University of Massachusetts Occasional Papers in Linguistics

Volume 15 *Papers on Quantification* Article 6

1989

Exception Phrases

Kai von Fintel

Exception Phrases[1]

Kai von Fintel

Object of investigation in this paper is the meaning of exception phrases (EPs) in quantified noun phrases as in (1) and (2).

(1) Every student but John attended the meeting.

(2) No student but John attended the meeting.

While I take English *but*-phrases to represent the core specimen of EPs, there is a number of different types of EPs with varying syntactic and semantic properties. My strategy in this paper will be to develop a more or less satisfactory account of *but* first and then to take a look at the properties of some variants.

The semantic properties of *but*-phrases that need explaining are first the complex truth-conditions they give rise to and second the restrictions on the main determiners that they can occur with. As for meaning, (1) conveys at least the following: all the students other than John attended, John is a student, John did not attend. The negative counterpart (2) says that no student other than John attended, again that John is a student, and that John did attend. Both therefore have a subset entailment that John is a member of the set of students. And there is a polarity reversal between (1) and (2) depending on the main determiner regarding the relation between the exception and the main predicate, I will call this phenomenon the *no*-flip.

As regards the co-occurrence restrictions, *but* can only occur with universal determiners, (3) is ungrammatical.

(3) Three students but John attended the meeting.

1. Syntax and previous accounts

The choice of a semantic analysis of the contribution of EPs to the meaning of a sentence is largely dependent on the prior choice of a syntactic analysis. I will assume without an argument that the internal constituency of EPs is that of a prepositional phrase.[2] Their external constituency, however, i.e. the question of the attachment site of EPs in a noun phrase, is as mysterious as the parallel problem of the attachment site of relative clauses.[3] Among the many possibilities, the three illustrated in (4) will play a role in this paper: (i) discontinous determiner, (ii) NP-modifier, (iii) N'-modifier.

---

[1]A previous version of this paper was presented at the Conference on Cross-Linguistic Quantification in Tucson 1989. For comments and suggestions, thanks go to Barbara Partee, Hotze Rullmann, Ginny Brennan, Jim McCawley, Eloise Jelinek, Claudia Casadio, Sue Schmerling, Polly Jacobson, and Jim Tyhurst. Research on this paper has been partially supported by the NSF under Grant No. BNS 87-19999.

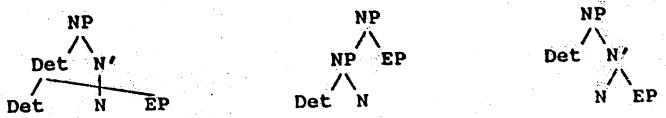
[2]This seems largely uncontroversial. Landman & Moerdijk (1979) defend this assumption explicitly for Dutch *behalve*. Azoulay-Vicente (1988) argues that *que* in the French *ne...que*-construction is a preposition. The English preposition *but* does not seem to be synchronically related to the conjunction *but*. The latter occurs in a construction type very similar to some uses of EPs but, significantly, shows absence of the characteristic *no*-flip of EPs. Compare the EP in (i) with the stripping construction in (ii).

(i) I met every student but John.

(ii) I met every student, but not John.

[3]The relative clause-attachment problem seems to have been neglected in recent research. Older references include: Stockwell, Schachter & Partee (1973), Bach & Cooper (1978).

(4)



The two analyses listed first in (4) have been employed in the two previous approaches to EPs in the semantics literature that I will now briefly review.

Keenan & Stavi (1986) treat EPs as forming a discontinuous constituent together with the main determiner. This is associated with the semantics in (5).

(5) $\|no...but\ John\|(A) = (P \subseteq E \mid P \cap A = (j))$.
$\|all...but\ John\|(A) = (P \subseteq E \mid \bar{P} \cap A = (j))$.

While this fulfills the requirement of getting the meaning of sentences with EPs right, it has considerable shortcomings as an explanatory account. The meaning of the discontinuous determiners in (5) is not in any obvious way derived in a compositional manner by combining the meanings of the main determiner and the EP. Furthermore, the co-occurrence restrictions of *but*-phrases are not captured, indeed nothing prevents us from defining a hypothetical discontinuous determiner *some...but John* to have the semantics in (6).

(6) $\|some...but\ John\|(A) = (P \subseteq E \mid P \cap A \neq \emptyset$ *and* $(j) \in \bar{P} \cap A)$.

Hoeksema (ms), discussing a variety of possible approaches, develops something like the view of Keenan & Stavi's, but in the end adopts an analysis that treats EPs as modifiers on the NP-level. This forces him to "go intensional": EPs operate on the model of evaluation and subtract entities from the domain of entities for the purposes of evaluating a universal claim. The formalization of this idea is given in (7).

(7) $\|NP\ but\ c\|_E = (X \subseteq E \mid X\text{-}(\|c\|) \in \|NP\|_{E\text{-}(\|c\|)}$ *and* $X \in \|NP\|_E)$
where NP is closed under submodels and additive.

This approach captures the truth-conditions of sentences with EPs and is very possibly compositional. The co-occurrence restrictions, however, are simply stipulated by demanding the NP combining with EPs to have the two model-theoretic properties of closure under submodels and additivity.

The two previous proposals then have no problem accounting for the truth-conditions of EPs but have to resort to stipulations to capture their co-occurrence restrictions.

2. The proposal

The proposal that I would like to put forward in this paper is based on a semantic intution similar to Hoeksema's, namely that EPs subtract entities from sets. It seems to me though that the set being subtracted from is the common noun set of the NP containing the EP and not the whole universe of discourse. The syntactic analysis that goes with this approach is one that treats EPs as modifiers on the N'-level, which is the third possibility

---

[4]These model-theoretic properties are defined as follows:
(i) Closure under submodels.
If $E' \subseteq E$ and $X \in \|NP\|_E$, then $X \cap E' \in \|NP\|_{E'}$.
(ii) Additivity.
If $(X \cap E) \in \|NP\|_E$ and $(X \cap E') \in \|NP\|_{E'}$, then $X \cap (E \cup E') \in \|NP\|_{E \cup E'}$.
These two properties together single out the universal determiners, as shown by van Benthem (1984: 455).

illustrated above in (4). The arguments for this constituency are not entirely decisive which will prompt us to consider a different strategy later.

By the standard constituency tests, we expect the N'-constituent supposedly consisting of the common noun and the EP to move as a unit and to be substitutable as a unit. Lappin (1988) presents N'-anaphora facts that show the expected substitution behavior. The most readily available reading for (8) is one where Mary saw several students other than John.

(8) Bill saw no student but John although Mary saw several.

This suggests that the null anaphora goes back to the set *students-{John}* denoted by *student but John.*[5]

Data where the N' in question seems to move can be found in German topicalization. As is well known, the topic position before a finite verb in second position in a German matrix sentence can only be occupied by a constituent. Consider then (9) which leads me to conclude that *studenten ausser Hans* is a constituent.

(9) [Studenten ausser Hans]$_i$ waren keine t$_i$ da.
students but Hans were none there

Before some counterexamples are discussed, I will develop this hypothesis further and give an attractive semantics for it, which will make it desirable to save the appearances rather than to discard our theory.

EPs under this view subtract entities from the set denoted by the common noun they are in construction with. Thus, we have the rather simple semantics in (10).

(10) $\|CN \text{ but } c\| = \|CN\| - \{\|c\|\}$.

Now, (10) is certainly too simple-minded. Hoeksema in discussing this possibility rightly points out that this does not at all capture the co-occurrence restrictions of EPs. The set of students minus John is not special in any way, so again nothing prevents the existence of NPs like *some student but John.*

But there is of course something special about the set of students minus John in the offending example: it is denoted by an expression that contains an EP. I propose to augment the basic semantics in (10) with a condition on the proper use of EPs as formalized in (11). This Minimality Condition demands that the set subtracted from the common noun set be the smallest set that subtracted makes the sentence true, in other words: the EP has to denote the smallest necessary exception.

(11) Minimality Condition
$\forall Y\ (P \in \|D(A\text{-}Y)\| \Rightarrow \{\|c\|\} \subseteq Y)$.

The status of this condition will be discussed later, for the moment I will assume that it is accomodated into the eventual meaning of the NP-cum-EP, giving as a result (12).

(12) $\|D(A \text{ but } c)\| = \{P \subseteq E \mid P \in \|D(A\text{-}\{\|c\|\})\| \ \&\ \forall Y\ (P \in \|D(A\text{-}Y)\| \Rightarrow \{\|c\|\} \subseteq Y)\}$.

The final semantics (12) can now be shown to capture all the pertinent facts about EPs. I will use as test cases the sentences (1) and (2), repeated here as (13) and (14).

(13) All the students but John attended the meeting.
(14) No student but John attended the meeting.

Basic meaning. The first conjunct in (12) gives the correct basic meaning for our test sentences: with the proper assignments, $P \in \|D(A\text{-}\{\|c\|\})\|$ will mean that in (13) all the students other than John attended the meeting and that in (14) no student other than John attended.

[5] There is an interfering but irrelevant reading where the N'-anaphora goes back to the smaller N'-node consisting only of *student*. This reading was pointed out to me by Jim McCawley.

Subset entailment. (12) ensures that (‖c‖) will be a subset of A. Suppose the contrary: the non-empty (‖c‖) ⊈ A. Then A-(‖c‖) ≠ A. But then P ∈ ‖D(A-(‖c‖))‖ <==> P ∈ ‖D(A)‖. But then by minimality (11) (‖c‖) = ∅, contrary to the assumption. Hence (‖c‖) ⊆ A. For (13) and (14), this means that John is entailed to be a student, as desired.
The no-flip. The Minimality Condition on the exception c means that A-(‖c‖) is the largest subset of A that makes the sentence true. Adding ‖c‖ to it would make the sentence false. Now, for (13) this means that John cannot be added to the set of students who attended the meeting, hence he didn't attend. For the negative (14), John cannot be added to the set of students who didn't attend, hence he did attend. Thus, the *no*-flip is correctly predicted.
Restriction to universal determiners. Uniqueness of the exceptions set as demanded by (11) is only guaranteed by universal determiners. Consider (15).
(15) Three students but John attended the meeting.
This would mean that if John is subtracted from the set of students then (exactly) three students attended the meeting. But note that we could have easily subtracted one of the other three attending students to make the sentence true. There is no unique smallest exception as demanded by (11). This is only possible with universal determiners.[6]

Having seen that the analysis proposed here has the very desirable properties of being rather intuitive and of explaining the facts that needed explaining, we can now turn to counterexamples. The N'-analysis for EPs encounters the same problems that an N'-analysis for relative clauses faces. There are cases where the EP clearly attaches to an NP because the NP is just one word.[7] Consider the sentences (16) and (17).
(16) Everybody but John attended the meeting.
(17) Nobody but John could have done this.
It would be rather absurd (although diachronically plausible) to claim that the word *everybody* is in fact syntactically composed of *every* and *body*. So an NP-level syntax seems unavoidable. Is there a way of retaining the N'-semantics at the same time?

There is a solution to this problem which was described by Bach & Cooper (1978), who showed that an N'-semantics for relative clauses is compatible with syntactic attachment higher in the tree (at the NP- or even S-level). The crucial device is the introduction of a free variable at the N'-level which can then later be filled in by the relative clause. We could adopt this solution and assume that any common noun routinely introduces a free variable. Now, the denotation of *everybody* could then be *every(person ∩ R)*[8] where R is the free variable. The EP can then fill in for the free variable and with some wizardry we get the denotation *every(person-{j})* for *everybody but John*.
The free variable approach might draw supportive evidence from the behavior of the word *other* which in constructions as in (18) could be argued to mark the fact that a specification of the free variable is given either in form of an EP or by contextual means.
(18) Except for John, all the other students had fun.

I will leave it open whether we should just adopt an NP-syntax across the board and couple it with an N'-semantics in the way outlined here, or whether the attachment site of EPs is in fact ambiguous. It seems that this problem should be solved together with the problem of relative clause attachment.

In the next two sections, I am going to look at types of EPs that make it necessary to modify and extend the approach formulated so far.

[6] I have no formal proof of this claim but do not anticipate any problems here.
[7] This problem was pointed out to me by Polly Jacobson.
[8] Instead of the Montague notation used by Bach & Cooper, I stick to the Barwise & Cooper notation used throughout this paper.

3. Liberal EPs

Not all EPs are restricted to co-occurrence with universal determiners. In fact, sentences like (15) are fine if liberal EPs formed with *besides*, or German *ausser* are used, as shown in (19).

(19) a. Three students besides John were present.
b. Drei Studenten ausser Hans waren da.

So the Minimality Condition does not seem to hold in all its force in these cases. Rather than abandoning it completely, I will propose a slightly more liberal version of it that can account for the behavior of *besides*. The proposal is given in (20).

(20) Minimality Condition II
$\forall Y\,(P \in \|D(A\text{-}Y)\| \Longrightarrow |(\|c\|)| \leq |Y|)$.

The exception set has to be the smallest set that makes the sentence true if subtracted, but smallest in the cardinal sense, not in the stronger subset sense of (11). This explains why EPs that are associated with the condition in (20) can co-occur with any cardinal determiner.[9]

It can be easily checked that apart from this difference in co-occurrence restrictions, there will be no further differences between EPs obeying (11) or (20). Note that, in particular, nothing in the semantics for *besides* prevents it from occurring with universal determiners. That is empirically adequate as (21) and (22) show.

(21) Everyone besides Richard attended.
(22) Noone besides Richard attended.

The meaning we predict for these cases is the same as if *but* or *except* had been used, since with universal determiners cardinal minimality of the exception is equivalent to subset minimality.

This is the point where a discussion of the status of the minimality conditions (11) and (20) becomes finally unavoidable. There are two options that I would like to consider. We could say that these conditions are found specified in the lexical entry for the various exceptive prepositions. Whether a type of EP only occurs with universal determiners or also with cardinal determiners is then determined by the lexicon. The other strategy would be to generalize the more liberal condition to hold of all EPs and to derive the stricter behavior of *but* and *except* in some other way. I am leaning towards the lexical approach without finally committing myself, thus leaving space for further research or at least deliberation.

4. Free EPs

The other phenomenon that needs addressing is the fact that EPs do not always occur NP-internally but can appear in extraposed or free positions. Among such EPs we have to distinguish further between the ones that are used purely appositively and those that are as essential for the truth of a sentence as the NP-connected ones discussed above.

Appositive EPs occur in sentences that make an almost universal claim and state the exceptions that make it impossible to make a truly universal claim. Thus, they are not constrained by any minimality condition. This is the state of affairs in cases like (23).

---

[9] In fact it predicts that they should occur with any precise determiner. Note that in (19) the determiners have their exactly-reading and not the at least-reading. At least-cardinals can occur with EPs, but then the EP is merely appositive, rather like in (23).

(23) Almost all politicians are corrupt, except for the socialists.
I will not develop an account of such appositive uses of EPs but do not anticipate serious problems from this direction.

In (24), however, the EP, although free, is used to make the sentence true.
(24) Except for John, every student has a car.
Since these EPs do not attach inside the NP, we have to somehow connect them with our N'-semantics developed earlier. First, however, some arguments against a different approach.

Hoeksema (ms) suggests to treat free EPs as S-level modifiers, again employing the strategy of making them subtract entities from the domain of entities, this time for the evaluation of the whole sentence. The exact formulation is given in (25).
(25) $\|$*Except c, S*$\|_E$ *= True iff (*$\|S\|_{E-\|c\|}$ *= True,* $\|S\|_E$ *= False & S is additive).* [10]

The additivity requirement is added to capture the co-occurrence restrictions that continue to hold for free EPs in the same way as for connected EPs. But it turns out that "being an S that contains a universal determiner" and "being an additive S" are not the same property. Thus we encounter problems with sentences that are not additive but do allow free EPs. A crucial example is (26).
(26) Somebody damaged every car, except for this Cadillac.
This sentence is of course perfect, but it is not additive. Hoeksema has no explanation except to suggest that "it must be due to scope factors" (9). But the sentence seems to be fine with both possible scopings. The next sentence (27), also not additive, is also fine.
(27) Except for me, everybody knows somebody who knows everybody.
Hoeksema notes that "perhaps additivity is not computed all the way through" (10).

Even if the additivity problem is solvable by further assumptions, (27) raises a more devastating problem. In the model manipulation approach the exception is taken out of the domain of entities for the computation of the meaning of the whole sentence. Now, for (27) this means that not only do I not know anybody who knows everybody but I am also not necessarily known by any of the somebodies who know everybody, which for me is not the right reading. This may be too complex to judge, but the same point is made more obvious by (28), given to me by Hotze Rullmann.
(28) Except for John, everybody loves John.
This sentence would be in fact uninterpretable if John were removed from the domain of entities for the interpretation of the whole sentence.

I would like to suggest instead that free EPs specify the free variable that we assumed above. This implies in particular that they semantically associate with an NP in their host sentence and not with the whole sentence.[11] This has to be generalized even further when we deal with cases where the surface sentence contains no universal quantifier as in (29) or (30) (examples from Hoeksema).
(29) Bob smokes, except in the car.
(30) Except for Thursday night, Bob is free.
We need a mechanism that unifies EP with some variable after the hidden universal quantification has been made visible by some previous interpretation routine. In so far as this is a generalization of the free variable approach, I consider it to be still superior to

[10] Additivity for sentences is defined as follows: $\|S\|$ is additive iff (if $\|S\|_E$ = 1 and $\|S\|_{E'}$ = 1 then $\|S\|_{E \cup E'}$ = 1).

[11] There is probably no prohibition against them specifying more than one free variable so that sometimes we may get the semblance of them subtracting entities across a whole sentence.

Hoeksema's more brute-force strategy of taking entities, or places, or times out of the model of evaluation.

Before pointing out some open questions, I would like to conclude that my proposal in this paper predicts the main properties of EPs to a larger degree than previous approaches and without many implausible stipulations.

5. Problems and open questions

Quite apart from the problems that became apparent in the above proposals, there remain a number of further open questions.

Calling for a rather technical solution (already sketched by Hoeksema) is the fact that full generalized quantifiers can occur in the EP instead of the proper names that I used in this paper. We have cases such as (31) and (32).

(31) All the first-year students except the foreigners attended.

(32) All the marbles except for at most seven red ones.

If an NP denotes a principal filter as in (31) we can lower it onto its generator set via the lower-operation of Partee (1987). The NP in (32) probably has to be interpreted as an e-type NP anyway.

A topic that merits further research is the curious fact that a number of languages express the meaning of English *only* by employing constructions with EPs. This is illustrated in (33) for Classical Arabic and in (34) for Modern Irish (from Chung & McCloskey 1987).[12]

(33) Mā qāma 'illā Zaydun.
NEG rise:PERF but Zayd-NOM
'Only Zayd rose'

(34) Ní tháinig ach beirt
NEG come(PAST) but two-people
'Only two people came'

At this point, I have nothing to say about such cases.

[12] Eloise Jelinek (p.c.) informed me that similar constructions are found in Appalachian English.

References

AZOULAY-VICENTE, Avigail. 1988. "La Syntaxe de *ne...que*". Lingvisticæ Investigationes. 12(2). 205-233.
BACH, Emmon. COOPER, Robin. 1978. "The NP-S Analysis of Relative Clauses and Compositional Semantics". Linguistics and Philosophy. 2. 145-150.
BARWISE, John. COOPER, Robin. 1981. "Generalized Quantifiers and Natural Language". Linguistics and Philosophy. 4. 159-219.
CHUNG, Sandra. MCCLOSKEY, James. 1987. "Government, Barriers, and Small Clauses in Modern Irish". Linguistic Inquiry. 18. 173-237.
HOEKSEMA, Jack. ms. "The Logic of Exception". University of Pennsylvania.
KEENAN, Edward L. STAVI, J. 1986. "A Semantic Characterization of Natural Language Determiners". Linguistics and Philosophy. 9. 253-326.
LANDMAN, Fred. MOERDIJK, Ieke. 1979. "*Behalve* als voorzetsel". Spektator: Tijdschrift voor Neerlandistiek. 9. 335-347.
LAPPIN, Shalom. 1988. "The Semantics of 'Many' as a Weak Determiner". Linguistics. 26. 977-998.
PARTEE, Barbara H. 1987. "Noun Phrase Interpretation and Type-Shifting Principles". In: GROENENDIJK, J. DE JONGH, D. STOKHOF, M. (eds.) Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers. 115-143.
STOCKWELL, R.P. SCHACHTER, P. PARTEE, B.H. 1973. The Major Syntactic Structures of English.
VAN BENTHEM, Johan. 1984. "Questions about Quantifiers". Journal of Symbolic Logic. 49. 443-466.
WRIGHT, William. 1964. A Grammar of the Arabic Language.

Amherst, November 1989

Kai von Fintel
Department of Linguistics
South College
University of Massachusetts
Amherst, MA 01002

[fintel@umass.bitnet]