Abstract

This paper introduces a general semantic procedure for the interpretation of sentences in English which contain a pronoun appearing linearly before its antecedent. This phenomenon is known as backward anaphora or Kataphora. We concentrate on cases of Kataphoric Quantification: sentences where the antecedent is interpreted as a generalized quantifier. A well-known case of crossed kataphoric and anaphoric quantificational links are the Bach-Peters' (BP) sentences. We consider a wide range of singular and plural BP sentences and other kataphoric constructions and show that in English there are two kinds of kataphora: S (set) Kataphora, in which the pronoun is interpreted as a set determined by the antecedent, and Q (quantifier) Kataphora, where the pronoun can be interpreted as a variable bound by its antecedent. Only the later can be considered a genuine case of Kataphoric Quantification. We show that S-kataphora appears only with plural and definite NPs, while Q-kataphora appears both in singular and in plural NPs but only with a limited class of determiners.

The distribution of various determiners that allow Q-kataphora readings is shown to be determined by the weak/strong distinction between quantifiers, which is known to affect the availability of similar readings in forward anaphora.

A successful procedure for the interpretation of certain BP sentences was presented by Higginbotham and May in their article "Questions, Quantifiers and Crossing" (1981). This work introduced an operator of Quantifier Absorption which produces a binary generalized quantifier from a pair of unary generalized quantifiers. We show that H&M's original definition does not yield the correct truth conditions for several plural BP sentences. H&M's definition also does not fully specify the way the absorbed quantifier applies to its arguments. We propose a generalization to H&M's operator which deals properly with these problems. The revised procedure of absorption yields correct truth conditions also for non-BP sentences with kataphoric quantification. This shows for the first time a general motivation for quantifier absorption, that may explain the complicated truth conditions of BP sentences as a special case.

1 Introduction

In most artificial logical languages with quantificational operators, variable binding operates in one, constant, direction, normally from left to right. Natural languages, on the other hand, provide a rather infrequent construction which can be interpreted as binding in the opposite direction: a quantifier binds a variable that appears linearly before it in the sentence. A famous example for such a case of "backward binding" are the Bach-Peters' (BP) sentences. For instance:

(i) *Every pilot who saw it hit some mig that chased him.*

*The part of the second author was partially funded by a grant from the Israeli Ministry of Science and by the fund for the promotion of research in the Technion.*
In one reading under a quite subtle judgement, which is widely accepted in the literature, the relation between the underlined NP’s in (i) can be considered, intuitively speaking, as a case where a variable, the translation for *it*, is bound by a quantifier, the translation for *some mig that chased him*. If this judgement is correct, as many speakers seem to agree, then this is a case where a quantified NP “binds backward” a pronoun that linearly precedes it. Throughout this paper we use the term *Kataphoric Quantification* to refer to this phenomenon. The interest in this kind of reading for sentences like (i) emerged from the inter-relations between this kataphoric link and the forward anaphoric link between the NP’s which are given in (i) in italics.

In this paper we aim at a closer look at backward quantification. Our interest is both in its distribution and in the ways it is logically interpreted. We consider three classes of additional linguistic data:

- **Singular BP sentences** with various determiners in the matrix verb’s object. For example:

(i) Every teacher who liked him praised \{one, the, *no, ? every\} student who admired her.

- **Plural BP sentences**. For example:

(ii) Most officers who knew them instructed \{some, most\} all soldiers who watched them.

- **Plural non-BP sentences** with kataphoric quantification. For example:

(iv) Some clerks who needed them used \{some, most\} all \{all, no, two\} folders.

The empirical status of these three classes of examples is hard to be definitely determined. This is because different speakers have notably different judgements with respect to their grammaticality. However, many speakers accept all the judgements above and no speaker was found to consider as grammatical sentences or readings marked with * or ?, above and henceforth in this paper. For this reason we adopt the most liberal judgements we got as the relevant for the study of kataphoric quantification. We hope that further research may determine why are judgements with respect to this kind of readings so diverse among English speakers.

In order to present our account of Kataphoric Quantification we use two important ideas from previous works on anaphora and quantification.

The first idea, following works of Evans ([Evans '80]), Reinhart ([Reinhart '86], [Reinhart '89]) and many others, is the distinction between two kinds of anaphoric links between plural pronouns and plural quantified NP’s:

- **Bound variable anaphora**. As in:
The first kind of anaphora is the same as the bound variable anaphora with singular NP’s. The second kind appears especially with plural and definite NP’s. In this kind of anaphora there is no simple way to represent the pronoun as a bound variable. For instance, in (vi) such a representation would mean that no teacher was gone, which is the contrary to what the second sentence in (vi) implies. A similar distinction between two kinds of anaphora is traditional also in infra-sentential anaphora as in plural “donkey” sentences (see the above references for further details). We will claim that different readings of plural kataphoric constructions are no more than a manifestation of the above distinction between quantificational and set anaphora. This will allow us to present a characterization for the class of quantifiers that permit kataphoric quantification and to relativize this classification to well-known distinctions considering the different anaphoric links with weak and strong NP’s.

Another central idea on which we elaborate is Higginbotham and May’s proposed procedure of Quantifier Absorption. This procedure is intended to provide general (different) operators for the interpretation of BP sentences and multiple Wh-Questions. These operators are functions that attribute a generalized binary quantifier to any two generalized unary quantifiers. We will show that Higginbotham and May’s absorption operator for BP sentences yields wrong truth conditions for certain plural BP sentences as in (iii). We will also claim that H&M’s notion of absorption does not allow an interpretation for other kataphoric constructions as in (iv). We then propose a generalized procedure for absorption, which includes H&M’s proposal as a special case for the class of examples discussed in their work. This revised procedure also yields the desired truth conditions for sentences in the form of (iii) and (iv). This suggests a broader motivation for quantifier absorption, which is shown here to be a general procedure yielding the complicated truth conditions of a wider range of sentences with kataphoric quantification, in addition to the BP sentences discussed by H&M.

The paper is organized as follows. Section 2 provides the essential linguistic data concerning kataphoric quantification and focuses on the distinction between Quantificational kataphora and “Set” kataphora. In section 3 we propose a characterization for the class of determiners that allow kataphoric quantification and show the connections between this class and the larger class of weak determiners. This provides a principled explanation for the distribution of various determiners in sentences with kataphoric quantification. In section 4 we present Higginbotham and May’s proposal and its drawbacks with respect to plural BP sentences and other kataphoric constructions. We show how a revision in H&M’s definitions provides a solution for these problems. We conclude with a discussion of the remaining open problems for the generality of absorption.

Appendix A contains some remarks on “Set” anaphora. Appendix B provides a proof for the correctness of interpretation using the revised absorption operator. Appendix C shows some problems for the analysis of kataphoric quantification using two versions of DRT and Dynamic Logic which were introduced to handle with this problem.

2 Kataphoric Quantification in English

2.1 The scope of discussion

By the name Kataphoric Quantification (backward quantification) we refer to sentences in English where a Quantified NP (QNP) can be interpreted as binding a pronoun that appears before it in
the sentence. Here are some examples from plural kataphora:

1. Every boy who liked *them* played with *some girls*.
2. Most teachers who wanted to see *them* invited (at least) five students.

Another (more complex) case which exhibits kataphoric quantification are the famous and puzzling Bach-Peters’ sentences (henceforth “BP sentences”):

3. [Every pilot who saw *it*; j] hit [some mig that chased him; i].
4. [Many women who like *them*; i] bring presents to [their; children; j].

Roughly, in both (1)-(2) and (3)-(4) the first pronoun in *italics* can be interpreted as a bound variable. The quantifier binding this variable is the representation of the matrix verb’s object NP. At Surface Structure, the pronoun is linearly placed before its antecedent hence the term *Kataphoric Quantification*. We henceforth persistently ignore other readings with non-restrictive relative clauses or with deictic interpretation of pronouns. Readings with object wide scope are discussed in sub-section 2.4.

Kataphoric quantification is possible in simple sentences like (1) or (2), where the pronoun is in plural form. Singular instances of kataphoric quantification include (presumably only) the well-known type of Bach-Peters’ sentences similar to (3). The same kind of “crossed binding” readings appear in the less investigated case of plural BP sentences exemplified by (4).

Sentences parallel to (1) or (2) with singular NP’s are often considered unacceptable in English (but see [Koster ‘91]). For example:

5. ? Every boy who liked *her* played with *some girl*.

This fact is usually attributed in GB theory to a weak crossover violation. In this paper we aim to deal only with the semantic problems that are raised by kataphoric quantification. For this reason we will not address here the question of why is it that sentences like (5) are considered ungrammatical by most speakers while (1) and similar sentences are OK for most speakers. See for example [May ‘85] or [Jacobson ’92] for a discussion of this point.

In order to concentrate on the interaction between kataphora and quantification, we do not consider examples of kataphora in discourse, where the pronoun and antecedent are in two different sentences. We tend to think that these cases are always analyzable as cases of coreference, and not as genuine quantification. For example:

6. I saw *them* helping the students. *These teachers* are wonderful.

An evidence which supports the classification of (6) as a case of coreference is that it is not known of any similar examples from English discourse where the antecedent in a “discourse kataphora” is a QNP, such as “every man”, “some students”, etc.

### 2.2 Two kinds of kataphora in plurals

Kataphoric sentences like (1), (2) or (4) with plural pronouns seem to involve two different readings. In order to see this consider the contrast between (1) and (7):

---

1. Henceforth we omit indices whenever possible and use *italics* to denote the pronoun in focus and its antecedent. When needed, we use the indices i and j, assuming i ≠ j.
2. Obviously, we refer only to the semantic content of the term ‘antecedent’.
(1) Every boy who liked them played with some girls.

(7) Every boy who liked them played with no girls.

As a kataphor, the pronoun them in (1) can be interpreted either as a constant set of girls or as a set which may vary for each boy quantified over. This fact was verified by asking our informants the following question: “Can them be different for each boy?” Most native speakers that were asked tend to answer this question affirmatively. By contrast, with respect to (7) the same question was persistently given a negative answer. We may express this contrast by tentatively assuming that while the kataphora in (1) is ambiguous between the paraphrases (1’a) and (1’b), sentence (7) only has the meaning of (7’):

(1’) a. Every boy who liked some girls played with some girls that he liked.

   b. Every boy who liked the girls played with some of the girls.

   (the set referred to by the girls is some salient, context depended and fixed set)

(7’) Every boy who liked the girls played with none of the girls.

More reason to assume these specific interpretations will be given in the discussion that follows. Similar observations with respect to plural NP’s with various determiners were discussed extensively in [Reinhart ’86] and [Reinhart ’89]. We get back to some problems addressed by Reinhart in the following sections.

What is the difference between determiners like some and no, that allows reading (1’a) in (1) and does not allow a parallel reading in (7)? This question is one of this paper’s main concerns. At this stage, however, it is only important to recognize that sentences like (1) admit two kataphoric readings.

We henceforth use the ad-hoc names Q-Kataphora (“quantifier” kataphora) for the ‘a’ reading of (1), and S-Kataphora (“set” kataphora) for the ‘b’ reading. These specific names are appropriate because in reading ‘a’ the interpretation of the pronoun them involves the whole QNP antecedent, whereas in reading ‘b’ only the set term referred to by the common noun girls is needed for the interpretation of the pronoun. The two readings are closely related to some well-known phenomena in (forward) anaphora. We refer to these connections in sub-section 2.5.

The same kind of readings as in (1) appear also with other determiners in the antecedent. For example:

(8) Every pilot who saw them hit (at least) three migs.

In the S-kataphora reading the pronoun them in (8) refers to a fixed set of migs. The sentence claims that every pilot who saw the migs in this set hit three migs of this set. There is also a (logically weaker) Q-kataphora reading in which the set of migs that the pilot saw may vary for each pilot quantified over. In this reading the sentence claims that every pilot who saw at least three migs hit at least three of the migs that he saw.

An interesting contrast is that between (9a) and (9b):

(9) a. Every pilot who saw them hit more than two migs.

   b. Every pilot who saw them hit less than two migs.

While most informants consider (9a) to be logically equivalent to (8), having two readings with the fixed/varying distinction between the set interpretation of the pronoun, sentence (9b) is judged to have only the S-kataphora reading, similar to (7), where the pronoun is interpreted as a fixed set.

Some more cases of plural kataphora with judgements on Q and S kataphora are summarized in (10):

5
(10) Every pilot who saw them hit \( \{ \text{some, several } Q, S \}
\text{all } Q=S
\text{no } S
\text{(at least) } n Q, S
\text{less than } n S
\text{more than } n Q, S
\text{exactly } n Q?, S
\text{between } n \text{ and } m Q?, S
\text{many } Q?, S
\text{few } S
\text{most } S
\text{more than } n\% \text{ of the } S
\text{less than } n\% \text{ of the } S
\text{all but } n S \} \text{ migs.}

Notice that in the case of the universal determiner all the two reading are hardly distinguishable and can be considered equivalent. We will get back to this point in section 3.

Some readers might find sentences like

(11) Every pilot who saw them hit more than 30% of the migs.

close in their interpretation to sentences like

(12) Every pilot who saw them hit more than 30 migs.

But after some reflection on their meaning one realizes that while (12) can be interpreted as a genuine case of Q-kataphora (them can be interpreted as more than 30 migs), in (11) the pronoun them can be interpreted only as some vague fixed set of the migs.

2.3 BP sentences and singular kataphoric quantification

With respect to the distinction between Q and S kataphora, plural BP sentences admit similar readings to the more simple plural kataphoric constructions discussed so far. For example, compare (13) and (14):

(13) Every pilot who saw them hit some migs that chased him.

(14) Every pilot who saw them hit no migs that chased him.

In (13), similarly to (1), the pronoun them may be interpreted for every pilot as representing the set of migs that chased the pilot (S-kataphora), or as some of these migs (Q-kataphora). In (14), as in (7), only the S-kataphora reading is available.

We may substitute in (13) (or in (14)) also other determiners from (10), to realize that plural DP sentences do not differ substantially from other kataphoric constructions with respect to the availability of the two readings discussed.

Evidently, the ambiguity in plural kataphora we discussed does not exist in the singular kataphoric case of DP sentences like (3), repeated here:

(3) [ Every pilot who saw it; ] hit [ some mig that chased him; ].

We would like to claim that the kataphoric reading in singular DP sentences should be classified as Q-kataphora. There are two arguments for this classification:
1. Determiners that do not allow Q-kataphora in plurals are not possible at all in singular BP sentences. For example:

\[ \text{(15) } [\text{ Every pilot who saw it }_i ; \text{ hit } [\text{ no mig that chased him }_i ]] \]

2. Forward “set anaphora” (“common-noun” anaphora) is a phenomenon that is restricted to plural anaphora (and definite NP’s, as we claim in subsection 2.6). It does not appear in singular anaphora with a QNP antecedent. For instance compare the following examples:

\[ \text{(16) We saw no student. } * \text{ He was gone.} \]

\[ \text{(17) We saw no student. } ? \text{ They were gone.} \]

\[ \text{(18) We saw no students. } \text{ They were gone.} \]

These two arguments support the claim that Q-kataphora is a phenomenon the distribution of which is controlled by differences between determiners in the antecedent QNP, with no respect to the number of the pronoun or antecedent. S-kataphora, on the other hand, appears with all determiners in plural kataphora and do not show up in singular kataphora with a QNP antecedent. In subsection 2.5 we will claim that S-kataphora is just a case of what is known as “set anaphora” in plurals.

A summary of the judgements for some determiners in singular BP sentences follows:

\[ \text{(19) Every pilot who saw it hit } \begin{cases} \text{some / one} \\ ?\text{every / ?each} \end{cases} \text{ mig that chased him.} \]

The possibility of an indefinite article in the antecedent NP will not be discussed in this paper. Cases of an antecedent NP with a definite article will be discussed in subsection 2.6.

2.4 Object wide scope readings

It might be tempting to classify at least some of the data discussed so far as a matter of object wide scope interpretation. In some respect, such an approach is proposed in [van Deemter '90] (see a summary and a discussion of van Deemter’s proposal in appendix D). However, there are some facts that show that this alternative is untenable. Reconsider sentence (1): in an object wide scope reading, it can be paraphrased as:

\[ (1'c) \text{ There are some girls s.t. every boy who liked them played with them.} \]

It is quite evident that (1’c) is not equivalent to (1’a) neither to (1’b); (1’c) entails that there is a set of girls A s.t. if a boy liked each member of A he also played with each member of A. However, (1) also can be interpreted as true when there is no such set and each boy liked and played with a different set of girls.

From another respect: BP sentences, if allow for object wide scope reading, are to be analyzed as ambiguous between a wide scope and a narrow scope Q-kataphora reading. So (3) actually can be interpreted also like the following different BP sentence:

\[ (3') \text{ There is } [\text{ some mig that chased him }_i ; \text{ that was hit by } [\text{ every pilot who saw it }_i ]_i . \]

It is hard to determine whether BP sentences allow for such wide scope readings. However, it is quite clear, that if such readings exist, they must diverge from the Q-kataphora and S-kataphora readings that were discussed.
2.5 Q and S Kataphora and forward anaphora

The readings for kataphoric sentences we referred to as Q and S kataphora are intimately related to other well-known ways of anaphoric links. “Set anaphora” (sometimes called “common-noun anaphora”) is easy to find in plural discourse anaphora as in (18), or in plural “donkey” sentences like:

(20) Every child who saw less than three chocolate bars wanted to eat them.

To become convinced that the pronoun them in (20) can refer to more than three chocolate bars it is convenient to consider the following variation:

(21) Every child who saw less than three chocolate bars wanted to eat all of them.

In the more accessible reading of (21) the pronoun them is interpreted as referring to some fixed set of chocolate bars (that normally is interpreted as including at least three items). It is quite tempting to try to attribute this “set anaphora” in (21) to an instance of simple coreference, where some salient context dependent set is assigned as the interpretation of the pronoun them. But an analysis along these lines would be over-simplistic. Consider for example:

(22) Every child who saw at least three of his brother’s chocolate bars wanted to eat all of them.

In (22), the pronoun them stands for a different set of chocolate bars for each child quantified over (the set of chocolate bars that belong to the child’s brother).

Similar observations hold with respect to S-kataphora in plural BP sentences. For example:

(23) Every pilot who saw them hit no migs that chased him.

Here also, the set that the pronoun them represents varies for each pilot, and includes only migs that chased the pilot.

We may conclude that in order to tell between S-kataphora and Q-kataphora in the general case, it is not sufficient to use the fixed/varying set distinction we used. In S-kataphora also, the set represented by the pronoun can vary when the antecedent contains a bound variable, so the set it forms for the anaphoric link is not fixed. However, when this set is fixed, as in the non-BP sentences (1) and (2), only Q-kataphora admits a varying set reading to the pronoun, so the test we have used is relevant for this kind of sentences.

Whatever analysis we propose for S-kataphora, it is clear now that this phenomenon is not related specifically to the kataphoric construction; it is just a special case of “set anaphora”, which is generally a possible anaphoric link between plural QNP’s and plural pronouns, and maybe also between singular or plural definite NP’s and pronouns (see sub-section 2.6). The discussion in [Heim ’90] brings some additional important points which are related to this issue. We tend to believe that “set anaphora” is a special case of coreference, which is better captured in the kind of “E-type” or “lazy pronoun” analyses that Heim discusses. However, this is not the main issue of the present work, so only a short discussion of the problem is included in appendix A.

Something rather different is going on with Q-kataphora. Here, no simple similarity exists between kataphoric sentences and their (forward) anaphoric parallels. Compare for example the Q-kataphora reading in (24) to the donkey sentence (25):

(24) Every pilot who saw them hit more than two migs.

(25) Every pilot who saw more than two migs hit them.
As was well observed in the literature (e.g. [Evans ’80], [Heim ’90], [Reinhart ’86], [Reinhart ’89]),
the sentence (25) has the following reading:

(25’) Every pilot who saw more than two migs hit all the migs that he saw.

In this reading, the pronoun *them* in (25) is classified by Evans as what he calls an “E-type” pronoun.
In order to interpret such a pronoun we cannot use only a substitution of its antecedent; we need also the predicate that applies to the antecedent (in (25) it is the two-place predicate *saw*).
In addition, an “E-type” reading for a pronoun requires some “universal force”: the pronoun *them* is not interpreted as *two migs that the pilot saw* but as *all the migs that the pilot saw*.

Definitely, there is no completely parallel reading in sentence (24). This sentence cannot mean:

(26) * Every pilot who saw all the migs that he hit, hit more than two migs.

The sentence is agreed to mean something entirely different:

(24’) Every pilot who saw more than two migs, hit more than two migs that he saw.

In section 4 we will discuss in detail the Q-kataphora readings for kataphoric constructions and BP sentences. At the moment our point is only to show that this kind of reading is not exactly the well-known “E-type” reading for pronouns. A short discussion of the relations between the interpretation of BP sentences and E-type pronouns may be found in [Evans ’77] (although we are not sure our informants’ judgement agree with all of Evans’s judgements).

However, in many senses these two kinds of anaphora resemble: interestingly, as witnessed by (26), the pronoun *them* in (24) cannot be replaced by a representation of its antecedent with the predicate applying to the antecedent, as in “E-type” pronouns. It is the interpretation of the antecedent that is affected by the predicate applying to the pronoun. The antecedent (the object of *hit* in (24)) should be interpreted using the predicate applying to the pronoun *them* (*saw*). In addition, no universal quantifier is needed for the interpretation of neither the pronoun nor the antecedent.

We conclude that Q-kataphora is related, but certainly not identical to the E-type anaphora as in Evans’s classification. What this relation might be will be proposed in section 3.

### 2.6 Definite descriptions in kataphora

A rather common case of kataphora, which has not yet been discussed here, is the case of an antecedent with a definite article. Consider:

(27) Every pilot who saw *them* hit the migs.

(28) Every pilot who saw *it* hit the mig that chased him.

(29) Every pilot who saw *them* hit the migs that chased him.

Such sentences are not easy to classify using the distinction we used between S-kataphora and Q-kataphora. Consider (27) for example: since the antecedent (*the migs*) represents a fixed set on the first place, the pronoun *them* stands for the same set of migs for all the pilots quantified.

---

3The use of the term “E-type” in the literature is quite confusing: sometimes (as in [Evans ’80]) it refers to a descriptive observation on readings of pronouns. Sometimes (as in [Heim ’90]) it refers to a certain kind of theories for anaphora resolution. For this reason we will use the term only in a rough fashion, in most places in its “descriptive” use.

4or a generalized quantifier which is the principal filter of a fixed set.
Table 1: Acceptability of determiners in Q-kataphora

<table>
<thead>
<tr>
<th>Determiner</th>
<th>Plural</th>
<th>Singular</th>
</tr>
</thead>
<tbody>
<tr>
<td>some/ several</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>all/every/each</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td>no</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>one</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(at least) n</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>more than n</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>less than n</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>exactly n</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>between n and m</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>(at least) n%</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>more/less than n%</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>most</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>few</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>many</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>all but n</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Over, so in contrast to similar sentences like (1), we cannot use the fixed/varying set distinction we used to decide between Q-kataphora and S-kataphora.

However, it seems plausible to classify “definite kataphora” as S-kataphora. The reason for this classification is that such a distinction allows a clear conception of the distribution of “set anaphora” in general—S-kataphora appears whenever a set-term is formed by the antecedent. This happens in the cases of plural and definite NP’s. Such a classification also allows an analysis of definite NP’s as representing set-terms, and not as generalized quantifiers. See [Loebner ’87] for a detailed argumentation in favour of such an analysis.

2.7 A summary of linguistic data

We may summarize the phenomena that were discussed and the distinctions we made as follows:

1. Two kinds of kataphoric links in English are:
   (a) Q-Kataphora: where the quantifier in the antecedent is part of the interpretation of the pronoun.
   (b) S-Kataphora: where the pronoun is interpreted using only a set-term that is formed by the interpretation of the antecedent.

2. Q-kataphora is similar, but not identical, to the “E-type” readings of pronouns (cf. [Evans ’80]). S-kataphora is identical to “set anaphora” that appears in the cases of (forward) plural anaphora and anaphora with definite NP’s.

3. Consequently: Q-kataphora appears both in plural and in singular(BP) kataphora. S-kataphora may appear with all determiners, but only in plurals or with definite antecedents.

4. Only a restricted set of determiners, summarized in table 1, allows for Q-kataphora in plural and in singular sentences.
We tend to believe that S-kataphora is to be accounted for differently than Q-kataphora. It seems to us that a discourse analysis, possibly in the lines of [Heim '90], is to be involved in the interpretation of pronouns in set anaphora. We provide a note on this possible analysis in appendix A.

In the rest of this paper we concentrate on the problem of BP sentences and other cases of Q-kataphora, which are a genuine case of kataphoric quantification. Two different questions will be addressed:

1. What is a proper characterization for the class of determiners that allow Q-kataphora? What can be the reason for this specific distribution?

2. What is a proper general semantic procedure for the interpretation of Q-kataphora?

An answer to the first question may exemplify once again the relevance of the mathematical features of various generalized quantifiers to the distribution of determiners in quantificational sentences.

An answer to the second question may introduce for the first time a general mechanism for the interpretation of BP sentences, which is responsible also for the interpretation of other forms of backward quantification in natural language.

3 Cardinality as the Major Restriction for Q-Kataphora

The linguistic facts brought in section 2 raised the following question: Why is Q-kataphora restricted only to certain plural and singular determiners in the antecedent? In this section we try to provide a proper characterization for the set of determiners that allow Q-kataphora. We propose that the restrictions that prevent Q-kataphora with other determiners is a manifestation of a well-recognized fact: the distinction between weak and strong determiners when they appear in (forward) anaphora such as donkey-sentences and discourse anaphora.

We start by a summary of definitions for Cardinality Quantifiers (CQ’s). Then we use these definitions in our proposed characterization of the set of determiners that allow Q-kataphora: the Kataphoric Quantification Constraint (KQC). Next, we propose a universal on the identity of cardinality quantifiers and weak quantifiers in natural languages, which is implied by several previous works. Using this universal it is shown that KQC is actually predicted by more general principles on the anaphoric behaviour of QNP’s with weak determiners.

3.1 Cardinality Quantifiers

An unary generalized quantifier \( \mathcal{Q} \) is called a cardinality quantifier if its definition is a certain condition on the cardinality of the intersection set of its arguments, as explained below.

For example: to say that two men sneeze is to say that the intersection of the set of men and set of sneezers includes at least two members. By contrast: to say that every man sneezes is no statement on the cardinality of this intersection set.

Formally:

An unary quantifier \( \mathcal{Q} \) over a finite universe \( U \) \(^6\) is in the class of Cardinality Quantifiers (CQ) iff there are two constant integers \( n \) and \( m \), \( 0 \leq n < m \) (with \( m \) possibly \( \infty \)) s.t.

\[^5\]An unary generalized quantifier \( \mathcal{Q} \) over a universe/domain \( U \) is a function \( \text{Pow}(U) \times \text{Pow}(U) \rightarrow \{0,1\} \). We say that \( \mathcal{Q}(A,B) \) holds for \( A,B \subseteq U \) if \( \text{Q}(A,B) = 1 \).

\[^6\]we henceforth do not mention the universe \( U \) whenever it is obvious from the context
for every two sets $A$ and $B$ it is true that:

$$Q(A,B) \text{ holds iff } n \leq |A \cap B| < m$$

We make the following distinctions between certain sub-classes that constitute CQ:

1. *Increasing* CQ (CQ$\uparrow$): $n > 0$ and $m = \infty$
2. *Decreasing* CQ (CQ$\downarrow$): $n = 0$ and $m < \infty$
3. *Mixed* CQ (CQ$\sim$): $n > 0$ and $m < \infty$
4. A special case: The “null” member of CQ: $n = 0$ and $m = \infty$

Here are some examples for determiners in English with the type of the QNP they form, with respect to the distinctions above:

- $CQ\uparrow$: *more than three* ($n=4$), *(at least) five* ($n=5$), *some* ($n=1$)
- $CQ\downarrow$: *less than three* ($m=3$), *at most five* ($m=6$), *no or exactly zero* ($m=1$)
- $CQ\sim$: *between two and six* ($n=2,m=7$), *exactly four* ($n=4,m=5$)
- The null CQ: *at least zero* $^7$
- Not in CQ: *every, most, all but five, more than 30%*

The following properties of CQ can be easily verified:

**Fact 1:** All the members of CQ$\uparrow$ and CQ$\downarrow$ are right monotone increasing ($mon\uparrow$) and right monotone decreasing ($mon\downarrow$), respectively $^8$. All the members of CQ$\sim$ are not monotone. The only member of CQ which is both $mon\uparrow$ and $mon\downarrow$ is the null CQ.

**Fact 2:** Every quantifier $Q$ in CQ$\sim$ is definable as conjunction of quantifiers from CQ$\uparrow$ and CQ$\downarrow$: $Q(A,B)$ holds iff both $Q_1(A,B)$ and $Q_2(A,B)$ hold, where $Q_1$ and $Q_2$ are quantifiers in CQ$\uparrow$ and CQ$\downarrow$, respectively.

The notion of cardinality quantifiers will help us to characterize the distribution of determiners in Q-kataphora.

### 3.2 The Kataphoric Quantification Constraint

We propose the following characterization for the results summarized in table 1:

**Kataphoric Quantification Constraint (KQC):** The availability of Q-kataphora readings varies as follows with respect to the interpretation of the antecedent QNP, representing the quantifier $Q$:

- Q-kataphora is *easy* when $Q$ is in CQ$\uparrow$.

$^7$Notice that the null CQ is the only member of CQ which is not a sieve, in the sense of [Barwise & Cooper '81]. Therefore, a formula of the form at least zero $Q(A,B)$ is automatically valid (a tautology) with no respect to the identity of $A$ and $B$. For this reason sentences like “at least zero students appeared” convey null content, although they are completely grammatical.

$^8$A quantifier $Q$ is in $mon\uparrow/mon\downarrow$ iff whenever $Q(A,B)$ holds then for every set $B$ s.t. $B \supseteq B$ / $B \subseteq B$, respectively, $Q(A,B')$ also holds.
Q-kataphora is *hard* when Q is in CQ~.

Q-kataphora is *impossible* when Q is in CQ| or when Q is not in CQ.

If we accept the empirical data in table 1, KQC characterizes properly the distribution of Q-kataphora in singular and plural sentences. Determiners in CQ| such as *some*, *more than three*, *at least four* allow easily for Q-kataphora readings. Q-kataphora is possible, but hard, when the antecedent includes a determiner in CQ~ like *exactly four* or *between three and five*. Q-kataphora is impossible with determiners that are in CQ| (*no*, *less than four*) or not in CQ at all (*every*, *all*, *more*/*less than 30%*, *all but two*).

The case of the determiners *many* and *few* is subtle: these determiners are sometimes interpreted in the sense of *more*/*less than n*, where n is a vague constant determined by the context. This is a CQ reading of these determiners. On the other hand, they are sometimes interpreted as the "proportional" determiners *more*/*less than n %*, which are not cardinal. KQC then predicts that *many* and *few* are sometimes possible in Q-kataphora and sometimes not. This agrees with their questioned status as antecedents in Q-kataphora, as brought in table 1.

The prediction of KQC that the universal determiners *every*, *each* and *all*, which are not in CQ, do not allow Q-kataphora readings is not self-evident, and deserves a closer examination. We must start here a short digression from our main theme to show that this prediction is harmless. For the singular case this prediction agrees with the intuitions of our informants (summarized in (19)), who consider BP sentences like the following unacceptable:

(30) ? Every pilot who saw it hit *every/each mig that chased him.*

However, one may tend to think that plural kataphora with *all* in the antecedent are ambiguous between Q and S kataphora readings. Consider for example:

(31) D pilot(s) who saw *them* hit *all migs*.

Sentence (31) might be considered ambiguous between the following paraphrases:

(31’) a. D pilot(s) who saw all migs hit all migs that he saw. (Q-kataphora)
   b. D pilot(s) who saw the migs hit all migs. (S-kataphora)

But, the plausible assumption that in a given context the NP *the migs* refers to a set which includes exactly all the individuals quantified over in the QNP *all migs* leads to equivalent formalizations of the alleged two readings. In appendix B this claim is proved to be correct for all the determiners that *live on* their arguments.\(^1^0\)

For example, consider the determiner D = *every*. Let the one- and two-place predicates \(p(x), m(y), s(x,y),\) and \(h(x,y)\) stand for the expressions *pilot*, *mig*, *saw*, and *hit*, respectively. We may formalize the two sentences in (31’) as follows:

\[
\begin{align*}
(\text{a}) & \quad \forall x \left[ \left[ p(x) \land \forall y (m(y) \rightarrow s(x,y)) \right] \rightarrow \forall y [ (m(y) \land s(x,y)) \rightarrow h(x,y)] \right] \\
(\text{b}) & \quad \forall x \left[ \left[ p(x) \land \forall y (m(y) \rightarrow s(x,y)) \right] \rightarrow \forall y [ m(y) \rightarrow h(x,y)] \right]
\end{align*}
\]

\(^9\)We henceforth abbreviate "a determiner D forms QNP’s which are represented as quantifiers in a class A” and simply say “D is in A”.

\(^{10}\)A semantic interpretation of a determiner is a function D from sets to generalized quantifiers. A quantifier \(Q = D(A)\) *lives on* A iff for every set B: \(Q(A,B)\) holds iff \(Q(A,A \cap B)\) holds.
It is not hard to see that under classical semantics for Predicate Calculus the \( a \) and \( b \) formulae are equivalent.

From these results we may conclude that the fact that KQC does not allow Q-kataphora readings with universal determiners yields right predictions in singular kataphora and vacuous predictions in plurals.

We may get back now to our main issue. The discussion above is to show that KQC is a proper restatement of the linguistic facts with respect to kataphoric quantification. However, sentences with Q-kataphora are a small portion of sentences with anaphoric relations in natural language, so KQC is no more than a restatement of the facts: it does not offer any explanation for why is the distribution of “Q-kataphoric” determiners as it is. A more general principle from which KQC is derived has to be found. In the next sub-section we present informally an idea for how can KQC be derived from the well recognized anaphoric features of weak NP’s. This might provide better intuitions for the relations between anaphora and kataphora, as a background for the formal procedure for interpretation of Q-kataphora presented in section 4.

### 3.3 Weak quantifiers and KQC

A linguistically interesting class of quantifiers is the class of weak quantifiers. The primary linguistic motivation behind the introduction of this class was to explain why certain determiners are allowed in existential there sentences (ETS), and other determiners are not allowed. For example:

\[
\begin{aligned}
\text{(32) There is/are } & \left\{ \begin{array}{c}
\text{some} \\
\text{* every} \\
\text{less than two} \\
\text{* most} \\
\text{only (a)}
\end{array} \right. \\
\text{dog(s) in the yard.}
\end{aligned}
\]

Traditionally, the name “weak quantifiers” is attributed to the class of quantifiers allowed in ETS. Three different formal semantic definitions for this class follow:

1. [Barwise & Cooper ’81]:
   a. A quantifier \( Q \) is positive/negative strong iff for every set \( A \) for which \( Q(A,A) \) is defined, it equals to 1/0 (respectively).
   b. The weak quantifiers in natural language are exactly the quantifiers in natural language that are not strong.

2. [Keenan ’89]:
   a. A quantifier \( Q \) (over a universe \( U \)) is existential iff for every two sets \( A \) and \( B \): \( Q(A,B) \) holds iff \( Q(A \cap B, U) \) holds. \(^{11}\)
   b. The weak quantifiers in natural language are exactly the existential quantifiers in natural language.

3. [Johnsen ’87]:
   a. A quantifier \( Q \) is intersective iff for every two sets \( A \) and \( B \): \( Q(A,B) \) holds iff \( Q(A \cap B, B) \) holds. \(^{11}\)

\(^{11}\) The definitions of existential and intersective quantifiers are originally Barwise and Cooper’s
b. The weak quantifiers in natural language are exactly the intersective quantifiers in natural language.

The three definitions agree on the “hard core” of the determiners in English: according to the three definitions the following classification holds:

\[\begin{align*}
\text{Weak: } & \text{ no, some, at least/most } n, \text{ more/less than } n \ (n \geq 1) \\
\text{Not-Weak: } & \text{ most, every, each, all but } n \ (n \geq 1)
\end{align*}\]

Keenan’s definition extends Barwise & Cooper’s class of weak determiners (in natural language!) to include also “trivial” determiners like \(\text{at least zero, less than zero}\). Johnsen’s definition extends Keenan’s (also formally) to include also determiners which create quantifiers that do not live on their arguments. Two examples for such determiners that Johnsen discusses are \(\text{only}\) and \(\text{mostly}\).\(^{12}\) For our purposes it is sufficient to consider only determiners which form quantifiers with the property “lives on”. For such determiners, Keenan’s and Johnsen’s definitions are equivalent, according to proposition C4 in [Barwise & Cooper ’81], which is repeated in [Johnsen ’87].

If we compare the empirical classification of weak determiners as summarized in (33) and the formal classification of cardinal quantifiers we are in a position to conclude the following universal, on which the three definitions above agree:

\(\text{U: The class of weak determiners in natural language (with the “lives on” property) is exactly the class of CQ determiners.}\)

Such a possible universal is implied by some previous works (e.g., [Reinhart ’86], [Reinhart ’89]). Moreover, Reinhart shows (in her later work) that the class of weak determiners, in addition to being the set of determiners allowed in ETS, is also the set of determiners that allow for extraposition from NP’s. Even more important for our purposes here is another distinction made by Reinhart (as well as other researchers): Weak determiners are exactly the set of determiners allowing “E-type” interpretation in plural “donkey” sentences. For example, compare the sentences in (34) to those in (35) (both taken from [Reinhart ’86]):

\(\begin{align*}
\text{(34) } & \text{Every vampire who invited } \left\{ \begin{array}{l}
\text{two} \\
\text{several} \\
\text{less than five} \\
\text{many} \\
\text{between ten and thirteen}
\end{array} \right\} \text{ guests}; \text{ for dinner was through with } \text{them; by midnight.}
\end{align*}\)

\(\begin{align*}
\text{(35) } & \text{Every critic who fell asleep during } \left\{ \begin{array}{l}
\text{most} \\
\text{all but two} \\
\text{at least twenty percent of the}
\end{array} \right\} \text{ pieces}; \text{ wrote enthusiastic reviews about } \text{them;}.\end{align*}\)

In (34), where the antecedent is weak, for each vampire \(x\) quantified over the pronoun \(\text{them}\) can be interpreted as \(\text{all the guests that } x \text{ invited.}\) By contrast, in (35), where the antecedent contains a non-weak determiner, the pronoun does not have a reading like \(\text{all the pieces during which } x \text{ fell asleep.}\)

\(^{12}\)It is arguable whether these particles are to be analyzed as determiners, syntactically and semantically.

\(^{13}\)An exception to this proposed universal may be the “trivial” determiner \(\text{less than zero}\) which according to [Keenan ’89] is existential hence weak and it not in CQ. However, only technical maneuvers are needed to overcome this, and this determiner is not a real empirical justification for such effort.
Now we come to our point: as we claimed in 2.5, Q-kataphora is related to the phenomenon manifested as the above “E-type” readings of donkey sentences. How exactly the differences that were introduced in 2.5 between the two cases are created is beyond the scope of this paper. However, we claim, Q-kataphora and “E-type” anaphora of this kind are subject to the same constraint that Reinhart and others focused on. Together with our observation on Q-kataphora, we may phrase this constraint as follows:

**E-type Quantification Constraint (EQC):** Only weak determiners allow Q-kataphora and “E-type” anaphora.

Evidently, EQC does not entail KQC. Why then do we claim that KQC is actually predicted by EQC? The universal $U$ is a good point of departure:

1. Determiners which are not in CQ are not weak according to $U$, and therefore EQC predicts, as KQC, that non-CQ determiners do not allow Q-kataphora.

2. Determiners in CQ are weak according to $U$, and EQC agrees then with KQC that Q-kataphora is allowed.

We have yet to explain the observation in KQC that CQ quantifiers are impossible as antecedents in Q-kataphora, and why CQ~ are not easily interpreted for Q-kataphora readings. We saw already in section 2 that sentences like (36) are interpreted as in (37):

(36) Every pilot who saw them hit two migs.

(37) Every pilot who saw two migs hit two migs that he saw.

In order to explain the Q-kataphora reading we propose in section 4 a general semantic procedure to assign sentences in the form of (38) an interpretation paraphrased as in (39):

(38) $D_1$ pilot(s) who saw them hit $D$ migs.

(39) $D_1$ pilot(s) who saw $D$ migs hit $D$ migs that he/they saw.

Such a semantic procedure can yield for determiners $D$ in CQ tautological and contradictory readings. Consider for example the determiner $D = \text{no}$, which is in CQ, and $D_1$ as every or all but one. The interpretations generated for (38) with these values would be paraphrased as follows:

(40) Every pilot who saw no migs hit no migs that he saw.

(41) All but one pilots who saw no migs hit no migs that they saw.

Notice that because CQ determiners are mon, the set of pilots who saw (e.g.) no migs is a subset of the set of pilots who hit no migs that they saw. Therefore (40) is a patent tautology and (41) is a contradiction. Other substitutions for $D$ and $D_1$ in (38) could create highly cumbersome sentences to convey simple statements. For example: consider $D_1 = \text{some}$, and $D = \text{no}$ as before. Then the statement behind (38) is equivalent to the very simple statement in a sentence like some pilot saw no migs. Evidently, a sentence in the form of (38) is not a reasonable way to convey such a statement.

The same line of reasoning explains why CQ~ determiners are not easily interpreted in Q-kataphora. For example:

(42) Every pilot who saw them hit between two and five migs.
(43) Every pilot who saw between two and five migs hit between two and five migs that he saw.

(44) Every pilot who saw between two and five migs hit at least two migs that he saw.

The interpretation procedure for Q-kataphora yields for (42) a statement as paraphrased by (43) which is equivalent to the one in (44) but contains redundant information. Unlike the former example with a determiner from CQ, the “distance” between (43) and (44) is not that large. It seems that this may explain why kataphora in (42) is hard, while in the former cases from CQ it seems impossible.

Essentially, in this line of analysis we follow a principle spelled out in [Johnsen ’87], borrowed from an implicit assumption of [Barwise & Cooper ’81]’s analysis of ETS:

“The degeneracy constraint” (Johnsen): Sentences are interpreted in such a way that makes them not degenerate (i.e. patent tautologies or contradictions).

Actually, we used some generalization of this principle, which seems less ad-hoc, but formally it is not well defined:

Sentences are interpreted in such a way so complex syntactic structures are hard to be interpreted as simple semantic representations (e.g. “degenerate” patent tautologies or contradictions).

So, instead of Johnsen’s clear-cut constraint: “Try to avoid interpretation of patent tautologies or contradictions”, we propose a more subtle relational approach: “The availability of an interpretation is proportional to the ‘distance’ between the complexity of the syntactic structure and the semantic representation in formula.”

This principle might be viewed as a manifestation of Keenan’s novel “fidelity principle”: “Semantic interpretation is faithful to syntax” (see [Keenan ’93]). This principle is inspired by Chomsky’s minimalist program.

Basically, in order to account for the relations between the constraints on “E-type” anaphora and Q-kataphora, we propose that a constraint on the grounds of the weak/strong distinction is the only constraint “hard coded” into the interpretation procedures of E-type anaphoric links, including Q-kataphora. The elimination of weak determiners in CQ and the hard to get Q-kataphoric interpretation in determiners in CQ~ should be explained because of a “degeneracy principle”, which is independently motivated by other works on quantifiers. 14

Back to earth. Here is a summary of our proposals:

1. KQC is an adequate description of the linguistic facts concerning Q-kataphora: only CQ determiners are easy to be interpreted in this construction.

2. The class of cardinality quantifiers (CQ) is exactly the class of weak determiners in natural language. (U).

3. Q-kataphora and “E-type” readings of donkey sentences are subject to the same constraint (EQC): they are available only with weak determiners.

4. The fact that in effect only a proper subset of weak determiners allow easily for Q-kataphora may be explained using a principle of “degeneracy”, motivated independently by some previous works. The general proper formalization of this principle is yet to be given.

---

14See also [von Fintel ’01], where a similar principle is used to explain the distribution of various determiners in exceptional constructions.
The next thing to see is what general semantic process can yield a proper interpretation for kataphoric constructions. In the next section we propose that a generalization of the procedure for quantifier absorption that was first introduced in [Higginbotham & May ’81] is such a proper procedure.

4 Generalized Quantifier Absorption as a procedure for Q-kataphora interpretation

Until this point we discussed only the question of which determiners allow kataphoric quantification. The natural question is now how are sentences with kataphoric quantification being interpreted. We start with what is semantically probably the most complicated case of kataphoric quantification, the case of BP sentences. One of the most successful accounts for this tough problem is the Quantifier Absorption solution proposed in [Higginbotham & May ’81] (henceforth H&M). We first briefly summarize the main ideas in H&M’s proposal. Then we discuss some problems for H&M’s definition to account for the truth conditions of plural BP sentences, which call for a generalization of their proposal. We present our extension of H&M’s theory and show how the very same revised procedure is applicable to the general case of Q-kataphora. This is the first definition we know of for an operator that handles the variety of BP sentences together with more ordinary English constructions. We then discuss a problem for absorption to operate correctly on simple sentences with multiple quantification. This problem was pointed out in [Clark & Keenan ’86] (C&K), and we propose that it can be overcome in a syntactic theory for Q-kataphora that uses our semantic definitions.

Our motivation is then twofold:

1. To extend H&M’s definition for quantifier absorption to all kinds of BP sentences.

2. To show that the same revised procedure yields the right predictions in the other cases of Q-kataphora, so absorption is not an ad-hoc device only for treating BP sentences. Neither should it yield false predictions, as in the problems for H&M’s proposal discussed by C&K.

4.1 Higginbotham and May’s quantifier absorption

As a paradigmatic case of quantificational BP sentences we follow H&M and consider (3), restated here as (45):

(45) [ Every pilot who saw it \(j\) ] hit [ some mig that chased him, \(j\) ].

As we have mentioned, BP sentences like (45) can be paraphrased as follows:

(46) Every pilot who saw some mig that chased him hit some mig that chased him and that he saw.

This paraphrase preserves the two important features in the meaning attributed to (45): the kataphoric binding of it by the object some mig that chased him, and the identity in reference entailed by this binding; any mig that was hit must be at the same time a mig that was also seen, hence the repetition of and that he saw in (46).

Evidently, a straightforward translation of (45) to classical Predicate Calculus cannot work here. Try for example:

(47) \(\forall x[(p(x) \land s(x, y)) \rightarrow \exists y(m(y) \land c(y, x) \land h(x, y))]\)
The antecedent of the implication in (47) contains a wrong free occurrence of \( y \).

Problems appear also in straightforward translations into other logical formalisms like Discourse Representation Theory, Dynamic Predicate Logic, and Predicate Logic with Flexibly Binding Operators (see the discussion in D).

The problem is basically a problem for a theory of quantifier scope: how can the universal quantifier formed by the matrix verb’s subject in (45) have scope over the pronoun in the object, while letting the quantifier in the interpretation of the object to have scope over the subject?

H&M’s answer to this question is technically complex, but essentially quite simple: (45) is not to be interpreted as in (47), with unary quantifiers that can bind only free occurrences of one variable at a time. The sentence should be interpreted using a binary quantifier, which can bind free occurrences of two variables simultaneously.

Binary quantification is a tool first presented as an extension to the theory of generalized quantifiers, which helps to represent sentences with quantification over binary relations. A simple sentence to exemplify this is:

(48) Every cat and dog are enemies.

This sentence can be represented by a formula with binary quantification:\(^{15}\):

(49) \( \text{every}\{< x, y > \text{\text{cat}}(x) \land \text{\text{dog}}(y)\}, \{< x, y > \text{\text{enemy}}(x, y) \land \text{\text{enemy}}(y, x)\} \)

which says something like:

(50) Every pair of a cat and a dog is a member in the symmetric relation of two individuals who are enemies of each other.

The quantifier \text{every} in (49) is of signature \(< 2, 2 >\). This means that it applies to two binary relations. In general, the theory of polyadic quantifiers allows quantifiers of any signature of natural numbers:

For \( n \) positive natural numbers \( i_1, i_2, ... i_n \), a quantifier of signature \(< i_1, i_2, ... i_n \) over a universe \( U \) is a function \( \text{Pow}(U)^{i_1} \times \text{Pow}(U)^{i_2} \times ... \times \text{Pow}(U)^{i_n} \rightarrow \{0, 1\} \).

(\( \text{Pow}(U) \) is the power set of \( U \) - the set of all the subsets of \( U \)).

H&M claim that in the interpretation for (45) is involved a quantifier of signature \(< 2, 2 >\) which applies to the following two binary relations:

\[ R = \{< x, y > \mid p(x) \land s(x, y) \land m(y) \land c(y, x)\} \]

( the set of all pilot-mig pairs s.t. the pilot saw the mig and the mig chased the pilot)

\[ S = \{< x, y > \mid h(x, y)\} \]

(the set of all the pairs \( x, y \) s.t. \( x \) hit \( y \))

The above \(< 2, 2 >\) quantifier is denoted by \text{every-some} and is defined as follows:

For each binary relations \( E \) and \( F \) \text{every-some}(E, F) = \text{every}(\text{dom}E, \{x|\text{some}(E_x, F_x)\})\(^{16}\)

\text{every} and \text{some} are the simple unary \(< 1, 1 >\) quantifiers.

---

\(^{15}\)Of course, there are ways to represent such sentences using a formula containing only unary quantifiers, but reducibility is not our point here. For a discussion of reducibility of polyadic quantifiers see for example [Keenan ’87]

\(^{16}\)For a binary relation \( E \):

1. The domain of \( E \) is defined as follows: \( \text{dom}E = \{x|\exists y(< x, y > \in E)\} \)

2. The range of \( E \) on an individual \( x \) is defined by: \( E_x = \{y|< x, y > \in E\} \)
When we apply every-some to the binary relations R and S above we obtain the right truth conditions for (45):

\[
\text{every-some}(R, S) \quad \text{holds iff}
\]

\[
\text{every}(dom R, \{x \mid \text{some}(R_x, S_x)\}) \quad \text{iff}
\]

\[
\text{every}(\{x \mid \exists y R(x, y)\}, \{x \mid \text{some}(\{y \mid R(x, y)\}, \{y \mid S(x, y)\})\}) \quad \text{iff}
\]

\[
\forall x [\exists y (p(x) \land s(x, y) \land m(y) \land c(y, x)) \rightarrow \exists y (p(x) \land s(x, y) \land m(y) \land c(y, x) \land h(x, y))]
\]

This formal derivation agrees with (46), the informal paraphrase for (45).

The use of binary quantifiers in the representation for BP sentences is the key for the interpretive procedure that H&M propose for this kind of sentences within the GB framework. Under standard Logical Form (LF) analysis, (45) has the following two LF representations after the QR (Quantifier Raising) movement applies:

(51) a. [ every pilot who saw it ]; [[ some mig that chased him; ] ; [ e; hit e ]]

b. [ some mig that chased him; ] ; [[ every pilot who saw it ]; [ e; hit e ]]

Both representations exhibit the problem of scope pointed out above: in (51a) it is bound by its antecedent, in (51b) him is left unbound.

H&M propose that in addition to QR there exists at LF a syntactic operator of NP absorption, the operation of which can be described roughly as follows:

... [ NP ] [ NP ] ... \rightarrow ... [ NP; NP ]i,j ...

The output of this operator is an LF structure with an absorbed NP, indexed as NP\textsubscript{i,j}:

\[
\begin{array}{c}
S \\
\text{NP}_i \quad \text{NP}_j \\
\hline
\text{NP}_{i,j} \\
\hline
\end{array}
\]

For example, the LF for (45) after absorption is:

(52) [[ every pilot who saw it ]; [ some mig that chased him; ] ; [ e; hit e ]]

The interpretation procedure on the absorbed NP, the output of syntactic absorption consists of two distinct procedures: 17

1. Quantifier Absorption: Creating the absorbed \(< 2, 2 >\) quantifier from the \(< 1, 1 >\) quantifiers at S-structure:

(53) If P and Q are \(< 1, 1 >\) quantifiers, and E and F are binary relations, then PQ, the \textit{absorbed quantifier} of P and Q, is a \(< 2, 2 >\) quantifier defined by:

\[
PQ(E, F) = P(\text{dom} E, \{x \mid Q(E_x, F_x)\})
\]

17H&M do not make this distinction explicitly, but it is needed for the sake of clarity, and especially in the discussion that follows.
We already saw how this definition works in (45) to create the absorbed \( <2, 2> \) quantifier every-some from the \( <1, 1> \) quantifiers every and some.

2. **Relation Absorption**: Although H&M do not mention that explicitly, the relation \( R \), to which the absorbed quantifier \( PQ \) applies, is actually an intersection of the two binary relations visible in the subject and object of the matrix verb at S-structure. For example: in (45) \( R = A \cap B \), where \( A \) and \( B \) are the denotations of the N's *pilot who saw it* and *mig that chased him*, respectively:

\[
A = \{ <x, y > | p(x) \land s(x, y) \} \quad \text{(the pairs of pilots and things that they saw)}
\]

\[
B = \{ <x, y > | m(y) \land c(y, x) \} \quad \text{(the pairs of migs and things that they chased)}
\]

An important technical question with respect to this process concerns the way the order of the arguments in the binary relations \( A \) and \( B \) above is determined. For example, in the interpretation of (45) nothing prevents the relation \( B \) from being:

\[
B' = \{ <y, x > | m(y) \land c(y, x) \}
\]

Actually, this is even a natural candidate for representing the N’s *mig that chased him*, where the CN *mig* precedes the pronoun *him*. However, it is clear that such a choice of \( B' \) instead of \( B \) would create an absurd interpretation for (45) using H&M’s definition, as the intersection of \( A \) and \( B' \) involves bizarre denotations of pilots who are migs. A rather simple answer to this new question is given in section 4.4.

The process that H&M propose is summarized in figure 1.

For H&M absorption is a free and general procedure at LF which occurs in any case of multiple quantification. As was pointed out in [Clark & Keenan ’86] (C&K), this claim on the generality of absorption yields many faulty results with respect to simple sentences with multiple quantification. We will discuss this kind of problems in sub-section 4.6.

For the cases of singular BP sentences, like the sentences in (19), H&M’s operator of absorption works fine – it creates the expected truth conditions for all the acceptable instances of determiners in the subject and in the object of the sentence. This is verified in many examples in the works of H&M and C&K. However, some problems arise once we consider also plural BP sentences.
4.2 Problems for H&M’s absorption procedure with plural BP sentences

Reconsider the following sentence:

(54) [ Every pilot who saw them; ] hit [ two migs that chased him; ]

H&M’s absorption operator predicts the following truth condition for this plural BP sentence:

\[ \text{every-two}(R, S) \text{ holds iff } \]
\[ \text{every}(\text{dom } R, \{x | \text{two}(R_x, S_x)\}) \text{ iff } \]
\[ \forall x [\exists y (p(x) \land s(x, y) \land m(y) \land c(y, x)) \rightarrow \exists \exists y (p(x) \land s(x, y) \land m(y) \land c(y, x) \land h(x, y))] \]

This PC formula can be paraphrased as:

(55) Every pilot who saw *some* mig that chased him hit *two* migs that chased him and that he saw.

And such a paraphrase does not agree with the kataphoric quantification reading speakers tend to assign to such sentences. Instead, (54) should be paraphrased as follows:

(56) Every pilot who saw *two* migs that chased him hit *two* migs that chased him and that he saw.

The difference between the statements conveyed by (55) and by (56) can be exemplified in the mini-model appearing in figure 2. According to H&M, (54) should be false in this model because pilot1, who saw *some* mig that chased him (mig1), did not hit two migs that chased him and that he saw (he hit and saw only mig1). According to most speakers intuitions, paraphrased in (56), (54) is true in this model because pilot1 is not understood as being quantified over by the *every* phrase (he did not see *two* migs that chased him) and pilot2 satisfies both the restriction (he saw two migs that chased him, mig2 and mig3) and the predicate (he also hit these two migs).

The same problem appears also in Q-kataphora readings in BP sentences with other numerals or with the determiner *more than n* at the matrix verb’s object. Similar problems show up also when we consider the S-kataphora readings in plural BP sentences. Consider for example the S-kataphora in:

\[ \text{We henceforth freely use the notation } \exists \exists (P(x)) \text{ as a shorthand for the expression } \exists x \exists y ((x \neq y) \land P(x) \land P(y)) \]
Every pilot who saw them hit all migs that chased him.

Similarly, a proper formalization for (57) is:

$$\forall x \{ [p(x) \land \forall y ((m(y) \land c(y, x)) \rightarrow s(x, y))] \land \forall y ((s(x, y) \land m(y) \land c(y, x)) \rightarrow h(x, y)) \} $$

while H&M's definition formalizes (57) as:

$$\forall x \{ [p(x) \land \exists y ((m(y) \land c(y, x)) \land s(x, y))] \land \forall y ((s(x, y) \land m(y) \land c(y, x)) \rightarrow h(x, y)) \} $$

Notice that it is not a genuine linguistic problem for H&M's account if our classification of S-kataphora as a different linguistic phenomenon is correct, which entails that actually Quantifier Absorption does not apply to sentences like (57). It is only a problem if we wish Quantifier Absorption to apply to all kinds of BP sentences. However, it is certainly a weakness of H&M's definition with respect to capturing a formal generalization for quantifier absorption in natural language.

These problems call for some revision in H&M's treatment of the procedure of absorption to generalize it also to the plural instances of BP sentences.

But, as we are going to see, there are still more important reasons to revise H&M's concept of absorption than to account for the semantics of plural BP sentences: one is to explain more common cases of kataphoric quantification, and another is to explain how a revised procedure of absorption can avoid the problems pointed out by Clark and Keenan (C&K) for simple non-kataphoric sentences. We begin by a discussion of revision needed in H&M's absorption operator, in order to account for the semantics of plural BP sentences. Consequently, we will see that this revision leads to a more general procedure of quantifier absorption.

4.3 A revised operator of absorption

In order to account for the truth conditions of plural BP sentences like (54) and the other examples in (10), it may help to notice the feature of H&M's operator which is responsible for the faulty predictions in such cases. The restriction domR for the absorbed quantifier PQ in definition (53) does not depend on the quantifier Q formed by the object of the BP sentence. For instance: in (54) domR is the set of pilots who saw some mig that chased them, instead of the required set: pilots who saw two migs that chased them. If instead of using the intersection R of the relations A and B from the S-structure of (54), we use A and B themselves, we can formulate the required set of pilots as follows:

$$\{ x | \text{two}(B_x, A_x) \} = \{ x | p(x) \land \exists y ((m(y) \land s(x, y) \land c(y, x)) \} $$

The process in question is what we called in H&M's account by the name Relation Absorption: using $R = A \cap B$ as the restriction of the absorbed quantifier PQ. If we renounce this process, we not only have a more simple way to formulate absorption, we can also resolve the problems with plural BP sentences for H&M's account. Moreover, as we will see later, this revised procedure will be easier to generalize.

Concretely: the absorbed quantifier PQ should apply directly to the relations A, B and S, and therefore it should be a binary quantifier of signature $< 2, 2, 2 >$. Consider for example the quantifier every-two, which is required in order to interpret (54). We claim that its right definition should be:

---

19 C&K point out this feature of H&M's definition, but do not consider it problematic, because they do not consider the problematic BP sentences for H&M's definition.

20 C&K also propose a (wrong) definition for a $< 2, 2, 2 >$ absorbed quantifier, but they reject it for reasons which will be shown to be quite irrelevant.
every-two(A, B, S) = every(\{x|two(B_x, A_x)\}, \{x|two((A \cap B)_x, S_x)\})

Notice that the second set to which every applies in this definition is exactly the same set as in H&M’s definition (53) (R in (53) is our A \cap B).

By the discussion above it is expected that the revised definition for every-two yields the correct truth conditions for (54), and indeed:

\[
\text{every}(\{x|two(B_x, A_x)\}, \{x|two((A \cap B)_x, S_x)\}) \quad \text{holds iff}
\]

\[
\forall x \exists y(p(x) \land s(x, y) \land m(y) \land c(y, x)) \implies \exists y(p(x) \land s(x, y) \land m(y) \land c(y, x) \land h(x, y))
\]

This is exactly the formal parallel for the paraphrase (56).

The general definition we propose for an absorbed quantifier PQ is therefore:

(58) Revised Quantifier Absorption: If P and Q are <1, 1> quantifiers and E, F and G are binary relations, then PQ, the absorbed quantifier of P and Q, is a <2, 2, 2> quantifier defined by:

\[
PQ(E, F, G) = P(\{x|Q(F_x, E_x)\}, \{x|Q((E \cap F)_x, G_x)\})
\]

As we are going to see, definition (58) yields the correct kataphoric quantification reading for BP sentences, and by section 3 it should apply only when Q is a weak quantifier (EQC).

We saw already that this definition works for P = every and Q = (at least) two. It is clear that for simple existential quantifiers like some, a (in singular), several (in plural), (58) yields the same (correct) result as H&M’s (53), since in these cases the restriction of P in (58) is reduced to H&M’s proposal. In formula:

\[
\{x|Q(B_x, A_x)\} = \{x|some(B_x, A_x)\} = dom(A \cap B) = domR
\]

In general, we claim, the definition in (58) yields for every acceptable BP sentence in the form of (59) a semantic interpretation which can be paraphrased as in (60):

(59) D_1 pilot(s) who saw it/them hit D mig(s) that chased him/them.

(60) D_1 pilot(s) who saw D mig(s) that chased him/them hit D mig(s) that chased him/them and that he/she saw.

This property of definition (58) is a conclusion from a corollary proved in appendix C:

(61) Corollary: For every unary quantifiers P and Q with the property “lives on”, where Q in CQ\perp or in CQ\sim, definition (58) yields the following formula as the interpretation of BP sentences in the form of (59).

(62) P(\{x|p(x) \land Q(\{y|m(y) \land c(y, x)\}, \{y|s(x, y)\})\}, \{x|Q(\{y|m(y) \land c(y, x) \land s(x, y)\}, \{y|h(x, y)\})\})

The formula (62) is an adequate translation to the paraphrase scheme (60). Since as we observed in section 3 (in KQC), only BP sentences as in (59) with a determiner D in CQ\perp or in CQ\sim are interpretable, we see that definition (58) is adequate to the paraphrase in (60), for which (62) is a formal translation. The meaning of this result is that if the observation in H&M is accepted, definition (58) yields the expected truth conditions for the whole class of BP sentences with various determiners as discussed in section 2.

To summarize this section. We propose to renounce the process of Relation Absorption in H&M’s procedure. Instead of being an <2, 2> quantifier, the absorbed quantifier is now represented as a <2, 2, 2> quantifier defined as in (58), which applies directly to the original relations without absorbing them. This generalized, yet more simple procedure is illustrated in figure 3 (compare 1).
4.4 The order of the arguments in the binary relations

The Absorbed Quantifier in H&M’s definition and in our proposed revision applies to two or three binary relations, respectively. As was mentioned before, it should be explained how the order of the arguments of these relations is determined. H&M do not discuss this problem although it has a straightforward solution. Assume that an argument of a relation is assigned the index of the NP from which it is derived. For example: consider the NP [every [N, pilot who saw it]2]. In the relation derived from the N ‘pilot who saw it’ an argument x is assigned an index i iff a predicate P representing the head of NP1 applies to x. In general, we denote by R(1,2,...,n) an n-ary predicate whose arguments are assigned the indices 1,2,...,n in this order. From the above N, then, two different indexed binary relations can be derived:

\[ A_{(1,2)} = \{ <x, y > | p(x) \land s(x, y) \} \]
\[ A_{(2,1)} = \{ <y, x > | p(x) \land s(x, y) \} \]

In the interpretation of a BP sentence, the indices originated in the main predicate of the sentence (e.g. hit), remain in their S-structure order, and the order of the other two relations is determined according to a simple indices unification process. Consider for example the BP sentence (45), restated here:

(63) [ Every pilot who saw it2 ]1 hit [ some mig that chased him1 ]2.

The indexed main predicate is:

\[ S_{(1,2)} = \{ <x, y > | h(x, y) \} \]

The reason we assume that the order of arguments in the main predicate remain as they appear in S-structure is because the main predicate is not a part of the process of absorption: in order for absorption to apply, we need only two NP’s: the object and the subject. It is reasonable to assume that only during this process arguments may change their order. Only then the absorbed quantifier may apply to the two relations extracted form the NP’s and to the main predicate “in situ”.

The two indexed relations that can be derived from the main subject are \( A_{(1,2)} \) and \( A_{(2,1)} \) as above, and the two relations derived from the main object are:

\[ B_{(1,2)} = \{ <x, y > | m(y) \land c(y, x) \} \]
The only triplet of relations $A, B, S$ with indices that match are $A_{(1,2)}, B_{(1,2)}$ and $S_{(1,2)}$. These are exactly the three relations used in H&M’s definition and in our revision, and which allow the derivation of the correct truth conditions for BP sentences.

This rather simple conception of the way the order of arguments in “BP predicates” is determined proves itself to be useful also for the account of other cases of Q-kataphora besides BP sentences.

### 4.5 Relation Inducement - Quantifier Absorption in the analysis of other cases of Q-kataphora

We may turn now to more simple cases of Q-kataphora than BP sentences, and see that an account for the general case of Q-kataphoric readings emerges naturally from the revised definition for Quantifier Absorption we presented. Reconsider the sentence:

(64) [ Every pilot who saw them], hit [ two migs ],

which is interpreted as in the paraphrase:

(64’) Every pilot who saw two migs hit two migs that he saw.

It may seem that Quantifier Absorption as presented in (58) is irrelevant for the interpretation of sentences in the form of (64): an absorbed quantifier applies to three binary relations, and (64) exhibits two binary relations and a one-place predicate:

$$A = \{ < x, y > \mid \text{p}(x) \land \text{s}(x, y) \}$$

$$M = \{ y \mid \text{m}(y) \}$$

$$S = \{ < x, y > \mid \text{h}(x, y) \}$$

However, an implicit assumption in both works of H&M and C&K is that an absorbed quantifier can apply to sets which are not binary relations. A description of how exactly this happens is not given in these works and we should try to work out the details here.

Consider the one place predicate $M$ in the example above. The most natural way in which the absorbed quantifier every-two derived from sentence (64) can apply to $A$, $M$ and $S$, is by lifting $M$ into a two place predicate. This shift can simply be done by cartesian product of $M$ with the domain of individuals $U$. The two possibilities to do this is by left or right multiplication:

a. $M \rightarrow M \times U = B'$

b. $M \rightarrow U \times M = B''$

If we use the same indexation procedure we proposed in the previous sub-section we can restate the process of the above cartesian product using an explicit indexation as follows:

1. **Original sentence**:

   [ Every pilot who saw them$_2$], hit [ two migs ],

2. **An indexed representation of sets**:
\[ A_{1,2} = \{ <x, y > | p(x) \land s(x, y) \} \]
\[ M_{2} = \{ y|m(y) \} \]
\[ S_{1,2} = \{ <x, y > | h(x, y) \} \]

3. Cartesian multiplication of sets:

a. \( M_{2} \times M_{2} \times U_{i} = B'_{i,2} \)
b. \( M_{2} \times U_{j} \times M_{2} = B''_{j,2} \)

Notice that we assume that arguments which originate in the domain \( U \) are assigned a free index, while other arguments are assigned the same indices which are on the NP’s from which they are derived.

The trivial assumption we made in the previous sub-section is that in order for quantifier absorption to apply to three binary relations they must agree on the indices. In the case of BP sentences this process determines the order of the arguments of the relations. In non-BP Q-kataphora sentences, like (64), the process can choose between the possible two binary relations (left or right product) according to the same index criterion. This boils down to simple unification of the indices on the binary relations.

In the case of \( B' \) such a unification fails: \( A, B' \) and \( S \) have the indices \((1,2), (2,i) \) and \((1,2) \). This produces the unsolvable equations \( 1=2=1, 2=i=2 \). Unification does succeed with \( B'' \): the indices \((1,2), (i,2) \) and \((1,2) \) lead to the solution \( i=1 \).

This simple process leads to choose \( B'' \) as the third binary relation for the absorbed quantifier which is exactly what we need in order to account for the truth conditions of (64) (as in the paraphrase (64’)) using quantifier absorption:

\[
\text{every} \two (A, B'', S) = \\
\text{every} \{x\two (B''_{x}, A_{x})\}, \{x\two (A \cap B''_{x}, S_{x})\}
\]

\[
\forall x[p(x) \land \exists y(m(y) \land s(x, y)) \rightarrow \exists y(s(x, y) \land m(y) \land h(x, y))] 
\]

It should not surprise us that a similar corollary to (62) exists also for the case of non-BP Q-kataphora, which shows that the procedure of quantifier absorption we proposed agrees with the intuitions behind paraphrase (56) and similar cases. (We save the details of this result, which is a direct consequence of the general claim in appendix C).

The process we propose consists then of three stages:

1. Creating all possible binary relations using cartesian product with the domain \( U \).

2. Index unification and eliminating impossible relations.

3. Application of Quantifier Absorption.

The first stage is of course optional and occurs only when there are no three binary relations ready for the absorbed quantifier. Let us call this process by the name Relation Inducement (RI).

The conclusion from this sub-section is a new result on an old problem: The definition of Quantifier Absorption we propose is no longer a special mechanism needed only for the interpretation of BP sentences. Using simple algebraic operations (cartesian product and index unification) we can use Quantifier Absorption as a general process for obtaining the truth conditions of the other cases...
of Q-kataphora as well. Notice that as we saw in the previous sub-section, the process of index unification is needed anyway for the interpretation of BP sentences.

We see that the revised procedure of Absorption we propose yields the correct truth conditions for the whole class of sentences with Kataphoric Quantification we considered. Does it mean that in addition to the discussed Q-kataphora sentences, absorption can be formulated in a way so it can apply freely also to other kinds of sentences with quantificational forward anaphora? Unfortunately, like C&K, we do not have a full answer to this question. In the next section we discuss the remaining problems and two alternatives for a solution.

4.6 C&K’s problem and the generality of Absorption

Theoretically, there is no reason to assume that quantifier absorption is a special semantic mechanism which applies only to BP sentences. Indeed, we have shown in the last section, the revised formula for absorption we propose is a useful mechanism to analyze the otherwise hard to get (see appendix D truth conditions of the variety of possible Q-kataphora constructions. Still, if we wish to keep absorption a free semantic procedure, as H&M first proposed, it is essential to consider the problem discussed in the work of C&K. C&K have shown that H&M’s definition for absorption produces the wrong truth conditions for simple sentences as in (65):

(65) Every pilot hit some mig.

In addition to such simple sentences with multiple quantification, it is not hard to show that H&M’s definition produces wrong truth conditions to simple anaphora as in (66) or “donkey” anaphora as in (67):

(66) Every pilot hit some mig that chased him.

(67) Every pilot that saw some mig hit it.

(we spare here the derivations).

Our proposal is no exception to these sad results. Let us exemplify it with respect to sentence (65):

1. Original sentence:

   [ Every pilot ]_1 hit [ some mig ]_2.

2. An indexed representation of sets:

   \[ M(1) = \{ x | p(x) \} \]
   \[ N(2) = \{ y | m(y) \} \]
   \[ S(1,2) = \{ < x,y > | h(x,y) \} \]

3. Cartesian product of sets:

   a. \( M(1) \times N(2) = A'_{1,2} \)
   b. \( M(1) \times U_{1,j} = A''_{1,j} \)
   c. \( N(2) \times U_{i,j} = B'_{i,2} \)
   d. \( N(2) \times U_{i,j} = B''_{i,2} \)
Unification of indices yields that the only possible matching triplet of binary relations is $A'$, $B''$ and $S$. Quantifier absorption then yields:

$$\text{every} - \text{some}(A', B'', S) = \text{every}(\{x \mid \text{some}(B'', A'_x)\}, \{x \mid \text{some}((A' \cap B'')_x, S_x)\}) =$$

$$\forall x [\exists y (m(y) \land p(x)) \rightarrow \exists y(p(x) \land m(y) \land s(x, y))]$$

This formula holds when there are some pilots and no migs, unlike (65), which is interpreted as false in such a situation.

In [May '90] it is proposed to deal with this kind of problem by requiring that three relations of the same arity must take place in quantifier absorption, thus restricting its operation only to the case of BP sentences. However, such a constraint is too specific to be theoretically interesting: what May proposes is simply to use a special semantic procedure for the interpretation of BP sentences. Also, this would not allow us to use the same procedure for other cases of Q-kataphora, for which Quantifier Absorption provides too an adequate analysis, and we should invent a new semantic mechanism for the interpretation of such sentences.

If we wish to keep the benefits of Quantifier Absorption it seems that we have only two possible directions of solving the problem for the analysis of non Q-kataphoric sentences:

1. Trying to keep absorption a general process and adjust our definition so it can correctly apply also to the non-kataphoric cases shown above, while reducing to our revised definition of absorption in the case of Q-kataphora.

2. Not allowing absorption to operate on non Q-kataphoric sentences. This should be imposed by some other principle.

The first possibility seems to be very hard to follow, although a-priori it is rather reasonable. Having an extremely complicated formula for absorption, which may result from this direction might not be a theoretically plausible way out of the maze.

We tend to believe that the second possibility is more promising. It depends on a strong formulation of the syntactic kataphoric circumstances possible in English. Such an enterprise is beyond the scope of this paper. However, a rough formulation of such a principle, which should be given more syntactic content, is the following:

"Try to apply all possible interpretative procedures without absorbing quantifiers. If no such procedure leads to an interpretation and a pronoun is coindexed with a weak NP that follows it: apply Relation Inducement and Quantifier Absorption"

Such a principle would mean that Quantifier Absorption is a “last resort” mechanism, which applies only to cases of kataphora, to yield the Q-kataphora readings. This might agree with the intuitions of speakers, who tend not to accept the Q-kataphora reading whenever there is another possible (forward) anaphoric reading for a pronoun.

Unfortunately, we must leave the details of such a syntactic condition open for further research.

5 Conclusions

We believe that a better understanding of the way quantifiers in natural language may “look backwards” may emerge from the ideas presented in this paper. First, it is proposed that kataphora is not essentially different from forward anaphora with respect the weak / strong distinction between determiners, and the different kinds of anaphoric links these two classes allow for. Second, special
semantic procedures seem to be needed in order to interpret sentences with kataphoric quantification. It seems that Quantifier Absorption as it is presented here is general enough to be able to account for all the kataphoric quantificational constructions we know about.

Some important problems are still left open. A comprehensive research is still needed in order to determine the empirical status of kataphoric quantification and kataphora in general. The results of such a research may require further theoretical explanations for the (in)accessibility of various kataphoric readings. In addition, a full syntactic account is yet to be given for the circumstances in which quantifier absorption can occur. This leaves us still with many open questions concerning the psycholinguistic and syntactic analysis of Kataphoric Quantification. However, we hope that the semantic results presented in this paper may become part of a comprehensive future theory of forward and backward anaphora.

Appendices

A A Note on “Set” Anaphora

We saw in section 2 that pronouns in what we called “set” anaphora do not always simply stand for a constant set of individuals. Sometimes a pronoun corefers with an antecedent that contains another pronoun, which is interpreted as a bound variables. For example:

(i) Every child who saw at least three of his brother’s chocolate bars wanted to eat all of them.

(ii) Every pilot who saw them hit no migs that chased him.

The pronouns them in (i) and (ii) do not represent a constant set of chocolate bars or migs, but refer to a different set for each child / pilot quantified over.

In [Heim ’90], Heim presents what she classifies as “an E-type theory of pronouns”, which is a possible explanation for what is going on here. According to Heim’s proposal, a singular pronoun can be viewed as a function from pairs of individuals and situations to individuals. It is shown how this is a possible way of accounting for the truth conditions of many of the almost classical “donkey” and “sage plant” sentences. If we extend Heim’s approach to cases of set anaphora like (i) or (ii), the plural pronoun can be interpreted as a function from individuals (and situations) to sets of individuals. For example, in (i) the pronoun them attributes for each child the set of his brother’s chocolate bars. In (ii), them attributes for each pilot the set of migs that chased him. As Heim puts it, the function that an “E-type” pronoun like this represents becomes salient in the context of a specific sentence. In examples like (iii) or (iv) where the pronoun refers to a constant set, this function is constant.

(iii) Every child who saw at least three chocolate bars wanted to eat all of them.

(iv) Every pilot who saw them hit no migs.

The fact that plural pronouns may be taken as representing this type of functions may dictate some changes in the proposal brought in [Heim ’90]. We refer the reader to this paper for further details concerning “E-type” theories of pronouns interpretation.

21Heim’s account preserves the ‘uniqueness’ requirement regarding the antecedent of such anaphora, relativizing this uniqueness to the situation argument of the function.
Recently, such changes were carried out in [Lappin & Francez ‘94], where the domain of the function has a finer algebraic structure than sets of individuals. Using the i-sum operation (cf. [Link ‘87]) it links denotations for plural NP’s. The main advantage of this approach is the possibility to provide a unified account of both universal and existential readings of donkey sentences. Heim’s original account provides universal readings only. We refer the reader to [Lappin & Francez ‘94] for the actual details.

B Q-kataphora with a Universal QNP Antecedent

In section 3.2 we claimed that the following two paraphrase schemes are equivalent for every determiner \( D \) that “lives on” its arguments:

\[
(31')
\begin{align*}
&\text{a.} & D \text{ pilot(s) who saw all migs hit all migs that he saw. (Q-kataphora)} \\
&\text{b.} & D \text{ pilot(s) who saw the migs hit all migs. (S-kataphora)}
\end{align*}
\]

This claim may be formalized as follows:

**Proposition:** For every quantifier \( Q \) with the property “lives on” the following formulae are equivalent:

\[
(\text{a}) & \quad Q(\{x[p(x) \land \forall y(m(y) \rightarrow s(x,y))\}, \{x[\forall y((m(y) \land s(x,y)) \rightarrow h(x,y))]\}) \\
(\text{b}) & \quad Q(\{x[p(x) \land \forall y(m(y) \rightarrow s(x,y))\}, \{x[\forall y[m(y) \rightarrow h(x,y)]\})
\]

**Proof:** Trivial, by definition of the “living on” property. (a) is equivalent to:

\[
Q(\{x[p(x) \land \forall y(m(y) \rightarrow s(x,y))\}, \{x[\forall y((m(y) \land s(x,y)) \rightarrow h(x,y))]\}) \cap \{x[p(x) \land \forall y(m(y) \rightarrow s(x,y))\}) = \\
Q(\{x[p(x) \land \forall y(m(y) \rightarrow s(x,y))\}, \{x[p(x) \land \forall y[m(y) \rightarrow (s(x,y) \land h(x,y))]\})
\]

Using the same “living on” property of \( Q \) it is easy to see that this is equivalent to (b). QED.

C Adequacy of Quantifier Absorption

It is assumed throughout this paper, according to speakers’ intuitions and previous works, that Q-kataphoric sentences as in a below can be interpreted as equivalent to the paraphrases in b, which reflect the semantics of the formulae in c:

\[
\begin{align*}
&(\text{i}) \quad \text{“Simple” Q-kataphora:} \\
&\text{a.} & D_1 \text{ pilot(s) who saw them hit D migs.} \\
&\text{b.} & D_1 \text{ pilot(s) who saw D migs hit D migs that he/they saw.} \\
&\text{c.} & P(\{x[p(x) \land Q(\{y[m(y)\}, \{y[s(x,y)]\})\}, \{x[Q(\{y[m(y) \land s(x,y)]\}, \{y[h(x,y)]\})]\})
\end{align*}
\]

\[
\begin{align*}
&(\text{ii}) \quad BP \text{ sentences:} \\
&\text{a.} & D_1 \text{ pilot(s) who saw it/them hit D mig(s) that chased him/them.} \\
&\text{b.} & D_1 \text{ pilot(s) who saw D mig(s) that chased him/them hit D mig(s) that chased him/them and that he/they saw.} \\
&\text{c.} & P(\{x[p(x) \land Q(\{y[m(y) \land c(y,x)]\}, \{y[s(x,y)]\})\}, \{x[Q(\{y[m(y) \land c(y,x) \land s(x,y)]\}, \{y[h(x,y)]\})]\})
\end{align*}
\]
In section 4 we used the denotations of the following binary relations:

\[ A = \{ <x, y > | p(x) \land s(x, y) \} \]
\[ B'' = \{ <x, y > | m(y) \} \]
\[ B = \{ <x, y > | m(y) \land c(y, x) \} \]
\[ S = \{ <x, y > | h(x, y) \} \]

Using definition (58) for Quantifier Absorption and the procedure we proposed, it was claimed that the above sentences schemes are interpreted as follows:

(i) \[ P(\{x|Q(B''_{x}, A_{x})\}, \{x|Q((A \cap B'')_{x}, S_{x})\}) = P(\{x|Q(\{y|m(y)\}, \{y|p(x) \land s(x, y)\}\}), \{x|Q(\{y|p(x) \land s(x, y) \land m(y) \land c(y, x)\}, \{y|h(x, y)\}\})\]

(ii) \[ P(\{x|Q(B_{x}, A_{x})\}, \{x|Q((A \cap B)_{x}, S_{x})\}) = P(\{x|Q(\{y|m(y) \land c(y, x)\}, \{y|p(x) \land s(x, y)\}\}), \{x|Q(\{y|p(x) \land s(x, y) \land m(y) \land c(y, x)\}, \{y|h(x, y)\}\})\]

We would like to show that the “c’s” (the desired truth conditions) and the “d’s” (the outputs of Quantifier Absorption) are equivalent for quantifiers in CQ\[ and CQ\~{}, as which we observed in section 3 are the relevant cases for Q-kataphora. It is evident that (i)(c) and (i)(d) are special cases of (ii)(c) and (ii)(d), respectively (consider the relation c(y, x) as the complete relation). Therefore, the following proposition is what we need to prove:

**Proposition:** For any unary quantifiers P and Q with the property “lives on”, where Q is in CQ\[ or in CQ\~, the following two formulae are equivalent:

\[ c. P(\{x|p(x) \land Q(\{y|m(y) \land c(y, x)\}, \{y|s(x, y)\}\}), \{x|Q(\{y|m(y) \land c(y, x) \land s(x, y)\}, \{y|h(x, y)\}\})\]
\[ d. P(\{x|Q(\{y|m(y) \land c(y, x)\}, \{y|p(x) \land s(x, y)\}\}), \{x|Q(\{y|p(x) \land s(x, y) \land m(y) \land c(y, x)\}, \{y|h(x, y)\}\})\]

**Proof:** We denote:

\[ M_{1} = \{x|p(x) \land Q(\{y|m(y) \land c(y, x)\}, \{y|s(x, y)\}\})\]
\[ M_{2} = \{x|Q(\{y|m(y) \land c(y, x)\}, \{y|p(x) \land s(x, y)\}\})\]
\[ N_{1} = \{x|Q(\{y|m(y) \land c(y, x) \land s(x, y)\}, \{y|h(x, y)\}\})\]
\[ N_{2} = \{x|Q(\{y|m(y) \land c(y, x) \land s(x, y) \land p(x)\}, \{y|h(x, y)\}\})\]

We should show that \[ P(M_{1}, N_{1}) \] and \[ P(M_{2}, N_{2}) \] are equivalent. First we show that \[ M_{1} = M_{2} \]:

1. \[ M_{1} \subseteq M_{2} \]:
   - Assume \( x \in M_{1} \).
   - Then \( p(x) \) holds by definition and therefore \( \{y|s(x, y)\} = \{y|p(x) \land s(x, y)\} \)
   - \( Q(\{y|m(y) \land c(y, x)\}, \{y|s(x, y)\}) \) holds since \( x \in M_{1} \)
   - and therefore we also get that \( Q(\{y|m(y) \land c(y, x)\}, \{y|p(x) \land s(x, y)\}) \) holds.
   - or \( x \in M_{2} \)

2. \[ M_{2} \subseteq M_{1} \]:
   - Assume \( x \in M_{2} \).
   - Then \( Q(\{y|m(y) \land c(y, x)\}, \{y|p(x) \land s(x, y)\}) \) holds.
   - and therefore \( p(x) \) holds by definition and therefore \( \{y|s(x, y)\} = \{y|p(x) \land s(x, y)\} \)
   - \( M_{2} \subseteq M_{1} \)
assume $x \in M_2$
then $Q(\{y|m(y) \land c(y,x)\}, \{y|p(x) \land s(x,y)\})$ holds
so because $Q$ is in $CQ|$ or $CQ\sim$ we have: $0 < |\{y|m(y) \land c(y,x)\} \cap \{y|p(x) \land s(x,y)\}| \leq \infty$
this entails that $p(x)$ holds and therefore $\{y|s(x,y)\} = \{y|p(x) \land s(x,y)\}$
from these facts it is easy to see then that from $x \in M_2$ we get $x \in M_1$ by definition.

We may denote then $M_1 = M_2 = M$.
It is left to show that $P(M, N_1)$ and $P(M, N_2)$ are equivalent. Since $P$ has the property “lives on” it is sufficient to show that $P(M, M \cap N_1)$ and $P(M, M \cap N_2)$ are equivalent. We will show this simply by proving $M \cap N_1 = M \cap N_2$.

1. $M \cap N_1 \subseteq M \cap N_2$:
assume $x \in M \cap N_1$.
from $x \in M = M_1$ we get that $p(x)$ holds
from $x \in N_1$ we get:

$0 < |\{y|m(y) \land c(y,x) \land s(x,y)\} \cap \{y|h(x,y)\}| \leq \infty$ (by the $CQ|$ or $CQ\sim$ property of $Q$)

we may conclude:

$0 < |\{y|m(y) \land c(y,x) \land s(x,y)\} \cap \{y|p(x) \land h(x,y)\}| \leq \infty$
or $x \in N_2$. So $x \in M \cap N_2$

2. $M \cap N_2 \subseteq M \cap N_1$: by similar considerations.

We may conclude that $P(M_1, N_1)$ and $P(M_2, N_2)$ are equivalent. QED.

The conclusion from this result is that the revised definition we gave for Quantifier Absorption is adequate for the interpretation of $Q$-kataphora with $CQ|$ and $CQ\sim$ quantifiers. According to the KQC classification this means that it is OK for all quantifiers that allow $Q$-kataphora in English.

D  Q-kataphora interpretation in other semantic frameworks

In this appendix we would like to point out some problems for two proposals to account for the semantics of kataphoric constructions. One is van Deemter’s version of “patient” DRT. The other is Pagin and Westerstahl’s Predicate Logic with flexibly binding operators.

D.1 “Patient” DRT

In [van Deemter ’90] a revised version of DRT is proposed to handle with kataphora in natural language. The modification van Deemter makes in the formalism of DRT allows a kataphoric pronoun to introduce an incomplete discourse entity. Such a Reference Marker (RM) may be completed later in the translation of the sentence by the antecedent in the kataphora. These “patience” principle and “completion” principle are responsible for the interpretation of sentences like the following (from [van Deemter ’90]):

(i) Whenever she was off duty the president spent her time in the swimming pool.
The sentence (i) is represented as in DRS-1. According to van Deemter’s “Optional Rise” principle the RM for the president may be introduced also in higher accessible boxes in the DRS, as in DRS-2.

Van Deemter maintains that the same procedures also apply in the representations of other cases of kataphora. For example:

(ii) Every farmer who admires her courts a widow.  

However, this claim is problematic: the procedure leading to DRS-1 leads to a “donkey” reading of (ii), which can be paraphrased as in:

(ii’) Every farmer who admires a widow courts her.

and after “optional rise” applies as in DRS-2, we get an object wide scope reading for (ii):

(ii”) There is a widow s.t. every farmer who admires her courts her.

(ii’) seems as a non-existent reading for (ii) (see 2.5), whereas (ii”) diverges from the Q-kataphoric reading for (ii) (see 2.4). Therefore, both ways to interpret (ii) and similar sentences in patient DRT do not produce the Q-kataphoric reading.

Van Deemter’s proposal also handles properly simple cases of BP sentences. For example:

Such sentences in singular are considered in general as cases of weak crossover. However, the same should be true with respect to plural sentences as discussed in section 2.
(iii) A woman who works in his department was interviewed by the manager who hired her.

On the other hand, the same procedure cannot deal with the H&M’s kind of BP sentences like (iii) where the indefinite NP is replaced by a universal QNP. The reason for this problem is the same as in the DRS’s for (ii): If the RM for the kataphoric pronoun in his department is introduced in the restricting DR of the implication as in DRS-1 we get a non-existent “donkey” reading for the BP sentence. Otherwise the RM is introduced in the principal DR, as in DRS-2 and then we are left with an unresolved RM for the pronoun her in a manager who hired her.

As a result, “patient” DRT may handle with non-quantification kataphora, but is not a solution for dealing with quantificational BP sentences and other cases of Q-kataphora.

D.2 Q-kataphora in Dynamic Semantics

The problem of Q-kataphora is basically a problem for common definitions of quantifier scope. It would be interesting to see how Q-kataphora be handled by dynamic semantics for natural language, which are tools especially introduced to handle with other problems concerning quantifier scope (“donkey” sentences and discourse anaphora). A discussion of this point with respect to Groenendijk and Stokhof’s Dynamic Predicate Logic (DPL) may be found in [van Deemter ’90]. Van Deemter’s conclusion is that it is hard to handle kataphora in DPL while retaining its left-to-right interpretation. Another dynamic formalism for the analysis of anaphora, Predicate Logic with Flexibly Binding Operators (PFO), was introduced in [Pagin & Westerstahl ’93]. PFO allows for straightforward translation of BP sentences like:

(i) A pilot who sighted it downed the mig that chased him.

Pagin and Westerstahl also propose a translation constraint that blocks backward quantification in ungrammatical BP sentences with a universal QNP matrix object:

(ii) ? Every pilot who sighted it downed every mig that chased him.

The following grammatical BP sentence is not blocked by the translation procedure into PFO:

(ii) Every pilot who sighted it downed some mig that chased him.

The translation yields the following PFO formula:

\[ (\text{pilot}(x) \land \text{sighted}(x,y)) \rightarrow (\text{mig}(y) \land \text{chased}(y,x) \land \text{downed}(x,y)) \]

This formula is interpreted in PFO as equivalent to the following PC formula (for details see [Pagin & Westerstahl ’93]):

\[ \forall x \forall y [(\text{pilot}(x) \land \text{sighted}(x,y)) \rightarrow (\text{mig}(y) \land \text{chased}(y,x) \land \text{downed}(x,y))] \]

This is clearly not a correct interpretation for (iii). In general, translation to PFO according to the procedure Pagin and Westerstahl propose leads to similar problems with other cases of Q-kataphora.

From the examples of patient DRT and PFO we conclude that Q-kataphora is hard to be handled using common semantic mechanisms that are designed to handle properly “donkey” and discourse anaphora. Consequently, how to interpret sentences with Kataphoric Quantification in a general semantic framework without introducing a special mechanism like Quantifier Absorption is still left as a tough open problem.
References


