DEFINITE DESCRIPTIONS WITH CROSSING COREFERENCE*

A Study of the Bach-Peters Paradox

0. There is an argument, due to Susumo Kuno, Emmon Bach and Stanley Peters, that one or more of the following familiar assumptions is wrong:1

(1) (a) There is a rule of pronominalization that operates on two identical noun phrases.
(b) The rule requires that the noun phrases in question be (i) structurally, (ii) morphemically, and (iii) referentially identical.
(c) Pronominalization is an obligatory cyclic rule.

Together they appear to lead to the unacceptable conclusion that there are sentences which cannot be derived from a finite underlying structure. (2a) would be an example of such a sentence, provided that (i) the pronoun it is understood to refer to the same individual as the object the Mig that chased him and (ii) the pronoun him is coreferential with the subject the pilot who shot at it. The intended surface structure is given in (2b).

(2) (a) The pilot who shot at it hit the Mig that chased him.

* This paper is part of my Ph.D. dissertation Problems of Reference in Syntax (Indiana University, 1969). I want to thank the following persons, who have provided assistance by discussing these matters with me: Fred W. Householder, Emmon Bach, C. LeRoy Baker, Frances Karttunen, Stanley Peters, Gerald Sanders, and Robert E. Wall. It was
Under the assumption (1b), it would seem that the pronoun *it* must come from a full-fledged noun phrase *the Mig that chased him*. But *him* in the latter must in turn come from *the pilot who shot at it*, which again contains an *it* that must come from underlying *the Mig that chased him*, and so on ad infinitum. The structure underlying (2b) would have to be as in (3), where both the subject and the object NP contain an infinite pile of relative clauses.

(3)

earlier circulated under the title *Migs and Pilots* (mimeographed, University of Texas, Austin, 1969).

1 James D. McCawley (1967) attributed the discovery to Kuno, apparently without knowing that Bach and Peters (Bach, 1967) had independently presented the same argument.
There are many reasons to believe that such deep structures should not exist. For example, they are incompatible with the idea of cyclic rule application because there is no deepest embedded S where the application of rules could begin.

1. This argument does not show that any particular assumption in (1) is wrong, but it demonstrates a difficulty which arises from maintaining them all simultaneously. To get out of the trouble, it might be enough to change any one of them. For example, instead of (1b) we might require only referential identity as a condition for pronominalization. But if pronominalization is not a cyclic rule as indeed argued by Lakoff (1968) and Postal (1971), there is also a procedure for deriving (2) without violating any of the assumptions in (1b). Let (4) be the structure underlying (2) just at the point where pronominalization applies.

\begin{align*}
(4) & \quad S \\
& \quad \downarrow \\
& \quad NP^5 \\
& \quad \downarrow \\
& \quad NP \quad S \\
& \quad \downarrow \\
& \quad the\ pilot \quad NP^3 \\
& \quad \downarrow \\
& \quad the\ pilot \quad VP \\
& \quad \downarrow \\
& \quad shot \ at \quad NP \quad S \\
& \quad \downarrow \\
& \quad the\ Mig \quad NP \quad VP \\
& \quad \downarrow \\
& \quad the\ Mig \quad V \quad NP \\
& \quad \downarrow \\
& \quad chased \quad the\ pilot \\
& \quad \downarrow \\
& \quad hit \quad NP \quad S \\
& \quad \downarrow \\
& \quad the\ Mig \quad VP \\
& \quad \downarrow \\
& \quad chased \quad the\ pilot
\end{align*}

Let us assume for the time being that pronominalization is subject to constraints of the kind proposed by Ronald W. Langacker (1966) and John Robert Ross (1967).

The following pair of transformations would do the trick:

\begin{enumerate}
\item[(5)] (a) NP$^1$ pronominalizes NP$^2$,
\item [(b)] NP$^3$ pronominalizes NP$^4$.
\end{enumerate}
The first step is unobjectionable as backwards pronominalization, since NP\textsuperscript{2} is in a subordinate structure that is commanded by NP\textsuperscript{1}. The second step seems like an acceptable instance of forwards pronominalization. But notice that NP\textsuperscript{3}, which serves as antecedent here, is dominated by another NP that is coreferential with it. This is NP\textsuperscript{5}. But although NP\textsuperscript{5} and NP\textsuperscript{3} are both supposed to be coreferential with NP\textsuperscript{4}, only NP\textsuperscript{3} may pronominalize NP\textsuperscript{4}, if structural identity is to be observed. To pronominalize NP\textsuperscript{4}, one has to bypass NP\textsuperscript{5} and pick the antecedent from its relative clause. Thus one might object to (5) by saying that it doesn’t really obey the constraints set down in (1b); it merely shows a way around them. But a general constraint on pronominalization that would rule out such tricks has yet to be proposed.\textsuperscript{2}

A derivation such as (5) is actually compatible even with the principle of cyclic pronominalization, provided that pronominalization is optional at every stage. We only need to assume that (4) reflects the situation on the very last cycle of a derivation where pronominalization has not applied on any previous cycle. On the last cycle it finally applies as specified in (5).

2. Another objection that might be raised is the following. If steps like (5b) are accepted, then there is also another underlying structure from which (2) can be derived much the same way. Namely, (2) could also be derived from (6) by following the steps in (7).

(6) (see page 161)

(7) (a) NP\textsuperscript{1} pronominalizes NP\textsuperscript{2} and
(b) NP\textsuperscript{3} pronominalizes NP\textsuperscript{4}.

The second step in (7) is the same as (5b) in the sense that the antecedent NP is dominated by a presumably coreferential, larger NP that is not structurally identical to the pronominalized NP. Since NP\textsuperscript{4} does not command NP\textsuperscript{3}, (7b) should be acceptable as backwards pronominalization. But (4) and (6) are different structures and, presumably, are associated with different meanings. If they do have different meanings, then the acceptance of step (5b) which makes it possible to derive (2) from both (4) and (6) commits us to the

\textsuperscript{2} For example, it is not possible to disqualify NP\textsuperscript{3} as antecedent simply because it is dominated by another coreferential NP. There are phrases, such as the pilot who shot at the Mig that chased him, where a pronoun is coreferential with the NP that dominates it. Assuming that pronominalization requires structural identity, it is obvious that the noun phrase as a whole cannot be the antecedent of him. The subject noun phrase of the relative clause, that itself turns up as a relative pronoun in surface structure, must be allowed to pronominalize another NP within the same relative clause at some earlier stage. What is suspicious about step (5b) is that NP\textsuperscript{3} is used to pronominalize another NP outside its own relative clause.
view that (2) is ambiguous – a claim that comes as something of a surprise to most people who think they understand this sentence. From (4) we can obviously also get (8a) and, similarly but passivizing, from (6) we get (8b). Therefore, if the sentence

(2a) The pilot who shot at it hit the Mig that chased him.

indeed is ambiguous in this way, we should be able to paraphrase its two senses as

(8) (a) The pilot who shot at the Mig that chased him hit it.
    (b) The Mig that chased the pilot who shot at it was hit by him.

It is by no means immediately clear whether this hypothesis is correct. To make the problem worse, there is also a third underlying structure from which (2) is derivable provided that it is derivable from both (4) and (6). This is the P-marker in (9).

(9) (see page 162)

With (9) as the underlying structure we need two applications of the questionable move of using an NP within a larger coreferential NP as antecedent
in pronominalizing outside that larger NP. But doing it twice shouldn’t be any worse than doing it once.

3. Whether or not the above speculation is correct it demonstrates that, even in the traditional framework, there appear to be ways to avoid raising the spectre of infinitely deep deep structures.\(^3\) While it is true that none of these ways is wholly satisfying, the Kuno–Bach–Peters argument remains far less conclusive than sometimes has been suggested. For example McCawley (1967) – as well as Jackendoff (1968) – presents as an argument for his treatment of pronominalization the fact that the problem of infinite regression does not arise, which presupposes that it could not be avoided in the traditional theory. The value of such an argument hinges on the truth of the presupposition which we have just shown to be questionable.

In McCawley’s grammar, (2) would have the following underlying representation.

\[(10)\]

\(^3\) Another way to circumvent the problem was charted by John Kimball (1967). While Kimball’s proposal does not include a trick such as (5b), it runs into insurmountable difficulties elsewhere.
The two noun phrases are then substituted sequentially for a variable that bears the same subscript. Different ways of substitution may result in different surface structures. By replacing the first term in the proposition by \( NP:x_1 \) and the second variable by the second NP we get, eventually:

(2a) The pilot who shot at it hit the Mig that chased him.

As the example shows, remaining variable symbols become pronouns. If we substitute the second NP not for the \( x_2 \) variable in the proposition but for the \( x_2 \) of the first noun phrase, we get:

(8a) The pilot who shot at the Mig that chased him hit it.

We also have a choice of substituting the first NP for the \( x_2 \) of the second NP rather than for the \( x_2 \) in the proposition. In this case, we have to passivize in order to come out with a well-formed sentence. That sentence is (8b).

(8b) The Mig that chased the pilot who shot at it was hit by him.

If we look at the intermediate structures of (2a), (8a), and (8b) following the attachment of noun phrases in (10) we have (11a), (11b) and (11c), respectively.

In McCawley's theory, the order of attachment does not have any semantic significance. Since (11a–c) are all derived from the same underlying structure, they all should have whatever meaning is represented by (10). Somewhat surprisingly, McCawley is committed to the view that, not only is (2a) a paraphrase of both (8a) and (8b), but that the latter are also paraphrases of each other. The validity of the second claim seems even less clear than that of the first.

(11) (a) – Intermediate representation of (2a)

\[
\begin{align*}
S &\quad NP: x_1 &\quad hit &\quad NP: x_2 \\
&\quad the \ pilot \ who \ shot \ at \ x_2 &\quad &\quad the \ Mig \ that \ chased \ x_1
\end{align*}
\]

(b) – Intermediate representation of (8a)

\[
\begin{align*}
S &\quad NP: x_1 &\quad hit &\quad x_2 \\
&\quad the \ pilot \ who \ shot \ at \ NP: x_2 &\quad &\quad the \ Mig \ that \ chased \ x_1
\end{align*}
\]
In Jackendoff's grammar also there is no rule of pronominalization. Pronouns are generated in the Base as NP's with the feature [+pronoun]. The relation between a pronoun and its antecedent is not marked in the deep structure. At some point in the derivation, there is a cyclic interpretive rule that marks noun phrases as having the same or a different referent. It follows from his approach that the deep structure of (2a) is pretty much the same as its surface structure except that the latter also contains information about coreference. It also follows that there is no particular relation between (2a), (8a), and (8b), which all have different structures, unless it is established by some higher interpretive rules, which Jackendoff does not discuss. But there definitely is no way to relate any two of them to a common underlying representation. If they should happen to have the same meaning, Jackendoff would not be able to account for it as anything more than a mere coincidence.

The three approaches to the Kuno-Bach-Peters problem we have sketched above seem to commit us to three rather different views about the semantic relations between these three sentences.

(2a) The pilot who shot at it hit the Mig that chased him.
(8a) The pilot who shot at the Mig that chased him hit it.
(8b) The Mig that chased the pilot who shot at it was hit by him.

Let us summarize briefly what these positions are. If the attempt to derive (2a) in the traditional framework assuming that there is a rule of pronominalization is correct, it should be the case that (2a) is derivable in at least two ways. On the one hand, (2a) and (8a) may have the same deep structure, on the other hand, (2a) and (8b) can also be derived from the same underlying representation. Therefore, (2a) not only should have the same meaning as (8a) and (8b) but it should be ambiguous provided that (8a) and (8b) have different meanings, which we would expect since their deep structures differ in a way one feels is likely to correspond to a semantic difference.

McCawley and Jackendoff agree in that there is no rule of pronominaliza-
tion. But while McCawley derives (2a), (8a), and (8b) from the same deep structure, in Jackendoff’s grammar they all have a different underlying representation as well as a different surface structure. Since, for McCawley, the deep structure represents the meaning of the sentence, (2a), (8a) and (8b) should have the same meaning. If they do not, this fact would count as evidence against his proposal. Jackendoff’s grammar postulates no semantic relation whatsoever between (2a), (8a), and (8b). If, however, it should turn out that any two of these, in fact, do have the same meaning, this would count as a deficiency in his proposal.

But what are then the correct semantic representations for (2a), (8a), and (8b)? In order to judge the relative merits of the three proposals, we should at least have some independent way of ascertaining whether they are the same or different. Plain intuition seems to be of no particular help in this. On one hand, what all these sentences state is that a certain Mig got hit by a certain pilot and if they all refer to the same Mig and the same pilot, isn’t the meaning the same, since they assert the same thing about the same individuals? That is, (2a), (8a), and (8b) seem reducible to a statement like *Pilot A hit Mig B*, and we have assumed all along that the intended referents are the same. Therefore, it can never be the case, that, say, one of the three is a false proposition while the other two are true. On the other hand, it may be suspected that it is not enough for an adequate semantic representation to take into account only what is being referred to and what is said about it and to disregard the question how it is being referred to. Since (2a), (8a) and (8b) differ slightly in the latter respect, it could still be said that they have different meanings. Although the referents are the same, they are described differently and, perhaps, this makes a difference in meaning.

4. Let us then look at these sentences more carefully and try to see if there are clear instances where there is a difference between them that is due to different ways of characterizing the referents. It seems easier to do that, if we give these sentences some content and get a few real pilots and Migs to talk about. By having some individuals and some information about them we can examine (2a), (8a), and (8b) in relation to a small universe of discourse, which we will call a data base. More formally, let us say that a data base consists of

(12) (a) a set of individuals,  
   (b) a set of predicates, and  
   (c) a set of assertions that specify which predicates hold for which individuals.

4 The term was used in the same sense by J. L. Kuhns (1967).
In all of our examples, the set of individuals and the set of predicates will remain the same. The data bases will differ only with respect to the (12c). The set of individuals shall consist of six members that are named: A, B, C, 1, 2, and 3. The set of predicates contains two one-place predicates, ‘x is a pilot’ and ‘x is a Mig’ and three two-place predicates: ‘x shot at y’, ‘x chased y’, and ‘x hit y’. What also is constant in all our data bases is that the predicate ‘x is a pilot’ is known to be true of A, B, and C; similarly, ‘x is a Mig’ is true of 1, 2, and 3. Therefore, we shall informally refer to our individuals as: Pilot A, Pilot B, and Pilot C, Mig 1, Mig 2, and Mig 3. In specifying our data bases we will only list the remaining propositions.

It is assumed that nothing more is known than what is explicitly stated in the data base. We are going to look at (2a), (8a), and (8b) as sentences that tell us something new about the individuals in the data base. The task is to interpret this new information and to restate it in the same elementary form as the rest of the data base.

The first data base we are going to look at consists of the following sentences.

(13) (a)  – Pilot A shot at Mig 1 and Mig 3.
– Pilot B shot at Mig 1 and Mig 3.
– Pilot C shot at Mig 2 and Mig 3.
– Mig 1 chased Pilot C.
– Mig 2 chased Pilot A.
– Mig 3 chased Pilot B.

Such a list is obviously equivalent to a chart where individuals are represented by points and predicates by arrows. An arrow between two individuals indicates that the predicate holds for them. Thus (13a) and (13b) describe the same data base.

(13) (b) (chased) (shot at)

For each pilot, there is in (13) a unique Mig that chased him. However, it appears that only Pilot B shot at such a Mig. Hence it would be true to say that Pilot B, and no one else, is the pilot who shot at the Mig that chased him. This Mig is obviously Mig 3. Looking at Migs, it seems that, for Mig 1 and Mig 3, there is no unique pilot who shot at the Mig, since both Migs were
shot at by more than one pilot. As far as Mig 1 and Mig 3 are concerned, the phrase *the pilot who shot at it* is inapplicable. The definite article in singular presupposes that there be one and only one such pilot. For Mig 2, there is a unique pilot who shot at the Mig, but it turns out that Mig 2 did not chase that pilot. There is no Mig at all that could be described as *the Mig that chased the pilot who shot at it*. Mig 3 comes closest to matching the description, since it chased *one* of the pilots that shot at it, but he was not *the* pilot.

Given the above facts about data base (13), it is immediately clear that, in this context, sentence

(8a)  The pilot who shot at the Mig that chased him hit it.

must refer to Pilot B and Mig 3. By substituting proper names for the subject and the object NP in (8a) we get

(14)  Pilot B hit Mig 3.

which restates the same fact in simpler terms.

On the other hand, it is also clear that sentence

(8b)  The Mig that chased the pilot who shot at it was hit by him.

is not interpretable in the context of (13), since the subject NP of (8b) fails to identify any individual in the data base. There is no such Mig, and therefore, no pilot either, that we know about. The third sentence,

(2a)  The pilot who shot at it hit the Mig that chased him.

obviously must be interpreted in the same manner as (8a), if it is going to be interpretable at all. (2a) must also refer to *the pilot who shot at the Mig that chased him*. It could not be about *the Mig that chased the pilot who shot at it*, since we do not know of such a Mig.

It turns out that, in the context of (13), (8a) and (2a) are easily interpreted. There are individuals in the data base that match the given descriptions. Furthermore, if all definite noun phrases and pronouns are replaced by corresponding proper names, they both reduce to (14). It is only (8b) that is uninterpretable with this data base. Its inappropriateness is due to the fact that the constraints on the use of definite articles are not met. Obviously, a definite noun phrase that contains a restrictive relative clause, such as *the pilot who shot at the Mig that chased him*, presupposes that (i) there is such an individual, and (ii) there is only one such individual.

---

5 One must be careful not to claim that (8b) is false or meaningless in order to avoid getting mixed in the old philosophical argument about *the present King of France* to which even more space has already been devoted than to the grammaticality of *colorless green ideas sleep furiously*. Given the fact that the sentence is not about any individual in the universe of discourse, it is then another question whether it should be called true, false, or neither.
The second data base is given in (15).

(15) (a) – Pilot A shot at Mig 2.
– Pilot B shot at Mig 1.
– Pilot C shot at Mig 3.
– Mig 1 chased Pilot A and Pilot C.
– Mig 2 chased Pilot A and Pilot C.
– Mig 3 chased Pilot A and Pilot B.

(15) (b) (shot at) (chased)

\[\begin{align*}
A & \rightarrow 1 \rightarrow A \\
B & \rightarrow 2 \rightarrow B \\
C & \rightarrow 3 \\
\end{align*}\]

In (15), Pilot A and Pilot C were chased by more than one Mig. Pilot B was chased by a single Mig, but he did not shoot at that Mig. Therefore, the phrase *the pilot who shot at the Mig that chased him* does not refer to any of the three pilots. For that reason, sentence

(8a) The pilot who shot at the Mig who chased him hit it.

is not interpretable in (15). However, for each Mig, there is in (15) a unique pilot who shot at that Mig. Only Mig 2 chased such a pilot, which makes it *the Mig that chased the pilot who shot at it*. In fact, Mig 2 chased two pilots, but the pilot who shot at it was Pilot A. Sentence

(8b) The Mig that chased the pilot who shot at it was hit by him.

can thus be restated as

(16) Pilot A hit Mig 2.

As for the third sentence,

(2a) The pilot who shot at it hit the Mig that chased him,

it apparently can be interpreted in the same manner as (8b). Since there was initially some doubt about this, a number of linguists at UT were presented with the sentences in (15a) and then asked to restate (2a) in a simpler way. No one failed to come up with (16).
as data base, (2a) must also be understood as referring to Pilot A and Mig 2.

(17) (a) – Pilot A shot at Mig 1.
   – Pilot B shot at Mig 1 and Mig 3.
   – Pilot C shot at Mig 2.
   – Mig 1 chased Pilot B.
   – Mig 2 chased Pilot A and Pilot C.
   – Mig 3 chased Pilot C.

(17) (b)  

\[
\begin{array}{c}
1 \\
2 \\
3 \\
\end{array}
\begin{array}{c}
\rightarrow A \\
\rightarrow B \\
\rightarrow C \\
\end{array}
\begin{array}{c}
\rightarrow 1 \\
\rightarrow 2 \\
\rightarrow 3 \\
\end{array}
\]

Looking at (17), we observe that, for Pilot A and Pilot B, there is a unique Mig that chased the Pilot, but only Pilot B shot at such a Mig. Therefore, Pilot B is the pilot who shot at the Mig that chased him. The Mig that chased Pilot B was Mig 1. Sentence

(8a)   The pilot who shot at the Mig that chased him hit it.

can be restated by

(18)   Pilot B hit Mig 1.

But this time it is also true that there are two Migs, Mig 2 and Mig 3, that were shot at by a unique pilot. Of these Migs, only Mig 2 chased such a pilot, which makes Mig 2 the Mig that chased the pilot who shot at it. Mig 2 was shot at by Pilot C. Sentence

(8b)   The Mig that chased the pilot who shot at it was hit by him.

can be restated by

(19)   Pilot C hit Mig 2.

But which pilot and which Mig are now referred to by

(2a)   The pilot who shot at it hit the Mig that chased him?

It seems that (2a) is open to both interpretations. In interpreting (2a) in the context of (17), one apparently has a choice of (i) finding first the pilot who matches the description the pilot who shot at the Mig that chased him

\footnote{To dispel any doubt, some linguists at UT, who were unaware of the purpose of the experiment, were asked to restate (2a) after reading through (17a). Some responded with (18), some with (19), but no one failed to find the other possibility when asked whether there was more than one answer.}
and letting the identity of the pilot determine the identity of the Mig, or (ii) finding first the Mig that qualifies as the Mig that chased the pilot who shot at it and letting the Mig determine the identity of the pilot. The first approach leads one to believe that the sentence is about Pilot B and Mig 1, the second method singles out Mig 2 and Pilot C.

The fact that (2a) can be understood in two ways in the context of (17) is no problem in a grammar that derives it from two underlying deep structures. But it would seem to be a counterexample for McCawley and Jackendoff who can only derive (2a) in one way. In their case, is it proper to say that (2a) is ambiguous? The answer depends on what one conceives referential indices to be. If the indices are intended as some kind of identifiers that point to individuals in the mind of the speaker, as McCawley seems to think, there are simply two versions of (2a) which are otherwise identical but each version has a different pair of referential indices. One version has indices that point to Pilot B and Mig 1, in the other one, indices are associated with Pilot C and Mig 2. Neither version is ambiguous. But if the function of referential indices is only to mark coreference within the sentence itself, as Jackendoff maintains, the two senses of (2a) are associated with exactly the same surface structure, including the same referential indices. It is true in both cases that the pilot who shot at it is coreferential with him and, similarly, the Mig that chased him with it. Therefore, (2a) must be ambiguous. But notice that this ambiguity is not of the kind in which, say,

(20) The pilot hit the Mig.

could be called ambiguous, because it can be used to refer to different individuals in different contexts. Following Jackendoff, (2a) is ambiguous within the context of (17) itself.

(21) (a) – Pilot A shot at Mig 2.
    – Pilot B shot at Mig 2 and Mig 3.
    – Pilot C shot at Mig 1.
    – Mig 1 chased Pilot A.
    – Mig 2 chased Pilot A and Pilot C.
    – Mig 3 chased Pilot B.

(21) (b)  \begin{figure}
\centering
\begin{tikzpicture}[node distance=1.5cm, auto]
\node (1) {1};
\node (2) [right of=1] {2};
\node (3) [below of=2] {3};
\node (A) [above right of=1] {A};
\node (B) [below right of=2] {B};
\node (C) [below right of=3] {C};
\draw[->] (1) -- (A);
\draw[->] (2) -- (B);
\draw[->] (3) -- (C);
\draw[->] (1) -- (B);
\draw[->] (2) -- (A);
\draw[->] (3) -- (B);
\end{tikzpicture}
\end{figure}
Data base (21) completes our set of examples. In this context, (8a) and (8b) are both interpretable, and it just happens that they also must refer to the same Mig and the same pilot.\(^8\) (2a), (8a), and (8b) can only be restated as

\[
\text{(22) Pilot B hit Mig 3.}
\]

5. In retrospect, it appears that the question 'Do these three sentences have the same meaning or different meanings?', that led to this investigation, was not the best of all possible questions. The words 'same meaning' and 'synonymous' suggest a timeless relation that is independent of any particular discourse. But what we really set out to do was to put these sentences into some very specific contexts. Even in cases where two of the sentences, say (8a) and (8b), must be understood as referring to the same individuals and come out the same when proper names are substituted for descriptive noun phrases, we may not wish to call them 'synonymous'. To describe our results we need a term that is less suggestive and that is restricted to a specific context. Let us define semi-formally a notion of equivalence as follows.

\[
\text{(23) Given a data base } D \text{ and two sentences } S_1 \text{ and } S_2 \text{ that are interpretable in } D, S_1 \text{ and } S_2 \text{ are equivalent in } D \text{ provided that they make the same assertion about the same individuals.}
\]

We also need to state more clearly what is meant by the word interpretable, that has been used throughout the discussion. Let us say that

\[
\text{(24) Given a data base } D \text{ and a sentence } S, S \text{ is not interpretable in } D \text{ unless every definite noun phrase in } S \text{ refers to some individual in } D.
\]

Given the definition in (23), it appears that the stronger notion of synonymy can be treated as a special case of equivalence.

\[
\text{(25) Two sentences are synonymous if and only if they are equivalent in any data base.}
\]

There is no question about the fact that the above definitions are somewhat careless and imprecise, but they seem adequate for the present purpose. Armed with these shiny notions we may now proceed to summarize the findings of our semantic explorations in the form of a single chart.

---

\(^8\) Pilot B and Pilot C were chased by a unique Mig, but only the former shot at such a Mig. This Mig was Mig 3. Similarly, Mig 1 and Mig 3 were shot at by a unique pilot but only Mig 3 chased such a pilot. This pilot was Pilot B.
<table>
<thead>
<tr>
<th>(26)</th>
<th>Data base</th>
<th>Interpretable</th>
<th>Equivalent pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>(13)</td>
<td>(2a), (8a)</td>
<td>(2a, 8a)</td>
<td></td>
</tr>
<tr>
<td>(15)</td>
<td>(2a), (8b)</td>
<td>(2a, 8b)</td>
<td></td>
</tr>
<tr>
<td>(17)</td>
<td>(2a), (8a), (8b)</td>
<td>(2a, 8a), (2a, 8b)</td>
<td></td>
</tr>
<tr>
<td>(21)</td>
<td>(2a), (8a), (8b)</td>
<td>(2a, 8a), (2a, 8b), (8a, 8b)</td>
<td></td>
</tr>
</tbody>
</table>

Although there are data bases, such as (21), where (8a) and (8b) are equivalent, this is not generally the case, as shown by (17) where (8a) and (8b) are both interpretable but not equivalent. Therefore, (8a) and (8b) do not have the same meaning according to the definition (25). But the relation between (8a) and (2a) is different. It is obvious, that, in any data base where (8a) is interpretable, (2a) is also interpretable and equivalent with it. The same seems true of (8b) and (2a). In the case where (8a) and (8b) are both interpretable but not equivalent, (2a) has two possible interpretations that correspond to (8a) and (8b). According to (25), we are justified in saying that one of the two versions of (2a) is synonymous with (8a), the other with (8b). In other words, sentence (2a) represents two different propositions.

It may take a long time to convince an ordinary speaker of English that (2a) is potentially ambiguous, but the difficulty of perceiving the ambiguity does not disprove its existence. The situation with (2a) is very much the same as with McCawley’s example (27).

(27) All but one of the boys danced with all but one of the girls.

For a logician, who knows that the commutativity of universal quantifiers does not extend to a case where all is replaced by all but one, it is immediately obvious that (27) can be translated into predicate calculus in two different ways that are not logically equivalent. For an ordinary speaker, the ambiguity is very difficult to grasp, unless it is demonstrated to him by presenting data bases where (27) is true under one interpretation but not under the other. As with (2a), it is also possible to construct a data base where (27) is true under both interpretations, but the two sets of boys and girls that are involved are different. Assume, for example, that we have three boys, called A, B, and C, and three girls, 1, 2, and 3. The following chart indicates who danced with whom.

---

9 The distinction between equivalence and synonymy drawn here is reminiscent of Carnap’s (1958) distinction between the intension and the extension of an individual expression. In Carnap’s terminology, the two phrases the father of Peter Brown and the mayor of Lexington have different intensions, although their extensions may turn out to be the same. That is, they have different senses, although both may occasionally be used to refer to the same individual. The intension of a sentence Carnap defines as the proposition designated by the sentence. The extension of a sentence, on the other hand, is its truth value. In his terminology, if (8a) and (8b) are true in data base (17), they have the same extension, although they are not equivalent in the sense of (23).

10 Personal communication.
In the context of (28), sentence (27) is true in both of the following senses:

(i) The set *all but one of the boys* consists of B and C. The two girls they danced with are 1 and 3, 2 and 3, respectively.

(ii) The set *all but one of the girls* consists of 2 and 3. The two boys that danced with them are A and C, B and C, respectively.

That there are two ways in which (27) is true with respect to (28) seems as hard to grasp as the fact that there are two possible interpretations of (2a) in (17). The difficulty of becoming aware of the ambiguity is, perhaps, partially explained by the fact that, in some databases, the two interpretations are equivalent. In this respect, (29) is analogous to (21).

That otherwise non-synonymous sentences may state the same thing about the same individuals in certain contexts may sometimes obscure their semantic difference. But that the grammar must provide two interpretations for (2a) and (27) should not be in doubt, even if people often cannot.

6. By demonstrating that

(8a) The pilot who shot at the Mig that chased him hit it.

(8b) The Mig that chased the pilot who shot at it was hit by him.

are not synonymous in spite of their occasional equivalence, we have given a fatal blow to deep structures of the kind presented in McCawley (1967). According to McCawley, both (8a) and (8b) are derived from the same deep structure (10).
(10)

\[ S \]

\[ \text{Proposition} \]
\[ x_1 \text{ hit } x_2 \]
\[ \text{NP: } x_1 \]
\[ \text{the pilot who shot at } x_2 \]
\[ \text{NP: } x_2 \]
\[ \text{the Mig that chased } x_1 \]

Since, for McCawley, deep structure equals semantic representation, all sentences that are derived from (10) should have the same meaning. But this is not true of (8a) and (8b). By failing to distinguish between them, (10) does not adequately represent the meaning of either one of them. Therefore, (10) also represents neither one of the two interpretations of

(2a) The pilot who shot at it hit the Mig that chased him.

Deep structures such as (10) are inspired by symbolic logic. In fact, (10) can easily be translated into ordinary logical notation. The following formula is a result of such translation:

(30) \[ \exists x \exists y \ (\text{hit} \ (x, y) \cdot x = (iz) \ [\text{pilot} \ (z) \cdot \text{shot at} \ (z, y)] \cdot y = (iw) \ [\text{Mig} \ (w) \cdot \text{chased} \ (w, x)] \]

By translating (10) into (30) we can point out clearly where the failure lies. (30) presumes that

(i) there is, for the Mig we are talking about, a unique pilot who shot at it, and also that

(ii) there is, for the pilot we are talking about, a unique Mig that chased him.

As we saw in (21), there are data bases in which both conditions obtain and, in such a case, (8a) and (8b) indeed are equivalent. But there are other contexts where only one of the presuppositions is true and two of the three sentences are still perfectly interpretable. Whatever is represented by (30), it is not the meaning of any of these English sentences.\textsuperscript{11}

However, the following formula seems to capture the essence of (8a) in logical notation:

(31) \[ \exists x \exists y \ (\text{hit} \ (x, y) \cdot x = (iz) \ [\text{pilot} \ (z) \cdot \text{shot at} \ (z, (iw) \ [\text{Mig} \ (w) \cdot \text{chased} \ (w, z)])] \cdot y = (iv) \ [\text{Mig} \ (v) \cdot \text{chased} \ (v, x)] \]

\textsuperscript{11} To translate (30) into ordinary English one needs an expression like \textit{A certain pilot hit a certain Mig and that pilot was the pilot who shot at that Mig and that Mig was the Mig that chased that pilot.}
The corresponding formula for (8b) is (32).

\[(32) \quad \exists x \exists y \text{hit}(x, y) \cdot x = (rz) \text{pilot}(z) \cdot \text{shot at}(z, y) \cdot y = (iw) \cdot \text{Mig}(w) \cdot \text{chased}(w, (v) \text{pilot}(v) \cdot \text{shot at}(v, w))\]

Translating (31) and (32) back into McCawley type trees, we get (33) and (34), respectively.

\[(33)\]

\[(34)\]

But while (33) and (34) seem unassailable on semantic grounds, in one interesting aspect, they defeat McCawley's purpose in assuming that this is what deep structures should look like. There is obviously no way to derive

\[(8a) \quad \text{The pilot who shot at the Mig that chased him hit it.}\]

from (33) without engaging in some kind of pronominalization to generate the pronoun *it*. Similarly, pronominalization is required, if (34) is the deep structure of

\[(8b) \quad \text{The Mig that chased the pilot who shot at it was hit by him.}\]

The constraints on pronominalization will also have to be stated in such a way that

\[(2a) \quad \text{The pilot who shot at it hit the Mig that chased him}\]

can be derived from both (33) and (34). The difficulties in doing this in McCawley's grammar seem quite similar to those encountered in a more traditional version of transformational theory.

The reason why there must be a rule of pronominalization in McCawley's grammar merits some attention. At first, (31) and (32) may seem unneces-
sarily complicated. In (31), for example, the definite description \((iw)\) [Mig \((w)\). chased \((w, z)\)] ‘the Mig that chased \(z\)’ is repeated again as \((iw)\) [Mig \((v)\). chased \((v, x)\)] ‘the Mig that chased \(x\)’. Given the identity of \(z\) and \(x\), the two descriptions cannot fail to refer to the same Mig. One is tempted to think that one of them could be eliminated by a more clever use of variables, especially when variables in predicate calculus are generally used very much the same way as pronouns in natural languages.\(^{12}\) After all, in English we don’t have to say something as clumsy as

\[(35)\]

The pilot who shot at the Mig that chased him hit the Mig that chased him

instead of a shorter but synonymous expression (8a). But, alas, this is just how we have to write it down in logic. The substitution of a variable for \((iw)\) [Mig \((w)\). chased \((w, z)\)] ‘the Mig that chased \(z\)’ in (31), which is what McCawley did, would presuppose that there is a unique pilot who shot at that Mig.\(^{13}\) The second appearance of the same description cannot be avoided, because, in predicate calculus, there is no way to refer back to the first. But in the corresponding English sentence (35), either one of the two occurrences of the Mig that chased him can be replaced by it without changing the meaning. Consequently, if base structures are as close to being formulas in symbolic logic as McCawley has proposed, there must be a rule of pronominalization to take care of all cases where anaphoric pronouns cannot come from underlying variables.

The existence of sentences such as (2a) provides no support for McCawley’s view of deep structures. On the contrary, they raise a new problem. Since (10) does not represent the meaning of either (2a), (8a), or (8b), they should not be derivable from such a structure. There apparently must be some constraints on the substitution of noun phrases for variables to block their derivation or, otherwise, (10) must be ruled out, say, by some restriction on possible configurations of variables in deep structures.

\(^{12}\) The close relation between pronouns and variables has been known to logicians ever since Frege. The extreme point of view is illustrated by the following quote from Geach (1968, p. 112): “For a philosophical theory of reference, then, it is all one whether we consider bound variables or pronouns of the vernacular”. Thus he completely overlooks the fact demonstrated above that pronouns are sometimes used in a way that is not possible with variables in the more restricted syntax of predicate calculus.

\(^{13}\) Notice that, in (8a), one cannot in general replace the phrase the Mig that chased him by some other NP that refers to the same individual. For example, in data base (13), the phrase the pilot who shot at the Mig that chased him refers to Pilot B and Mig 3. However, the substitution of Mig 3 for the Mig that chased him changes this to the pilot who shot at Mig 3 which no longer singles out Pilot B. There is no such unique pilot at all. In Quine’s (1960) terminology, this indicates that, in (8a), the Mig that chased him does not occur in a ‘purely referential position’. The context is ‘referentially opaque’. See Karttunen (1969) for a further discussion of this problem.
7. As mentioned above, in Jackendoff’s grammar, the deep structure of (2a) is essentially the same as its surface structure, except for indices that mark noun phrases as coreferential or non-coreferential with each other. But as we have seen, it is not enough for semantic interpretation just to indicate that, in

(2a) The pilot who shot at it hit the Mig that chased him.

the pronoun it is coreferential with the Mig that chased him while him is coreferential with the pilot who shot at it. This does not tell us whether we should look for the pilot who shot at the Mig that chased him and the Mig that chased him or whether the sentence refers to the Mig that chased the pilot who shot at it and the pilot who shot at it. Without this information, one cannot claim to have a semantic interpretation for (2a).

In order to be adequate in this respect, Jackendoff’s interpretive theory has to be augmented by further interpretive rules that make (2a) synonymous with either (8a) or with (8b). It seems quite possible to invent such rules. But it is worth pointing out that they are needed only in a theory which does not contain a rule of pronominalization and which, therefore, cannot relate (2a) to (8a) or to (8b) transformationally.

8. It would seem, from the above criticism of McCawley and Jackendoff, that what comes closest to being an adequate solution to the Kuno-Bach-Peters problem is the somewhat facetious proposal, in the beginning of this article, to derive (2a) in the traditional transformational framework with a rule of pronominalization. According to that proposal, it is possible to derive (2a) from two sources. One of them is also the deep structure of (8a), the other one is also the deep structure of (8b). This proposal predicts correctly that (2a) should be interpretable in these two ways and it does not make the erroneous claim that (8a) and (8b) are synonymous. So far so good, but unfortunately, accepting the crucial step in pronominalization that is necessary to permit the derivation of (2a) from the desired underlying structures (4) and (6) also makes it possible to derive (2a) from a third underlying P-marker, namely

(9) (see page 178)

From (9) one should also be able to derive, by not pronominalizing, a sentence like

(36) The pilot who shot at the Mig hit the Mig that chased the pilot.

where the Mig is coreferential with the Mig that chased the pilot and the pilot
is coreferential with the pilot who shot at the Mig. (36) is thus identical to

(2a) The pilot who shot at it hit the Mig that chased him.

except that (2a) has undergone pronominalization. If (2a) is derivable from three seemingly quite different deep structures, one would expect it to be, not two, but three ways ambiguous. But apparently there is no such third interpretation that would correspond to (9). But just the same assumptions about pronominalization that made it possible to relate (2a) to (8a) and (8b) also entail that it be derivable from (9). They predict non-existent ambiguities.

Incidentally, sentence (36), which itself should be unambiguous considering that it has a unique deep structure, seems hardly acceptable at all under the suggested coreference relations. For some people, (2a) is already a little difficult to accept, and, for some reason, (36) is much worse.\textsuperscript{14} Suppose that we decide to consider it as grammatical. What kind of pilot and what kind of Mig would (36) refer to? In all likelihood, it would turn out to be ambiguous just as (2a) did – synonymous with (8a) or (8b).

(8a) The pilot who shot at the Mig that chased him hit it.
(8b) The Mig that chased the pilot who shot at it was hit by him.

Although the deep structure of (36) is different from those of (8a) and (8b), its meanings apparently would not be different. We would need interpretive semantic rules of some sort to state these relationships. But this would

\textsuperscript{14} Many people find (36) quite impossible. However, there are some who consider it marginally admissible provided that it is uttered with certain intonation and stress patterns (Fred W. Householder, personal communication).
defeat the argument that

(2a) The pilot who shot at it hit the Mig that chased him.

must be syntactically related to (8a) and (8b) through pronominalization. Then, whatever rules of interpretation are needed for (36) will take care of (2a) as well. If sentences like (36) were acceptable, they would be a very strong argument for Jackendoff's position. However, (37) sounds so anomalous that it probably should be marked as deviant by some new output condition. It is not ruled out by any of the conditions proposed by Lakoff (1968).

Whatever the status of (36) is, the fact that our proposal makes it possible to derive (2a) from (9) must be taken as evidence against it. The advantage that this solution seemed to have over McCawley's and Jackendoff's proposals, therefore, was more apparent than real.

9. Thus far we have shown the following. The Kuno-Bach-Peters argument against pronominalization is far less conclusive than what sometimes has been believed. In particular, it depends crucially on the assumption that pronominalization is an obligatory cyclic rule. Of course, this was generally accepted at the time they presented it. But if pronominalization is not both obligatory and cyclic, there exists a not too implausible procedure for deriving sentences such as

(2a) The pilot who shot at it hit the Mig that chased him.

where it is coreferential with the Mig that chased him and him is coreferential with the pilot who shot at it, without violating the other standard conditions: structural and morphemic identity of the antecedent and the pronominalized NP. This revised version of standard theory even appears superior to some recent proposals by James D. McCawley (1967) and Ray S. Jackendoff (1968), which do not include a rule of pronominalization. The existence of sentences such as (2a) provides no evidence for the correctness of these alternative theories. On the contrary, we were able to show that McCawley's treatment of (2a) commits him to the false claim that the following two sentences are synonymous.

(8a) The pilot who shot at the Mig that chased him hit it.
(8b) The Mig that chased the pilot who shot at it was hit by him.

If deep structures are as similar to propositions in symbolic logic as McCawley assumes, (2a) must be derived by a rule of pronominalization. Jackendoff's interpretive theory was also shown to be deficient in its present form. By failing to establish any connection between (2a) and (8a) or between
(2a) and (8b), it does not provide any adequate semantic interpretation for (2a). The semantic flaw in McCawley's theory and the inadequacy of Jackendoff's proposal are both due to the fact that, in a construction such as (38), the circled NP is not in a purely referential position.\footnote{See Quine 1960, §§30–31.}

(38)

```
NP
  \  / \\
NP \ S /\
   \   / \
    NP_j/\
      NP  \\
       \S/ \\
        NP_i/
```

E.g., the pilot who shot at the Mig that chased him.

It is not the referent of NP_j as such that must be known for the interpretation of the top NP in (38) but the description of this referent. The semantic representation of the top NP is not reducible to

(39)

```
NP
  \  / \\
NP \ S /\
   \   / \
    NP  \\
```

under either McCawley's or Jackendoff's concept of referential indices. That could only be done by distinguishing between indices that actually refer to individuals and indices that merely point to descriptions. Suppose that we want to represent (2a) in the sense in which it is synonymous with (8a) by (2b):

(2b) (see page 181)

This representation is not adequate unless it is made explicit that the index on the circled NP refers, not to the individual described as the Mig that chased him but, to this description itself. On the other hand, the index on the boxed NP must refer, not to the description the pilot who shot at it but, to the individual so described. At present it is not clear how these problems
should be handled, but it is evident that they will remain with us no matter what the fate of pronominalization is as a syntactic rule.

University of Texas, Austin

BIBLIOGRAPHY

Karttunen, Lauri: 1969, ‘Pronouns and Variables’, in Papers From the Fifth Regional Meeting of the Chicago Linguistic Society, Department of Linguistics, University of Chicago.