cases.) The case in (40b) does involve two clauses, but they form part of a larger sentence (and one that is not simply a conjunction). B's reply in (40c) directly answers A's question, and therefore A's explicit question counts as the most specific common QUD for the two explicit moves. Finally, the two relevant sentences in (40d) are simply too different from one another: their most likely common QUD probably has nothing to do with fruit.

5 Technical Details

The technical details of the proposal given here are a little more complex than so far presented, because we must define identity explicitly. This is potentially difficult, because Roberts' implicit QUDs are not full syntactically realized natural language objects. Instead, they are merely semantic objects, sets of propositions (the same as the denotations of explicit questions). We have already seen that GIVENNESS is based on semantic – not syntactic – identity, though, and there-

fore this lack of syntactic realization should not present a large problem. In the remainder of this section, I will propose a formal definition of GIVENNESS suitable for antecedents in spoken and implicit discourse moves.

One prerequisite to the definition of identity will be a notion of generalized existential closure (cf. Schwarzschild 1999, Ex. (26)). As defined in (41), the function ExClo turns (almost) any denotation into a proposition, via various types of existential closure:

- (41) Generalized Existential Closure: ExClo
 For any denotation α ,
- a. if α is type t , $\text{ExClo}(\alpha) = \alpha$.
 - b. if α is type e , $\text{ExClo}(\alpha) = \exists P_{et}[P(\alpha)]$.
 - c. if α is type $\langle a, b \rangle$, $\text{ExClo}(\alpha) = \exists x_a [\text{ExClo}(\alpha(x))]$.

To see how this works, consider the closures of nodes from (4) and its QUD:

- (42) a. $\text{ExClo}([\mathbf{Who\ insulted\ whom?}]) = \exists p_{st}[[\mathbf{Who\ insulted\ whom}]](p)$
 $\approx [\mathbf{Someone\ insulted\ someone}]$
 b. $\text{ExClo}([\mathbf{John}]) = \exists P_{et}[P([\mathbf{John}])]$
 $\approx [\mathbf{Something\ happened\ involving\ John.}]$
 c. $\text{ExClo}([\mathbf{insulted}]) = \exists x_e \exists y_e [[\mathbf{insulted}]](x)(y)$
 $\approx [\mathbf{Someone\ insulted\ someone.}]$
 d. $\text{ExClo}([\mathbf{told\ Mary\ she\ was\ ugly}]) = \exists x_e [[\mathbf{told\ \dots\ ugly}]](x)$
 $\approx [\mathbf{Someone\ told\ Mary\ she\ was\ ugly.}]$

In (42a), a question is converted into the proposition that one of its answers is true. In (42b), an individual is converted into the proposition that holds if the individual has at least one property – e.g., if the individual exists. In (42c), the transitive verb *insulted* becomes the proposition that someone insulted someone. Finally, (42d) shows that the closure of the VP *told Mary she was ugly* is the proposition that someone told Mary she was ugly.

With ExClo in place, we can define the our identity condition for GIVENNESS:

- (43) GIVENNESS
 A phrase denoting α is GIVEN iff its antecedent denotes a β such that
- a. $\text{ExClo}(\beta) \rightarrow \text{ExClo}(\alpha)$, and
 - b. there is no denotation γ of the same type as α such that:
 - (i) $\text{ExClo}(\beta) \rightarrow \text{ExClo}(\gamma)$ and
 - (ii) $\text{ExClo}(\gamma) \rightarrow \text{ExClo}(\alpha)$ but
 - (iii) $\text{ExClo}(\alpha) \not\rightarrow \text{ExClo}(\gamma)$.

Clause (a) of (43) requires that (the closure of) a GIVEN node be entailed by (the closure of) its antecedent⁸. Clause (b) disqualifies a node from being GIVEN if

⁸I'm assuming that a GIVEN node's antecedent will be determined by discourse structure. Although the precise determination of every node's antecedent is beyond the scope of this squib, at a first approximation, we can consider a node's antecedent to be either the current QUD or the denotation of a relevant recent sentence. (Non-sentences can take the denotations

a strictly stronger alternative is also entailed by the antecedent. This second clause ensures that only the “most identical” nodes count as GIVEN relative to a particular antecedent.

This stricter definition of GIVENness also correctly handles cases like (4) and (5), once implicit QUDs are assumed for these cases (shown in ⟨angle brackets⟩):

- (44) ⟨Who insulted whom?⟩
John told Mary that she was ugly and then SHE insulted HIM.
- (45) ⟨Where did people see animals today?⟩
A: I saw some GORILLAS in the SUBWAY today.
B: Oh really? WE saw some animals in the ZOO today.

As shown in (42), $\text{ExClo}(\llbracket\text{Who insulted whom?}\rrbracket)$ entails $\text{ExClo}(\llbracket\text{insults}\rrbracket)$ and there is no stronger transitive verb also entailed by the closure of this QUD (because the two mutually entail one another). Similarly, $\text{ExClo}(\llbracket\text{Where did people see animals today?}\rrbracket)$ entails $\text{ExClo}(\llbracket\text{animals}\rrbracket)$ but does not entail any one specific animal (taking people to be separate from animals). Therefore, this implicit question provides a suitable antecedent for the word *animals*.

One wrinkle to this definition comes from words triggering uniqueness pre-suppositions.⁹ Consider a slight variant of the sentences above:

- (46) A GORILLA appeared AHEAD, but I couldn’t catch UP to the animal.

In this (impoverished) context, the denotation of *the animal* is clearly taken to be the gorilla mentioned in the first clause, since no other (non-human) animals have been mentioned. Although the details might be tricky, it makes sense then that the full DP *the animal* is GIVEN, since its denotation is actually the previously mentioned gorilla. However, the single word *animal* itself seems to be GIVEN here as well, since a focused word whose immediate parent node is GIVEN must carry an accent (Schwarzschild 1999). This is unexpected under the definition in (43), since while $\text{ExClo}(\llbracket\text{A gorilla appeared ahead}\rrbracket)$ entails $\text{ExClo}(\llbracket\text{animal}\rrbracket)$, there are many more specific denotations whose closures this clause entails: $\text{ExClo}(\llbracket\text{primate}\rrbracket)$, $\text{ExClo}(\llbracket\text{gorilla}\rrbracket)$, $\text{ExClo}(\llbracket\text{gorilla who appeared ahead}\rrbracket)$, etc. (And in fact these phrases all may appear unaccented in this position.)

Variations of certain cases from §2 exhibit the same pattern when a definite description refers to the same individual as some earlier DP:

- (47) a. A: What did the owner of that dog buy?
B: A PICTURE of the animal/dog (she owns).
- b. A: There some nice APPLES over there!
B: I’ll take a BAG of that fruit/those apples (over there).

And further cases allow reference to a hypernym kind term:

of sentences as their antecedents.)

⁹My gratitude again to an anonymous reviewer for the comment that led to this discovery.

- (48) a. The HOT DOG eating champion SELLS that food for a LIVING.
 b. A: I saw some GORILLAS today.
 B: There was just a DOCUMENTARY about that animal on PBS.

Before I present my analysis, recall that definite descriptions carry a uniqueness presupposition, essential that the description portion of the definite is unique in the relevant, salient context. So, in order for a definite such as *the animal* to be felicitous, there must be only one salient animal in the relevant context of discourse. If this animal happens to be a gorilla, then there will only be one gorilla in the context; if it is a gorilla who appeared in the distance, there will be only one gorilla who appeared in the distance in the context. With this in mind, let g be the gorilla mentioned in (46). In a context where (46) – which includes the definite description *the animal* – is felicitously uttered, the following equality holds (taking the contextually circumscribed denotations of the properties):

$$(49) \quad \llbracket \mathbf{animal} \rrbracket = \llbracket \mathbf{gorilla} \rrbracket = \llbracket \mathbf{gorilla\ who\ appeared\ ahead} \rrbracket = [\lambda x.x = g]$$

In such a context, the following equality holds as well:

$$(50) \quad \text{ExClo}(\llbracket \mathbf{animal} \rrbracket) = \text{ExClo}(\llbracket \mathbf{gorilla} \rrbracket) = \text{ExClo}(\llbracket \mathbf{gorilla\ who\ appeared\ ahead} \rrbracket) = \exists x[x = g]$$

So, once the presupposition of the definite description is taken into account, the denotation $\llbracket \mathbf{animal} \rrbracket$ actually “entails” $\llbracket \mathbf{gorilla} \rrbracket$ and $\llbracket \mathbf{gorilla\ who\ appeared\ ahead} \rrbracket$, ruling out these denotations as alternatives under clause (biii) of (43). In short, these ostensibly stronger alternatives are actually synonymous in this context, and therefore they do not prevent *animal* from being GIVEN.

This odd notion of entailment is not unprecedented. First, Roberts’s (not to mention Rochemont’s) system already employs the similar notion of contextual entailment. Second, von Stechow (1999) uses a similar relation (which he terms “Strawson entailment”) to account for the appearance of NPIs in the scope of the word *only*. We can adopt a similar strategy as von Stechow, reflected in the following revised definition of GIVENNESS:

- (51) GIVENNESS
 A phrase denoting α in a sentence S is GIVEN iff its antecedent denotes a β such that, when $\llbracket \mathbf{S} \rrbracket$ is defined,
- a. $\text{ExClo}(\beta) \rightarrow \text{ExClo}(\alpha)$, and
 - b. there is no denotation γ of the same type as α such that:
 - (i) $\text{ExClo}(\beta) \rightarrow \text{ExClo}(\gamma)$ and
 - (ii) $\text{ExClo}(\gamma) \rightarrow \text{ExClo}(\alpha)$ but
 - (iii) $\text{ExClo}(\alpha) \not\rightarrow \text{ExClo}(\gamma)$.

6 Subsentential cases

I close with a brief discussion of some interesting cases below the sentential level, which hopefully point the way for future research on focus and discourse. We have already seen some (barely) subsentential cases, where two coordinated clauses act as separate discourse moves in Roberts’s system. It seems as though this pattern can in fact be extended beyond simple conjunctions:

- (52) a. JOHN called MARY a REPUBLICAN, because he thought that SHE had insulted HIM.
 b. JOHN called UGLY the woman who had insulted HIM.

As in (4) and (28) above, the verb *insulted* here may remain unaccented after a hyponymous action, such as calling someone ugly, has been mentioned. This is somewhat unexpected, since it is not entirely clear that the sub-clauses of the sentences in (52) should count as separate discourse moves in Roberts’s system.

To take this idea one step further, consider example (53), due to Rooth (1992), who calls it the beginning of a joke¹⁰:

- (53) An AMERICAN farmer was talking to a CANADIAN farmer.

Notice that here even the first utterance of the word *farmer* is unaccented. Rooth calls these “contrasting phrases” and proposes that the lack of accent on *farmer* is in anticipation of the later use of the word – i.e., the two are mutual or symmetric antecedents.¹¹

Now, there are several interesting parallels between Rooth’s contrasting phrases and the cases presented in this paper that support non-identical antecedents, especially those like (4) and (28). First, we can construct clausal cases that exhibit the same, ostensibly symmetric, antecedence relation as Rooth’s sub-clausal case:

- (54) a. JOHN insulted HER, and then SHE insulted HIM.
 b. A: JANE saw some animals in the SUBWAY, and FRED saw some animals in the ZOO.

Notice that the first occurrences of the words *insulted* and *animals* in (54) may remain unaccented (and therefore GIVEN), seemingly in anticipation of the later occurrence of the same word. Although a detailed analysis of this phenomenon might prove tricky, it is not entirely unexpected under the QUD analysis presented in the sections above. A careful speaker, who perhaps has thought out her entire utterance before speaking, might realize that both clauses can answer the same implicit question and pronounce them accordingly. Such an analysis would not require the kind of symmetric antecedence relation that

¹⁰Rooth later suggests “They got into an argument about canola” as the second line of the joke, but the punchline is not revealed.

¹¹See Wagner (2012) and Katzir (2013) for an interesting debate – beyond the scope of this squib – about the restrictions on the focussed items in such pairs – e.g., *American* and *Canadian* here.

Rooth posits for his example.¹²

A second parallel between the cases considered here and Rooth's is the fact that we can construct sub-clausal cases that allow non-identical antecedents:

(55) An AMERICAN SQUIRREL was talking to a CANADIAN animal.

Here, the word *animal* may be unaccented after the utterance of its hyponym *squirrel*. A third parallel is the fact that contextual effects occur in Rooth-style cases similar to those seen in (28) above, where calling someone a Republican had to be taken as an insult for the pronunciation shown to be felicitous. For instance, consider (56), where the unaccented pronunciation of *idiot* is only acceptable if politicians are contextually agreed to be idiots:

(56) A NATIONAL POLITICIAN was talking to a LOCAL idiot.

These parallels suggest several intriguing analyses, which I will mention here briefly, leaving most of the details to future work. For instance, we could pursue an analysis that maintains Rooth's symmetric account of contrasting phrases. Under such an analysis, we would have to chalk the parallels noted here up to coincidence (or, more likely, deeper generalizations that encompass both clausal and sub-clausal cases). If we wanted instead to make the analysis of contrasting phrases closer to the QUD-based account given in this paper, we could attempt to analyze them without making any major alterations to Roberts's system. This would require allowing an implicit QUD like the one in (57) for an utterance of (58a), but not allowing this same implicit QUD for an utterance of (58b):

(57) ⟨What happened involving which nationality farmer?⟩

(58) a. An AMERICAN farmer was talking to a CANADIAN farmer.
b. #An AMERICAN farmer was TALKING.

As things stand, (57) is ruled out as the QUD for both the sentences in (58), due to *IMPLICITANTECEDENT. Therefore, the trick would be to propose why (58a) actually allows this implicit QUD. One possible approach would be to claim that (58a) actually answers (57) in two separate ways, because it describes what happened with an American farmer and also a Canadian farmer. If this somehow were to count as two discourse moves for the purposes of calculating

¹²Even if Rooth's symmetry story is correct, there is still much more to be explained, because not every case where a word appears twice leads to such symmetric GIVENNESS. First, cases lacking the parallel structure of Rooth's example resist symmetric GIVENNESS:

(i) #A dog OWNER bought a PICTURE of a dog. (cf. ✓A dog OWNER is always a dog LOVER.)

It seems to be crucial that some structure larger than the word in question be parallel for Rooth's symmetry to go through. Second, the distance between the symmetric words must be rather minimal, as shown by the following further attempts at humor:

(ii) a. ?An AMERICAN farmer thought that he had played a trick on a CANADIAN farmer.
b. #An AMERICAN farmer walked into a BAR. Then, a CANADIAN farmer walked in.

SPECIFICCOMMONQUD, we would be in business: this higher-ranked constraint would negate the violations of *IMPLICITANTECEDENT.

Finally, a most radical proposal could involve major changes to Roberts's QUD system. If we truly wanted to call Rooth's contrasting DPs discourse moves, we would have to find a way to categorize items other than clauses as discourse moves. This would require QUDs whose answers are individuals instead of propositions. Furthermore, we would have to allow discourse moves within other discourse moves, like the subject of sentence being a separate discourse move from the sentence itself.

References

- Büring, D.: 2003, On D-trees, beans, and B-accent, *Linguistics and Philosophy* **26**(5), 511–545.
- Carlson, L. H.: 1982, *Dialogue games: An approach to discourse analysis*, PhD thesis, Massachusetts Institute of Technology.
- von Stechow, K.: 1999, NPI licensing, Strawson entailment, and context dependency, *Journal of Semantics* **16**(2), 97.
- Green, G. M.: 1968, On too and either, and not just too and either, either, *Chicago Linguistics Society (CLS)* **4**, 22–39.
- Katirji, R.: 2013, A note on contrast, *Natural Language Semantics* pp. 1–11.
- Lakoff, G.: 1971, Presupposition and relative well-formedness, *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology*, Cambridge: Cambridge University Press, pp. 329–340.
- Roberts, C.: 2012, Information structure in discourse: Towards an integrated formal theory of pragmatics, *Semantics and Pragmatics* **5**(6), 1–69.
- Rochmont, M.: 1986, *Focus in Generative Grammar*, Studies in Generative Linguistic Analysis, Vol 4, John Benjamins Publishing Company.
- Rooth, M.: 1992, A theory of focus interpretation, *Natural Language Semantics* **1**(1), 75–116.
- Schwarzschild, R.: 1999, GIVENness, AvoidF and other constraints on the placement of accent, *Natural Language Semantics* **7**(2), 141–177.
- Wagner, M.: 2012, Focus and givenness: A unified approach, *Contrasts and positions in information structure* pp. 102–147.