

1 Entailment and Givenness

Theories of focus structure define GIVEN words and phrases as those whose surrounding discourse context provides an appropriate 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?
 - a. JAIME is Joffrey's father.
 - b. #Jamie is JOFFREY's father.
- (2) Whose father is Jaime?
 - a. #JAIME 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 noted that a GIVEN phrase and its antecedent needn't be identical. First, an antecedent may be quite different syntactically and phonologically, as long as 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.

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).

- (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:

- (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* may be GIVEN in this example, 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 §4, 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.

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 between antecedents and GIVEN phrases. Since the solution presented here will involve the Question Under Discussion model of discourse, §3 briefly summarizes this system. Next, §4 argues 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 discusses a few examples below the sentential level.

2 Counterexamples

The cases in (4) and (5) both involve a pair of sentences that share some material and differ in other material; the shared material counts as GIVEN in the second sentence. Rooth (1992) terms such cases “contrasting phrases.” There are plenty of other ways to set up antecedents for GIVEN material, though. The fol-

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

lowing examples are parallel to (5), but involve different discourse relationships. In both (6) and (7), an antecedent statement is followed by three alternative replies: in the first version of the reply, all the GIVEN material is identical to material in the antecedent statement; in the second version, the word *animals* replaces the word *gorillas*; in the third version, the word *animals* still replaces *gorillas*, but this time it is accented. As shown, the pattern of felicity in these cases is different from the pattern in (5):

- (6) A: Where did you see some GORILLAS today?
B: (i) I saw some gorillas in the ZOO today.
(ii) #I saw some animals in the ZOO today.
(iii) I saw some ANIMALS in the ZOO today, but not gorillas.
- (7) A: Jane saw some GORILLAS in the SUBWAY today.
B: (i) No, she saw some gorillas in the ZOO today.
(ii) #No, she saw some animals in the ZOO today.
(iii) No, she saw some ANIMALS in the ZOO today, but not gorillas.

Example (6) is a question/answer pair, another classic focus environment. In the (i) version of B's reply, the word *gorillas* is felicitously unaccented. Presumably this is because it is GIVEN, finding an identical antecedent in A's statement. The option in (6Bii), though, shows that the word *animals* can not felicitously be left unaccented. (The reply in (6Biii), where *animals* is accented, is acceptable.) A similar pattern obtains in (7), which shows a correction. Again, the identical word *gorillas* may be unaccented, but the hypernym *animals* is odd when unaccented.

One potential reply a defender of Rochemont's analysis might raise to these two examples is that due to their special contexts, the versions of B's replies labeled (iii) above – where the hypernym *animals* is accented – have slightly different meanings from the (intended) meanings of the replies labeled (ii). For instance, (6Bi) and (6Bii) seem intended to directly answer (6A), while (6Biii) instead seems to reject part of A's question. Similarly, (7Bi) and (7Bii) only seem to correct A's mention of the subway as the location of the animal sighting while (7Biii) also corrects the type of animal seen. Thus, while these cases raise a question for Rochemont's analysis, perhaps a full analysis of question/answer pairs and corrections could explain these cases.

There are counterexamples, however, for which this reply is not applicable:

- (8) A: I saw some GORILLAS in the SUBWAY today.
B: Hmm. That reminds me.
(i) There's a DOCUMENTARY on gorillas on PBS tonight.
(ii) There's a DOCUMENTARY on #primates/PRIMATES on PBS tonight.

In (8Bi), the term *gorillas* – which has an identical antecedent in A's statement – is felicitously unaccented (and therefore GIVEN). However, in (8Bii), the hypernym *primates* is infelicitous when accented and felicitous when unaccented. Furthermore, all versions of B's reply convey similar ideas: A's statement reminds

B of a documentary on a particular animal-related topic. Further examples are given in (9):

- (9) a. Sean STUDIED Calculus; he doesn't KNOW Calculus/#higher math.
b. What's worse than finding a WORM in your apple? Finding HALF (a worm/#a bug) in your (apple/#fruit).
c. I chose JOHN_i because I thought (he_i/#a man) would understand my situation.

Based on cases like these, I will argue in §4 that Rochemont's and Lakoff's examples 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 (6) - (9) do not fit Rochemont's pattern is that they project a different discourse structure from Rochemont's example. 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:

- (10) 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. Stalnaker 1978). In practice, most conversations start with a subquestion³ of this larger question, as illustrated by Roberts's example:

³Roberts defines a subquestion as one whose complete answer (contextually) entails at least a partial answer to its superquestion.

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

Here, the first question under discussion (QUD) – namely *Who ate what?* – is a subquestion of the Big Question. In this idealized example, each utterance (or “move” as Roberts terms them) either answers or poses a subquestion to the most recent unanswered QUD. For instance, move (11a) (*What did Hilary eat?*) poses a subquestion to the initial move, which happens to be the previous question asked. Move (11b) 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 (12), where questions, subquestions, and answers are all mixed up:

- (12) 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, (13) sounds fine, despite the fact that (13b) is not technically a subquestion of (13a)⁴:

- (13) (Roberts's (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 (13), the implicit move is the question *What reasons would John have for not eating clams?*, which is a technically correct subquestion of (13a) and a technically correct superquestion of (13b). In other words, the reason that (13) sounds better than (12) is that there is a recoverable implicit QUD which brings (13) into compliance with the proper rules of discourse. (See Büring (2003) for more evidence of, and restrictions on, implicit QUDs.)

⁴As Roberts explains, a complete answer of “No” to (13b) does not even partially answer (13a), going against the definition of questions and subquestions.

4 Implicit QUDs and Givenness

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

- (14) Mary called Sue a Republican, and then SHE insulted HER.
(Roberts's (52))

Roberts's analysis of this case is that the verb'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 (14) answer: *Who insulted whom?* As shown in (15), this question may be felicitously spoken aloud before an utterance of (14):

- (15) 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 (15A), 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 (14) (which lacks an explicit preceding question) can also be understood as maintaining an identity definition of GIVENness but 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 (16) where a preceding (explicit) question has been added:

- (16) 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 that this question is the implicit QUD in Rochemont's original example (5) will also allow us to maintain a purely identity-based relationship between the GIVEN node and its (unspoken) antecedent.

Unlike these cases, though, the examples in §2 are not felicitous answers to the same question. Example (6) is already a question/answer pair and therefore cannot answer a single question. And (7) is a correction, which in Roberts's system often addresses a different QUD than the statement it corrects.⁵ Last, (8) is a case where speaker B changes the topic of conversation away from the (implicit) QUD being answered by speaker A. In a system requiring an identical antecedent for each GIVEN node, then, the patterns of GIVENness in these three examples makes perfect sense: since there is no implicit QUD with an appropriate antecedent for the hypernyms *animals* and *primates*, these nouns do not count as GIVEN and therefore must be accented.

⁵ “[I]n corrections generally, the corrector addresses a different question ... than that addressed by the corrected interlocutor” (Roberts 2012, 44).

More evidence for this connection comes from modifications of these three examples, shown below, where an explicit antecedent for the noun *animals* or *primates* is actually uttered:

- (17) A: Where did you see some animals today? For instance, where did you see some GORILLAS?
B: I saw some animals in the ZOO today, but they weren't gorillas.
- (18) A: Where did people see animals today? Jane saw some GORILLAS in the SUBWAY.
B: No, she saw some animals in the ZOO today, and they weren't gorillas.
- (19) A: Who can tell me something about primates? For instance, I saw some GORILLAS in the SUBWAY today.
B: Well, there's a DOCUMENTARY on primates on PBS tonight.

Here, with explicit questions that mentions animals or primates, these words are in fact GIVEN in B's later statement.

Of course, the technical details are a little more complex, because we must define identity explicitly. This is potentially difficult, because Roberts' implicit QUDs are not full syntactically realized natural language objects. Instead, they 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 therefore 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 questions, as well as declarative sentences.

Under this new definition, the antecedent for any GIVEN phrase will be (allowed to be) an entire statement or question.⁶ One reason for this shift has to do with cases Schwarzschild points out, shown in (20):

- (20) (Schwarzschild 1999, Ex. (52))
A: Did Karen get the money or did Marc get the money?
B: KAREN/*Karen got the money.

Even though the word *Karen* is explicitly spoken in A's question, B's answer still must accent this name; it doesn't count as GIVEN. As we will see in a moment, this phenomenon is much easier to explain if we assume that phrases in an answer (such as *Karen* in (20B)) prefer to take the entire question as their antecedent, rather than singling out phrases within this question (such as *Karen* in (20A)).⁷

⁶This is possible, too, in the system from Schwarzschild (1999).

⁷See also Schwarzschild's (50), where choice of antecedent is similarly important:

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

One prerequisite to the definition of identity will be a notion of generalized existential closure (cf. Schwarzschild 1999, Ex. (26)). As defined in (21), the function ExClo turns (almost) any denotation into a proposition, via various types of existential closure:

- (21) 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 various nodes from (20):

- (22) a. $\text{ExClo}(\llbracket \text{Did Karen get the money or did Marc...?} \rrbracket)$
 $= \exists p_{st} [\llbracket \text{Did Karen get the money or did Marc...?} \rrbracket(p)]$
 $= \llbracket \text{Karen or Marc got the money} \rrbracket$
- b. $\text{ExClo}(\text{Karen}) = \exists P_{et}[P(\llbracket \text{Karen} \rrbracket)]$
 $\approx \llbracket \text{Something happened to Karen.} \rrbracket$
- c. $\text{ExClo}(\text{got}) = \exists x_e \exists y_e [\llbracket \text{got} \rrbracket(x)(y)]$
 $\approx \llbracket \text{Someone got something.} \rrbracket$
- d. $\text{ExClo}(\text{the money}) = \exists P_{et}[P(\llbracket \text{the money} \rrbracket)]$
 $\approx \llbracket \text{Something happened to the money.} \rrbracket$

In (22a), a question is converted into the proposition that one of its answers is true. In (22b) and (22d), an individual is converted into the proposition that holds if the individual has at least one property – e.g., if the individual exists. And in (22c), the transitive verb *got* becomes the proposition that someone got something.

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

- (23) GIVENness
 A phrase with denotation α is GIVEN iff its antecedent⁸ has denotation β such that
- a. $\text{ExClo}(\beta) \rightarrow \text{ExClo}(\alpha)$, and
 - b. there is no (“stronger”) alternative denotation γ of the same type as α such that $\text{ExClo}(\gamma) \rightarrow \text{ExClo}(\alpha)$ and $\text{ExClo}(\beta) \rightarrow \text{ExClo}(\gamma)$.

Clause (a) of (23) requires that (the closure of) a GIVEN node be entailed by (the closure of) its antecedent. Clause (b) disqualifies a node from being GIVEN if a 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.

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 previous sentence.

Under the definition in (23), *got* and *the money* count as GIVEN with the question *Did Karen get the money or did Marc get the money?* as their antecedent, but *Karen* does not. This is because the closure of this question – the proposition that Karen or Marc got the money – entails that someone got something and that something happened to the money. However, it does not entail that something happened to Karen.⁹

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

(24) ⟨Who insulted whom?⟩ John told Mary that she was ugly and then SHE insulted HIM.

(25) ⟨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.

ExClo(⟦Who insulted whom?⟧) entails ExClo(⟦insults⟧) and there is no stronger transitive verb also entailed by ExClo(⟦Who insulted whom?⟧). Therefore, pronouncing (24) as shown indicates that the hearer should recover this question as the implicit QUD – it is the only question that (a) is answered felicitously by both clauses and (b) licenses this pronunciation. Similarly, ExClo(⟦Where did people see animals today?⟧) entails ExClo(⟦animals⟧) but no one specific animal (taking people to be separate from animals). Therefore, this question is the implicit QUD recovered in (25).

When such an implicit QUD is not recoverable, though, the new definition of GIVENNESS does not allow similar phrases to be GIVEN. For instance, consider the counterexamples from §2, repeated here:

(26) A: Where did you see some GORILLAS today?

B: #I saw some animals in the ZOO today.

(27) A: Jane saw some GORILLAS in the SUBWAY today.

B: #No, she saw some animals in the ZOO today.

(28) A: I saw some GORILLAS in the SUBWAY today.

B: #There's a DOCUMENTARY on primates on PBS tonight.

The (closure of the) question in (26) does entail ExClo(⟦animals⟧); however, it also entails the stronger (“more identical”) ExClo(⟦gorillas⟧), disqualifying *animals* from being GIVEN. The closure of speaker A’s statements in (27) and (28) similarly entail this stronger meaning ExClo(⟦gorillas⟧), again disqualifying *animals* and *primates* from being GIVEN.

⁹This is a tricky point. The spoken question *Did Karen get the money or did Marc get the money?* seems to presuppose that Karen exists, in that it would be odd to say if there was no one salient named Karen. However, the question does not *entail* that Karen exists, in that some felicitous answers to this question are compatible with worlds where Karen doesn’t exist.

5 Subsentential cases

In lieu of a summary for this short squib, I close with a discussion of some interesting cases below the sentential level. First, consider the case from Rooth (1992) in (29). Rooth calls this example the beginning of a joke¹⁰, and so I have shown a request for this joke as the implicit QUD. Notice that even the first utterance of the word *farmer* is unaccented, presumably because a QUD like this (which mentions farmers) may be recovered upon hearing this sentence.¹¹ This is not the case in (30), though, where the two nouns differ from one another. In this example, the only recoverable QUD would mention citizens rather than farmers and thus would not license GIVENNESS for the word *farmer*.

- (29) ⟨Tell me the one about the two farmers.⟩
 An AMERICAN farmer was talking to a CANADIAN farmer.
- (30) An AMERICAN FARMER/#farmer was talking to a CANADIAN citizen.

Interestingly, similar contextual effects hold in this type of case as were seen in (20) above (where calling someone a Republican was taken to be insulting). For instance, consider (31), where the second occurrence of *farmer* is replaced by *woman*. Here, the word *woman* may in fact be unaccented/GIVEN despite being distinct from the word *farmer*. However, when *woman* is unaccented, the most salient interpretation available is one where the American farmer is also a woman. This makes perfect sense if hearers are allowed to recover an implicit QUD for this sentence that mentions women. (Such a QUD is shown in (31).) Last, the sentence is a bit odd if both the nouns *farmer* and *woman* are unaccented, as in (32), but to the extent that it is felicitous, both the people are taken to be female farmers. Perhaps this indicates a (somewhat hard to recover) implicit QUD that mentions female farmers.

- (31) ⟨Tell me the one about the two woman.⟩
 An AMERICAN FARMER was talking to a CANADIAN woman.
- (32) ⟨Tell me the one about the two female farmers.⟩
 An AMERICAN farmer was talking to a CANADIAN woman.¹²

References

Büring, D.: 2003, On D-trees, beans, and B-accent, *Linguistics and Philosophy* **26**(5), 511–545.

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

¹¹Without assuming such an implicit QUD, it would be quite hard to find an antecedent for the first occurrence of *farmer* in (29) that precedes it.

¹²See 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.

- Carlson, L. H.: 1982, *Dialogue games: An approach to discourse analysis*, PhD thesis, Massachusetts Institute of Technology.
- Green, G. M.: 1968, On too and either, and not just too and either, either, *Chicago Linguistics Society (CLS)* 4, 22–39.
- Katzir, 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.
- Rochemont, 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.
- Stalnaker, R.: 1978, Assertion, *Pragmatics*, Vol. 9 of *Syntax and Semantics*, New York: Academic Press, pp. 315–332.
- Wagner, M.: 2012, Focus and givenness: A unified approach, *Contrasts and positions in information structure* pp. 102–147.