VP-Deletion, Obligatory *too*, and Focus Semantics*

Raj Singh, MIT

December 31, 2008

The additive particle *too* seems to be obligatory in VP-ellipsis contexts: *John came to the store. Bill did* {# ∅/ *too*}. Pesetsky [25] attributes to Kai von Fintel (p.c. to Pesetsky) the suggestion that this fact might be derivable from Heim's [12] *Maximize Presupposition!* Amsili and Beyssade [1] suggest an implementation of von Fintel's idea. The implementation involves certain assumptions that do not follow from a general theory of *Maximize Presupposition!* I suggest a way to eliminate the stipulations by spelling out a precise set of assumptions needed to generate structures with elided VPs (Rooth [26], Heim [14]). This set of assumptions, along with a modification of Rooth's [27] theory of focus to allow the existence of multiple focus interpretation operators, suffice to derive the obligatoriness of *too* in these contexts. To the extent that the arguments here are sound, they also support such a modification to the theory of focus semantics.

1 Obligatory *too* and VP-Deletion

It has long been observed that *too* (among other discourse particles) is often obligatory in certain contexts (eg. Green [9, 10], Kaplan [17], Krifka [19], Zeevat [36], Sæbø[28], Amsili and Beyssade [1], Chemla [4]). Consider, as a special case, VP-deletion contexts:

1. (a) John came to the store, and Bill did too
   (b) # John came to the store, and Bill did

One thing to note immediately is that the effect appears in arbitrarily embedded contexts, below the level of speech act:

2. (a) If John came to the store and Bill did too, then Mary will be quite pleased
   (b) #If John came to the store and Bill did, then Mary will be quite pleased

*Thanks to Pascal Amsili, Claire Beyssade, Kai von Fintel, Danny Fox, Ted Gibson, Irene Heim, Roni Katzir, Benjamin Spector, and Bob Stalnaker.

3. (a) (Either John came to the store and Bill did too) or Mary doesn't know what she's talking about
   (b) #(Either John came to the store and Bill did) or Mary doesn't know what she's talking about

4. (a) Sue believes that John likes Mary and Bill does too
   (b) #Sue believes that John likes Mary and Bill does

The same asymmetry has also been noted with respect to certain ACD constructions. For example, Pesetsky [25] notes that the infelicity of (5a), noted by May [23], is eliminated by insertion of *also*:

5. (a) *Dulles suspected Philby, who Angleton did
   (b) Dullies suspected Philby, who Angleton also did

The question is: why?

In a footnote, Pesetsky attributes the following suggestion to Kai von Fintel (p.c. to Pesetsky).[1] The suggestion is that this contrast might be related to an observation von Fintel attributes to Irene Heim (p.c. to him). Imagine you have a guest over at your place. You can of course offer her a cup of coffee by asking, *Would you like a cup of coffee?* If the guest finishes her cup, and you want to offer her another, you cannot do so by asking, *# Do you want a cup of coffee?* You have to ask, *Do you want another cup of coffee?*. This feels entirely like an effect of Heim's [12] *Maximize Presupposition!* (henceforth MP). MP is a principle of language use, a competition-based blocking principle that says something like the following:[2]

Maximize Presupposition! If ϕ, ψ are members of a formally defined set of alternatives, and the presuppositions of ψ are stronger than those of ϕ, and are met in the context of utterance c, and ϕ and ψ add the same new information to c, then one must use ψ in c.

For our current purposes there are two things to note about this principle. First, one needs a well-defined set of alternatives. Generally speaking, algorithms that have been proposed for generating alternatives[3] ensure that the alternatives are of equal syntactic complexity.[4] Second, one needs to appeal to a theory of presupposition projection in order to associate a presupposition with each alternative. If MP is to be used, there should (ideally) be no case by case stipulations concerning the class of alternatives, or concerning the presuppositions of alternatives. These should follow from the theory of alternatives, and the theory of projection.

[1] FN. 32 of Pesetsky [25].

[2] See Heim [12], Sauerland [29, 31], Percus [24], Schlenker [32], Chemla [3, 4], Magri [22], Singh [33] for proposals concerning the proper formulation of the principle.

[3] cf. The references in FN. 2.

[4] The algorithms are very similar, possibly identical, to algorithms for generating alternatives for scalar implicature (eg. Horn [15], Sauerland [30], Katzir [18], and much other work).

Turning to the Heim/von Fintel/Pesetsky sentences, MP would work as follows. The two questions, *Do you want a cup of coffee?* and *Do you want another cup of coffee* are alternatives (they are of the same complexity, differing only in the position of one terminal node), and they are equivalent in the context of use (they ask the same question, namely, whether the guest wants a cup of coffee), but the sentence with *another* carries a stronger presupposition (namely, that the guest already had a cup of coffee), a presupposition that is met in the context. Under MP the question with *another* is therefore optimal, and so the one with the indefinite article is blocked. The suggestion that Pesestky attributes to von Fintel is that the VP-deletion sentences with *too* block the ones without for the same reason that the question with *another* blocks the one with *a*, namely, because of MP. Amsili and Beyssade [1] follow this suggestion and attempt to use it to derive the obligatoriness of *too* in these and other environments. They stipulate that sentences with *too* compete with the sentences without, and that MP determines that the former are better competitors than the latter.

The reader will have noted that, as it stands, this account faces at least three difficulties.[5] First, in the relevant constructions, the sentences are all presuppositionless. For example, focussing now on the basic contrast in (1), neither sentence carries any presupposition. Sentence (1a) contains a constituent, *Bill did [come to the store] too*, which does carry a presupposition (roughly, that there is some salient individual different from John who came to the store). But this presupposition is not inherited by the sentence as a whole, given the projection properties of conjunctive sentences. Percus [24] first discovered cases of this sort. I have argued elsewhere (Singh [33]) that such difficulties can be avoided by checking MP in the local context (Heim [11]) of each constituent, rather than at the root. In (1), for example, we would need to check the presuppositions of the second conjunct in the context as it's updated by the first conjunct. I believe this effectively solves the first difficulty.

The second difficulty concerns the set of alternatives. Recall that it is a general feature of the alternatives that enter into MP competitions that they are equally complex. But in our VP-ellipsis cases, the 'winner' of the competition, *Bill did too*, is more complex than its losing competitor *Bill did*, given the existence of *too*. Such a stipulation is clearly undesirable. At the same time, if an MP story is to account for the contrasts in (1)-(5), we need the sentences with and without *too* to compete with one another. Chemla suggests that *too* might rest on a scale with a null morpheme $\emptyset$. Such a scale could then be input to the various algorithms that have been proposed for generating alternatives. But without motivation for this scale, and without understanding the properties of this null morpheme, the proposal lacks the desired generality.

The third difficulty concerns the presuppositions of the competing alternatives. What does it mean to say that *Bill did too* presupposes something stronger than *Bill did* in its local context? What *are* the presuppositions of

[5]There are actually four. Irene Heim points out to me (p.c.) that *too* is not exactly obligatory. The right pronunciation can obviate the obligatoriness. I will not have anything to say about this here, though I hope the line of analysis can be made consistent with this fact.

these sentences, and how do they relate to a general theory of VP deletion?

My goal in this squib is to address the second and third difficulties mentioned just above. More specifically, by spelling out some of our assumptions about VP-deletion, I hope to show that we can provide an MP account for the above facts without any stipulation concerning alternatives while also providing determinate answers concerning the presuppositions of the alternatives. The approach to ellipsis I follow is one first proposed by Rooth [26], and formalized and defended in Heim [14]. This approach to ellipsis itself depends on Rooth's [27] theory of focus interpretation. I turn to some of the basics of this approach to ellipsis in the next section.

2 Focus and VP-Ellipsis

Rooth [27] assumes that focus is syntactically represented, so that the LF of a sentence like *JOHN came to the store* (pitch accent on *John*) is: ϕ = [John$_F$ [came to the store]]. The focus semantic value of ϕ, $[[\phi]]^F$, is: $[[\phi]]^F = \{x$ came to the store: x an individual$\}$. This is an awfully big set, and focus effects usually show up only with highly restricted subsets Γ, or particular members γ, of this set. To capture this, Rooth offers the following as the main principle governing the interpretation of focus:

Focus Interpretation Principle In interpreting focus at the level of phrase ϕ, add a constraint that: (i) $\Gamma \subseteq [[\phi]]^F$, or (ii) $\gamma \in [[\phi]]^F$, where Γ is a variable with the type of a set of objects each of which matches ϕ in type, and γ is a variable matching ϕ in type.

Focus-sensitive semantic effects will be limited to either Γ or γ, a subset or member of the focus semantic value of ϕ, and it will be up to the context to determine the values of these variables. Rooth formalizes this context-dependency by introducing a squiggle operator $\sim$ to LFs. The function of this operator is to introduce the above constraints as presuppositions. Thus, the LF of *JOHN came to the store* would actually be one of: (i) [John$_F$ [came to the store]] $\sim \Gamma$, or (ii) [John$_F$ [came to the store]] $\sim \gamma$. Rooth argues that these LFs have additional presuppositions about Γ/γ. Here is the full statement of the presuppositional constraints on LFs containing the squiggle operator:

Presuppositional Constraints Introduced by the Squiggle (i) **Set Case:** $\phi \sim \Gamma$ presupposes that Γ is a subset of $[[\phi]]^F$, and contains both $[[\phi]]$ and an element distinct from $[[\phi]]$, (ii) **Individual Case:** $\phi \sim \gamma$ presupposes that γ is an element of $[[\phi]]^F$ distinct from $[[\phi]]$.

Now we need to say something about VP-deletion. Here I will follow Heim [14]. First, we need to define a notion of *appropriate contrast*:

Appropriate Contrast An LF ψ *contrasts appropriately* with LF ϕ iff: (i) ϕ is distinct from ψ, (ii) $[[\phi]] \in [[\psi]]^F$, i.e. ϕ is an element of the focus semantic value of ψ.

With this definition, the following condition is imposed on the licensing of VP-deletion:

Licensing VP-Deletion A VP can be deleted if it is contained in a phrase that contrasts appropriately with some phrase that contains the antecedent VP.

3 The Alternatives and their Presuppositions

So far, I have introduced some fairly innocuous assumptions from focus semantics and VP-ellipsis. Note that, already, these assumptions potentially give us a handle on the first of our concerns, namely, the complexity of the alternatives involved. For note that whatever the precise structure assigned to *BILL did too*, the LF of *BILL did* will have more complexity than the phonology reveals, since there will be a squiggle operator appended to it. Here is the LF of (1b):

6. $[PAST[\text{John come to the store}]]_7$, and $[PAST[[\text{Bill}]_F \text{ came to the store}]]\sim_7$

The above LF satisfies the appropriate contrast condition, and so receives the phonology indicated in (1b). But why is it odd, if it satisfies the relevant condition? If we want to follow von Fintel's intuition and extend MP to account for this fact, we seem to be forced to say that the oddness comes because it is blocked by the following structure, which we take to tbe LF of (1a):

7. $[PAST[\text{John come to the store}]]_7$, and $[PAST[[\text{Bill}]_F \text{ came to the store}]]too_7$

In other words, the parse with a squiggle operator must be competing with a parse with *too*. Note that the structure in (7) satisfies the appropriate contrast condition, and so licenses deletion of the VP. But the reader will no doubt have noticed that (7) contains no squiggle operator of its own. Is this problematic, given that $\sim$ is what interprets focus? We will see. For now, I have to make this assumption to ensure that the competitors are of equal complexity. To run an MP story here, I also need to ensure that the second conjunct in (7) carries stronger presuppositions than its variant in (6). Note that the presupposition of (6) is entirely determined by the theory of focus, as encoded in Rooth's *Presuppositional Constraints Introduced by* $\sim$ just above. To determine the presupposition of the LF in (7), we need to provide a lexical entry for *too*. Modelled after Kripke [20] and Heim [13], I propose the following entry:[6]

Lexical Entry 1 (Definedness Conditions of *too*)

Let ϕ *too*$_i$ *be an LF, with* too *co-indexed with LF* ψ_i*. Then* $c + \phi$ *too*$_i$ *is defined iff: (a)* $[[\psi]] \neq [[\phi]]$*, (b)* $[[\psi]] \in [[\phi]]^F$*, (c)* $c + \psi = c$ *(i.e.* ψ *is true in* c*).*

[6]Note that Kripke and Heim envision *too* as being anaphoric to the subject of the first sentence (when there is focus on the subject), not the entire clause. I have made this modification because it will make it easier for me to state the relevant competition principles.

We have everything we need now to use MP to account for the blocking effect. First, our structures are equally complex.[7] Second, we have a principled account of the presuppositions of the competing structures. The presuppositions of the structure in (6) follow entirely from the theory of focus and ellipsis assumed here. The presuppositions of (7) encode three definedness conditions that come from our (standard) entry for *too*: conditions (a) and (b) are identical to the presuppositions of $\sim$ (Individual Case), and condition (c) has an added requirement, namely, that the co-indexed LF be true in the local context. In our example, this would be the presupposition that John did, in fact, come to the store. Structure (7) thus carries a stronger presupposition than the one in (6). Given this, MP prefers (7) over (6), hence (1a) over (1b).

To summarize, the assumptions we needed to address the problem of alternatives and the problem of projection were:

- the theory of focus and ellipsis (Rooth [27, 26], Heim [14])
- the standard entry for *too* borrowed from Kripke [20] and Heim [13]
- the assumption that the structure with *too* does not contain a squiggle operator of its own

Of these assumptions, the last one is potentially controversial, since, under Rooth's [27] theory of focus interpretation, it is $\sim$ (and nothing else) that interprets focus. We needed there to be no $\sim$ in the LF in (7) so that it would be of equal complexity as the LF in (6). For this account to work, then, we would need to allow *too* to interpret focus directly, without $\sim$. But would such a move distort Rooth's [27] theory of focus?

4 The Theory of Focus

Rooth [27] states the theory of focus as consisting of the following axioms:

Rooth's Theory of Focus The theory of focus consists of: (a) Rules describing the phonological interpretation of the focus feature F, (b) Two-dimensional alternative semantics, defining focus semantic values with reference to F and ordinary semantic values, (c) The semantic clauses for the squiggle operator $\sim$, (d) The rule introducing $\sim$ in LF

What I am contemplating is allowing the existence of focus interpretation operators other than $\sim$. Recall that we needed the structures corresponding to (1a) and (1b) to be equally complex. We achieved this by eliminating $\sim$ from the structure for (1b). But then something needs to interpret focus, and I cannot see any reason why *too* cannot interpret focus directly without $\sim$. If we

[7] This could be derived by positing a scale with $\sim$ and *too* on it, or using Katzir's [18] procedure for tranforming trees into trees. I will later argue that *too* should be taken to be a focus interpretation operator like $\sim$, which would make the idea that they enter into competitions much more natural.

were to allow this, the set of focus interpretation operators **FIO** would include at least the following as a subset: {∼, too}. To keep the theory as constrained as possible, then, I suggest that any FIO be forced to obey the semantic clauses for ∼. In other words, each focus interpertation operator must obey the schema outlined in the Focus Interpretation Principle and, more generally, must include at least the presuppositions currently assigned to the squiggle operator ∼:

Presuppositional Constraints Introduced by Focus Interpretation Operators

Let O be a Focus Interpretation Operator. Then: (i) **Set Case** $\phi\ O\ \Gamma$ presupposes that Γ is a subset of $[[\phi]]^F$, and contains both $[[\phi]]$ and an element distinct from $[[\phi]]$, (ii) **Individual Case** $\phi\ O\ \gamma$ presupposes that γ is an element of $[[\phi]]^F$ distinct from $[[\phi]]$.

In addition to this, different FIO's may or may not bring additional presuppositional statements. However, it is important to note that focus effects would still be derived as theorems of Rooth's theory. The only modification would be the inclusion of a set of FIO's in addition to the squiggle. Thus, the change to the theory would consist of a modification to axiom (c): The semantic clauses would not be for ∼, but for any FIO O (invariant across the members of O). And, within the set **FIO**, variation would be limited to only arbitrary properties of the phonology and the semantics, but would make no reference to focus. For instance, our entry for *too* differs from the entry for ∼ only in that it is overtly realized and it carries an additional presupposition.[8] There are no stipulations related to focus, other than that it is an FIO, as with ∼.

If this much is accepted, we might contemplate expanding **FIO** to include more items: {∼, too, only, even, again, still, exh, ...}.[9] This would require a modification to the LFs assumed by Rooth. For instance, a sentence like *only JOHN came to the party* would have LF $only(\Gamma)(\phi)$, instead of Rooth's $only(\Gamma)(\phi) \sim \Gamma$. Our definedness condition for such an LF would then be forced (by the schema for all FIO's) to include the presuppositions introduced by clause (i) of the general schema (the set case). However, since we allow FIO's to encode phonological/semantic properties not dictated by the theory of focus, the entry for *only* could carry stronger presuppositions without distorting the general schema. In Singh [34], I develop a presuppositional analysis of *only* and *exh* that obeys this general schema, and argue that these entries, in conjunction with an appropriately localized version of MP, provides a solution to various puzzles concerning the (in)felicity of Sobel/Reverse Sobel Sequences (eg. Lewis [21], von Fintel [6]) and Hurford/Reverse Hurford Disjunctions (eg. Hurford [16], Fox [7, 8], Singh [35], Chierchia, Fox, and Spector [5], Fox and Spector [5]). I leave discussion of this complex of ideas for a future paper.

[8] The additional presupposition is the one listed as (c). The first two presuppositions ((a) and (b)) are inherited by virtue of *too*'s being a FIO.

[9] Though see Beaver and Clark [2] for arguments that not all focus sensitive operators can be treated uniformly.

References

[1] Pascal Amsili and Claire Beyssade. Compulsory presupposition in discourse. In *Proceedings of the Second Workshop on Constraints in Discourse*. Maynooth, 2006.

[2] David Beaver and Brady Clark. *Always* and *Only*: Why not all focus-sensitive operators are alike. *Natural Language Semantics*, 11:323–362, 2003.

[3] Emmanuel Chemla. French *both*: a gap in the theory of antipresupposition. *Snippets*, 15:4–5, 2007.

[4] Emmanuel Chemla. An epistemic step for anti-presuppositions. *Journal of Semantics*, 25:141–173, 2008.

[5] Gennaro Chierchia, Danny Fox, and Benjamin Spector. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Paul Portner, Claudia Maienborn, and Klaus von Heusinger, editors, *Handbook of Semantics*. Mouton de Gruyter, New York, NY, 2008.

[6] Kai von Fintel. Counterfactuals in a dynamic context. In Michael Kenstowicz, editor, *Ken Hale: A Life in Language*. MIT Press, Cambridge, MA, 2001.

[7] Danny Fox. Scalar implicatures: Pragmatics or grammar or both. Lecture Notes, 24.959: Topics in Semantics, MIT, 2006.

[8] Danny Fox. Too many alternatives: Density, symmetry, and other predicaments. *Proceedings of SALT 17*, To Appear.

[9] Georgia M. Green. On *too* and *either*, and not just on *too* and *either*, either. *CLS*, 4:22–39, 1968.

[10] Georgia M. Green. The lexical expression of emphatic conjunction. theoretical implications. *Foundations of Language*, 10:197–248, 1973.

[11] Irene Heim. On the projection problem for presuppositions. In *WCCFL 2*, pages 114–125, 1983.

[12] Irene Heim. Artikel und definitheit. In *Semantics: An International Handbook of Contemporary Research*. de Gruyter, Berlin, 1991.

[13] Irene Heim. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9:183–221, 1992.

[14] Irene Heim. Predicates or formulas? evidence from VP-ellipsis. In Aaron Lawon and Eund Cho, editors, *Proceedings of SALT 7*, pages 197–221. CLC Publications, Ithica, NY, 1997.

[15] Laurence R. Horn. On the semantic properties of logical operators in english. Doctoral Dissertation, UCLA, 1972.

[16] James Hurford. Exclusive or inclusive disjunction. *Foundations of Language*, 11:409–411, 1974.

[17] Jeff Kaplan. Obligatory too in english. *Language*, 60:510–518, 1984.

[18] Roni Katzir. Structurally defined alternatives. *Linguistics and Philosophy*, 30:669–690, 2007.

[19] Manfred Krifka. Additive particles under stress. In *Proceedings of SALT 8*, pages 111–128. CLC Publications, Ithica, NY, 1999.

[20] Saul Kripke. Presupposition and anaphora: Remarks on the formulation of the projection problem. Manuscript, Princeton University, 1990.

[21] David Lewis. *Counterfactuals*. Oxford University Press, Oxford, 1973.

[22] Giorgio Magri. A theory of individual level predicates based on blind scalar implicatures. Manuscript, MIT, 2007.

[23] Robert May. *Logical Form: Its Structure and Derivation*. MIT Press, Cambridge, MA, 1985.

[24] Orin Percus. Antipresuppositions. In Ayumi Uyema, editor, *Theoretical and Empirical Studies of Reference and Anaphora: Toward the Establishment of Generative Grammar as an Empiricial Science*, pages 52–73. Report of the Grant-in-Aid for Scientific Research (B), Project No. 15320052, Japan Society for the Promotion of Science, 2006.

[25] David Pesetsky. *Phrasal Movement and its Kin*. MIT Press, Cambridge, MA, 2000.

[26] Mats Rooth. Ellipsis redundancy and focus redundancy. In Steve Berman and Arild Hestvik, editors, *Proceedings of the Stuttgart Ellipsis Workshop, Report Series Sprachtheoretische Grundlagen fur die Computerlinguistik, no 29, SFB 340*. IBM Heidelberg, 1992.

[27] Mats Rooth. A theory of focus interpretation. *Natural Language Semantics*, 1:75–116, 1992.

[28] Kjell Johan Sæ bø. Conversational contrast and conventional parallel: Topic implicatures and additive presuppositions. *Journal of Semantics*, 21:199–217, 2004.

[29] Uli Sauerland. Implicated presuppositions. Handout for a Talk Presented at the University of Milan Bicocca, Milan, Italy, 2003.

[30] Uli Sauerland. Scalar implicatures in complex sentences. *Linguistics and Philosophy*, 27:367–391, 2004.

[31] Uli Sauerland. Implicated presuppositions. In A. Steube, editor, *Sentence and Context: Language, Context, and Cognition*. 2008.

[32] Philippe Schlenker. Maximize presupposition and gricean reasoning. Manuscript, UCLA and Institute Jean-Nicod, Paris, 2006.

[33] Raj Singh. *Maximize Presupposition!* and local contexts. Manuscript, MIT, 2008.

[34] Raj Singh. Modularity and locality of interpretation. Doctoral Dissertation, MIT, 2008.

[35] Raj Singh. On the interpretation of disjunction: Asymmetric, incremental, and eager for inconsistency. *Linguistics and Philosophy*, 31:245–260, 2008.

[36] Henk Zeevat. Particles: Presupposition triggers, context markers, or speech act markers. In Reinhard Blutner and Henk Zeevat, editors, *Optimality Theory and Pragmatics*. Palgrave Macmillan, 2003.