COREREFERENCE RULES DESCRIBED IN TERMS OF FUNCTIONAL SENTENCE PERSPECTIVE (FSP)

Milan Bily

Part I of this paper gives a survey of existing hypotheses (without referring, however, to such works as Grinder and Postal (1971), Brennan (1971), and Postal (1972b) which, in my opinion, just 'create' theoretical problems in order to use them as arguments in the controversy between interpretativists and supporters of generative semantics); Part II shows a new way to formulate constraints on pronominalization; Appendix A gives a short introduction to the theory of Functional Sentence Perspective, which is generally unknown in the West (except for some distorted fragments presented, for example, by Kuno (1972b) and Hinds (1975a)); in Appendix B some problems concerning English reflexives are discussed.

Part I – Earlier Studies of 'Pronominalization'

The standard transformational theory of pronouns and reflexives (Lees and Klime 1963, Chomsky 1965, Ross 1967, Langeacker 1969) assumes that pronouns originate as fully specified NPs identical to their antecedents (and their 'postcedents') in deep structure. Transformations change these NPs into pronouns on the basis of morphological identity and intended coreference with other NPs. Thus a deep structure which can be roughly represented as 'John$_3$ shaves John$_3$' becomes 'John shaves himself.' (In Chomsky (1955) each NP has an associated index in deep structure and two NPs are coreferential if they have the same index.)
The following constraints on 'pronominalization' are generally accepted:

I) NP^a may not be used to pronominalize NP^p if NP^a and NP^p are elements of separate conjoined structures and NP^p precedes NP^a.

II) NP^a may pronominalize NP^p unless (1) NP^p precedes NP^a; and (2) NP^p commands NP^a.

(Langecker (1969, p. 167). Also Ross (1967) independently arrived at a practically identical condition on 'backward pronominalization'.)

'Command' is a technical term used to indicate the role of 'depth' in an embedded structure. Thus, the node A commands B, if a) neither A nor B dominates the other and b) the S-node that most immediately dominates A also dominates B. In the following structures A commands B and B commands A:

(The symbol 'S' always stands for a clause, the other symbols stand for any other part of an 'S' than 'S'.)

But in the following structures only A commands B, B does not command A:

Sentences (1) and (2) exemplify Constraint I; (3) - (6) exemplify Constraint II. (Asterisk-prefixed sentences are meant not to be able to express coreference between a pronoun and a NP, i.e. nothing is claimed
about their grammaticality. Nearly all example sentences are taken from the authors quoted in the bibliography. Either the original sentences are used or the sentence types - for example some sentences have been changed so that certain semantic abnormalities may be avoided for the sake of naturalness.)

(1) Peter has a lot of talent and he should go far.
(2) He has a lot of talent and Peter should go far.
(3) John left town after he robbed the bank.
(4) He left town after John robbed the bank.
(5) After John robbed the bank, he left town.
(6) After he robbed the bank, John left town.

However, Constraints I and II cannot explain many facts. There have been many attempts to mend these constraints. Lakoff (1968) showed that ‘backward pronominalization’ from a subordinate clause to a main clause (which is against Constraint II) is possible if the pronoun is not the subject of the main clause. Thus, it is possible to say (at least for some speakers):

(A: Why didn’t Peter defend himself?) B: They silenced him every time Peter tried to speak.

Another (patently wrong) attempt was made by Kuno (1972a), using so-called ‘direct discourse analysis’, to explain sentences like (7) (with unstressed ‘him’):

(7) ‘That John was the best boxer in the world was claimed by him.
The deep structure of (7) is said to be something which can be roughly described as ‘John claimed: I am the best boxer in the world’. Since the subject of the embedded sentence is a pronoun from the beginning, there is no possibility to realize it as ‘John’ and there-
fore it is not possible to derive the ungrammatical (7). Unfortunately we can find sentences like (8) - (10), which should be equally wrong according to "direct discourse analysis" but which are not. The fact that (7) is correct with stressed "him" makes things even worse.

(8) That John was the best boxer in the world was never claimed by him.

(9) That John was the best boxer in the world was loudly and repeatedly claimed by him.

(10) That John was the best boxer in the world was claimed by him but nobody would believe such nonsense.

Postal (1972a, p. 48) postulates a so-called Global Constraint on Pronominalization in order to be able to make the difference between sentences like (11), (12) and (13), (14).

(11) Who killed his wife?

(12) It was Peter who killed his wife.

(13) "Who did his wife kill?"

(14) "It was Peter who his wife killed.

"The Wh Constraint

Mark as ill-formed any derivation in which:

a. there are two nominal constituents, A and B, in the input structure of a Wh Movement rule, where: (i) A is a pronoun, (ii) B is a wh form, (iii) A is to the left of B; and

b. the corresponding constituents of A and B in the output structure of the Wh Movement rule, call them A' and B' respectively, are aligned such that B' is to the left of A'; and

c. in the semantic representation, A and B (or, more precisely, their corresponding elements) are marked as stipulated coreferents."
However, there have been serious doubts as to whether there are any linguistic phenomena which need explanation in terms of a global constraint (e.g. Emonds 1973) and the global constraint on pronominalization in particular (Cole 1974). Besides, Cole (1974) shows that Postal’s informants’ judgements of his sentences are not at all fully representative for English speakers.

Another attempt was made by Postal (1971, p. 23) to save Constraint II with the following amendment: ‘Backward pronominalization is banned across a copular verb of referential identity,’ (emphasis in original). This restriction was motivated by the fact that there was no other way to prohibit ‘pronominalization’ in (15) while allowing it in (16):

(15) *What annoyed him was my punching Bill.
(16) It was my punching him that annoyed Bill.

Bickerton (1975) believes that Postal’s constraint is incorrect as it would predict noncoreferentiality for (17). However, Postal would probably claim that the copula in (17) is ‘a predicate set of inclusion’ (as he did for ‘The man who wrote to her was a friend of Mary’s.’). But then it remains to draw a clear borderline between these two possible meanings of a copula, which can be rather difficult, at least in certain cases.

(17) My punching him was Bill’s major gripe.

There are more other ‘bad cases of pronominalization’ for which Constraint II does not suffice (cf., for example, Jackendoff 1972, Kuno 1972b etc.).

As for the above-mentioned Constraint I, it has been neglected by
the majority of linguists. Only Postal (1971, p. 20) shows that one has to distinguish between "true coordination, where the conjuncts are, for instance, reversible without change of meaning, and pseudo-conjunction . . .". This restricts the operation of Constraint I radically, because even 'and' is very seldom a 'pure conjunct' without any causal, temporal or other implications affecting reversibility. Our sentences (1) and (2) cannot, then, be explained with Constraint I either. There is still another problem - there do exist structures of 'true coordination' where Constraint I should work but it does not:

(18) His wife and the woman Peter is living with have just met.
(18) is okay for many (or a majority of) speakers. Hinds (1975b, p. 332) tried to complete Constraint I with " . . . and NP is stressed.", which would make the right prediction about (18) with stressed 'HIS' but it cannot explain the difference between (1) and (2) if 'he' is pronounced with reduced stress in both cases.

The most discussed sentence in papers on 'pronominalization' must be (19). This so-called Bach-Peters paradox (Bach 1970) is hard to explain with the pronominalization hypothesis. One can, of course, do as Karttunen (1971) did, i.e. one can suppose (in order to get rid of the infinite deep structure which would be necessary) that (19) has the following deep structure:

(19) The pilot who shot at it hit the Mig which chased him.
The NP₁ pronominalizes NP₂ and NP₃ pronominalizes NP₄. But there are three possible meanings of this sentence, (Kuroda 1971) for which the transformationalists have to find three different deep structures. (According to the transformationalists every meaning of a given sentence must have a different deep structure.) Unfortunately for the transformationalists, it has been shown by Wasow (1973) that there is an algorithm for constructing an infinite number of deep structures for (19). Therefore transformationalists should discover infinitely many distinctive readings for (19)! The obvious impossibility of this means a hard blow to the transformational hypothesis of pronominalization.

Sentence (19) means either

a) 'The pilot who shot at the Mig that chased him hit it,'  
or b) 'The Mig that chased the pilot who shot at it was hit by him,'  
or c) something which is neither a) nor b) and has the surface structure of (19). (Kuroda (1971) tried to explain this third possibility with reference to a hypothetical world where c) is possible without a) and/or b) being true, but in fact he himself proves, without real-
izing it, that c) is something else than a) and b). It is not necessary to refer to a non-existing 'data base'. The fact that there do exist three meanings of (19) can be shown in the following 'data bases' taken from Kuroda (1971) and Karttunen (1971):

'Data base 1'
plane chased pilot shot at plane

Only the pilot B and the plane 3 qualify for the meaning a), b) and c).

'Data base 2'

Only the pilot A and the plane 2 qualify for the meaning b) and c), but no pilot and plane qualifies for meaning a).

'Data base 3'
plane chased pilot shot at plane

Only the pilot B and the plane 3 qualify for the meaning a) and c), but no plane and pilot qualifies for b).
'Data base 4'

plane chased pilot shot at plane

The pilot B and the plane 1 qualify for both a) and c).
The pilot C and the plane 2 qualify for both b) and c).
As we can see, the meaning c) is sometimes 'synonymous' with a), sometimes with b), sometimes with both of them. a) and b) stand in a sort of hyponymical relation to c).

An attempt to save the pronominalization hypothesis was shown and criticized in Wasow (1975). He discusses the use of so-called bound variables (in sense of McCawley (1968, 1970)) which would make it possible to derive (19) from a deep structure which can be simplified as

\[ [x: \text{the pilot}], [y: \text{the Mig}], [x [\text{shot at } y]] S1 \text{ hit } y[y \text{ chased } x]] S2] S3. \]

He shows that pronominalization in such a deep structure would generate certain ungrammatical sentences and it could not generate all grammatical sentences, because several syntactic rules in English are sensitive to the difference between pronouns and full NPs. If pronouns and their antecedents are derived from such variables, then these rules cannot apply correctly. Wasow gives, among other things, the following argument. If bound variables are accepted, the sentence with the simplified deep structure '[x: some burglars], [x shot a man who discovered [that x were in his house]] S1] S2' will give 'Some burglars shot a man who discovered that they were in his house.' (via substitution of 'x' in accordance with Constraint II) or 'A man who discovered that some burglars were in his house was
shot by them.' Passivization must precede the substitution of 'x' to prevent generation of '*They shot a man who discovered that some burglars were in his house.' As is known, 'there-insertion' in existential sentences is possible only when the NP in question is an indefinite one ('There is the man at the door,'). Wasow examines '*They shot a man who discovered that there were they in his house,' and 'A man who discovered that there were some burglars in his house was shot by them.' On the lowest transformation cycle in the latter sentence, there-insertion is applied, then passivization comes on the next cycle, and after that 'x' is substituted. But if passivization does not occur on the second cycle but only substitution, the former ungrammatical sentence is the unavoidable result; that is, the latter sentence can be generated only if the former ungrammatical sentence can.  

Wasow proposes an interpretative approach to the problem to save McCawley's bound variables. But this is already the method advocated by interpretativists (Dougherty 1969, Jackendoff 1968, 1972) who don't need to postulate such abstract deep structures like those proposed by Wasow. Interpretativists (led by the later works of Chomsky (Chomsky 1970, 1971 etc.) have returned to the position taken by Chomsky (1957). They do not postulate any 'fancy' abstract deep structures, their 'moderate' deep structures are syntactically motivated. They claim that

a) Pronoun forms are inserted into deep structures like any other NPs,
b) Coreferentiality between the full NPs and the pronominal NPs is a semantic judgement on the part of speakers/listeners and thus statements involving coreference shall not appear in the formulation of any transformation.
Describing 'pronominization' amounts to discovering rules of semantic interpretation of the type /NPa/ is α coref /NPb/ +pro/ just in case certain conditions are fulfilled.

d) Deep structures with incorrect forms which cannot get any semantic interpretation are semantically ill-formed.

Among the arguments interpretativists use against transformationalists (besides the Bach – Peters paradox) are these:

a) At least some pronouns must be generated by the base anyway (e.g., in 'She is beautiful,' combined with pointing at the person who was not talked about before).

b) There are NPs like 'the bastard', 'the bum' etc. These 'pronominal epithets' can occur in certain environments where they function more or less as specialized pronouns. ('We asked Tom, but the bastard is too lazy to do anything.') There is no sensible way to describe, in a transformational framework, when a 'pronominalized NP' becomes a pronoun and when it becomes one epithet or another.

c) Dougherty (1969) gives many other examples which are difficult to generate via a pronominalization transformation – e.g., 'Each of Mary's sons hated his brothers,' 'his brothers' cannot be derived from 'Each of Mary's sons' brothers' which would mean that every son hated all other sons and himself.2

(Dougherty has unnecessarily committed himself to a very strong claim: "... the set, Σ_A, of surface structure sentences which contain a proform that is understood anaphorically is a subset of the set Σ_N of surface structure sentences which contain a proform that is not understood anaphorically. ... There is no a priori reason
why this should hold in English, but the fact that it does hold is linguistically significant." (Dougherty 1969, p. 511). There is, indeed, no reason why this should hold and in fact it does not — neither for reflexives, nor for such sentences where the semantics demands obligatory coreference ('Except for her laziness, Mary is an amiable person,'). This has been criticized by McColl (1972), who is for the interpretative approach, and by Postal (1972b), who is against it. But Dougherty's statement just needs a slight reformulation to be correct. If we substitute 'syntactic surface structures' (i.e. an abstraction from the semantics of actual sentences) for Dougherty's 'surface structure sentences' and 'a proform that can be understood' for 'a proform that is understood', the claim is correct. Of course, the semantics of real sentences can determine that certain sentences get only the coreferent or noncoreferent reading.

Jackendoff (1972) expresses coreference in a "table of coreference". Each entry in the table consists of a pair of NPs and one of the relations coreferential or non-coreferential. Every possible pair of NPs in the sentence is included in the table. After the table is completed, it is subjected to so-called well-formedness conditions which determine whether it is consistent both internally and in relation to the rest of semantic representation. Jackendoff's rules for pronominal coreference are supposed to work as follows: the rules apply at the end of each transformational cycle and enter relations between pairs of NPs in the table of coreference. After the last cycle the noncoreferentiality rule, which says that every pair of NPs that have not been related by a rule of coreference will be marked as noncoreferential, ensures that every pair of NPs appear in
the table. Even Jackendoff's coreference rules contain a condition similar to Condition II on pronominalization, although he develops the concept of command from the earlier formulation to a more general one. For Jackendoff, the node A commands B, if a) neither A nor B dominates the other and b) the S-node or the NP-node (i.e., every node that defines a transformation cycle — according to Jackendoff, even NPs do that) that most immediately dominates A also dominates B. This makes it possible to give the right coreference interpretation to (20).

(20) His BROTHER visited John

There are also some other attempts to solve the problem of pronominalization with a combination of interpretative and transformational hypotheses, e.g. Harada - Saito (1972) where, by a mixture of the above-mentioned hypotheses, reflexives come about via transformations, while personal pronouns are generated in deep structure, or Teleman (1970), who proposes to derive pronouns and even other coreferent NPs from coreference-marked dummies in deep structure. These dummies get such surface structure representations as are able to express the intended coreference relations.

Generally speaking, the bulk of the works on pronominalization try to establish syntactic constraints on pronominalization. These basic syntactic constraints (the above-mentioned or some other — e.g. Culicover (1976), who uses the notion 'in construction with' instead of 'command') are then successively complemented (since they do not work) by new syntactic or other constraints. For example, a number of 'exception rules' are formulated for constituents bearing sentence stress. However, while the relations between sentence stress and 'pronominalization' are at times noted (Lakoff 1968, Postal 1971, Akmajian 1973,
Akamajian and Jackendoff 1970, Jackendoff 1972, Hinds 1975b), these amendments lie at the periphery of the study of coreference. They have never made an integrated coreference system; they are usually mere ad hoc patchworks.

One should not forget that the discussed rules of 'pronominalization' can be valid for a sort of 'normal language' only, i.e., a language in its 'basic functions', an abstraction from, for example, the metalinguistic function ('He says that Johnny wants a cake,' is okay as a compound of two languages - the baby language and the metalanguage of the adult speaker,), or the poetic function (A writer can certainly write: 'He thought that Peter behaved stupidly' to indicate that Peter looked at himself from the outside, as a neutral observer.) Restrictions on 'pronominalization' are also valid only for a sort of coreferentiality intended by the speaker. The sentence 'But it was HE who said that Peter was not here.' is okay in this situation: Peter, who I did not know, pretended he was somebody else and said that Peter was absent. I told it later to a friend of mine who knew Peter and the friend pointed Peter out. I objected: But it was HE etc., by which I showed my reluctance to accept that Peter and the person I saw were identical. Then 'HE' and 'Peter' have the same referent, but I do not know it/ refuse to believe it.

Part II - Coreference Rules Described in Terms of Functional Sentence Perspective

In this attempt to show a new means of explaining the problem of coreference, the interpretative theory is combined with the theory of Functional Sentence Perspective (FSP). The aim of this paper is not to present a complete theory of pronominal coreference, e.g. the
question of coordinate sentences and coordinate NPs is avoided and English reflexives are just mentioned in Appendix B as they represent a language specific problem (for example, see Spangler (1970), while 'pronominalization' works practically identically in many (perhaps all) languages, however, an attempt is made to show that it is possible to do without the generally accepted 'precede and command' rule. For those who may be unaccustomed with the theory of FSP, as it has been developed by the Prague school during the last forty years, Appendix A is included in this paper. This paper's bibliography also contains several papers by Firbas, Daneš, Svoboda, Pala, and Dvořáková concerning FSP.

Although FSP belongs to text linguistics, it is possible to speak about the distribution of Communicative Dynanism (CD) even in an isolated sentence which we analyse out of its context. When no context is known, everybody 'reconstructs' a type of context in which the sentence in question could be used. It is always the 'normal' use of an isolated sentence, i.e. the sentence taken at its 'face value' is the sentence uttered with the 'normal', 'unmarked' intonation. That's why only the 'marked' stresses in example sentences are indicated by capital letters, otherwise all sentences are supposed to have 'unmarked' sentence stress. Two otherwise identical sentences with different sentence stresses (i.e. with different theme - rheme structures) are NOT identical sentences. Now and then, generativists realize that certain stresses change coreference possibilities; generally they ignore these differences and analyze the ambiguous written sentences only (i.e. they behave as if the written language were prior to - or, at least, equal with - the spoken language; in fact the written language is a simplification of the spoken one), the consequence of which is that
they may say whether coreference is possible (for at least one way of reading the written sentence), but they have no chance to discover why.

It is not always easy to analyze the gamut of CD in a given sentence. But the intuition and reasoning of linguists can be checked in several objective ways. One of these is a careful translation into a language in which FSP is the leading principle of word-order, e.g. into one of the Slavic languages. The rheme proper (the most rhematic word or phrase) bears also sentence stress. Another method of discovering theme-rheme structure is the method of questions. One can construct series of questions which the given sentence (with the given intonation) suits as an answer. (Except for such general questions like 'What has happened?') The element(s) which must be present in all questions belong to the theme of the analysed sentence and the element which is absent in all the possible questions is the rheme proper.

There have been several attempts to incorporate certain notions as 'theme', 'rheme' (or 'old information', 'new information', or 'presupposition', 'assertion') in the rules of 'pronominalization' (Kuno (1972b) and its enlarged version Kuno (1975), Hinds (1975a), and above all Bickerton (1975) who has completely discarded the old syntactic framework). However, these attempts have not been very successful because their theoretical equipment is insufficient for this aim. Hinds and Kuno (who both retain the old discredited syntactic framework) try to use the theory of FSP in their analyses, but their knowledge of Functional Sentence Perspective is minimal and distorted. For example, Hinds believes that every 'by-phrase' in passive sentences
must be rhematic because of its final or 'near final' position! The fact that most 'agent-nodes' are deleted (even when the agent is known) and that the agent can be expressed by an unstressed pronoun (i.e. by a thematic element) shows the fallacy of such purely mechanical judgement based on sentence linearity only. (Hinds is not alone, however. Since the American generativists started thinking about 'presuppositions' and similar notions, they have produced a lot of incredible statements the fallacy of which is obvious to everybody who knows the theory of FSP. One has to regret that these works begin where the Prague school was about forty years ago instead of using the accumulated results of Czechoslovak linguists.) Of course, nobody can stop Hinds if he wants to use some terms in a way defined by him, but Hinds believes that he applies the Prague theory of FSP, which he praises and makes propaganda for!

One of Hinds' most astonishing statements is "A function (i.e. purpose) of pronominalization is to indicate that the referent [!] of the pronoun is considered thematic material." (Hinds 1975a, p. 91). According to Hinds, 'things in our world' are thematic or rhematic! Even Kuno (1975, p. 280) seems to believe in existence of 'predictable' and 'non-predictable' referents. There may live some 'thematic birds' or 'predictable butterflies' in the USA and Japan but they have never been seen in Central Europe. Kuno's lack of insight into FSP causes the terms he uses to be both numerous (theme, contrastive listing, exhaustive listing, neutral description, predictable theme, unpredictable theme, contrastive theme, known part of sentence, unknown part of sentence, old information, new information etc.) and extremely vaguely defined. ("I can only say that the theme is what the rest of sentence is about." - Kuno 1975, p. 277.) Of course, even the Czecho-
slovak linguists use at times such expressions when trying to popularize the concepts of FSP but one should (e.g. as Firbas has done) make some exact definitions, too. (In FSP terminology it is sufficient to speak about degrees of Communicative Dynamism (CD) as the other terms are used for the sake of convenience only:

'theme' = element(s) carrying a low degree of CD
'theme proper' = element(s) carrying the lowest degree of CD
'rheme' = element(s) carrying a high degree of CD
'rheme proper' = an element carrying the highest degree of CD etc.)

According to Kuno's definitions, it would be possible to claim that 'John' in 'Look, John is coming here.' is the 'theme' (more exactly: 'the unpredictable theme') as 'John' is what the rest of sentence is about. But Kuno suddenly says that this sentence is themeless since "there has been no previous mention of him [= John] or no expectation of his coming . . ." (Kuno 1975, p. 278). On the other hand Kuno (1972b, p. 308) says that "unpredictable themes appear when new topics are introduced". But this would be exactly the case of the above-mentioned sentence! Besides, Kuno (1975, p. 277) states that "The theme must be discourse anaphoric . . ." but it does not have to be "old predictable information". The first part of this statement clashes with the quotation about 'unpredictable themes'. What is meant by 'predictable themes' is mysterious, too. The fact that 'Tom' in the following dialogue is called 'unpredictable contrastive theme' is another example of Kuno's elastic and contradictory definitions:

"Speaker A: I understand that John, Bill, and Tom all teach high school. Speaker B: John does, and Bill does, too, but Tom does not." (Kuno 1975, p. 277). 'Tom' is as much (or as little) 'predictable' in this
context as 'He' is in the dialogue on the same page about which Kuno
claims that it is a 'predictable theme'!

"Speaker A: What does John like? Speaker B: He likes fish."

Both 'Tom' and 'He' are 'predictable' in the same way: If these words
were made unidentifiable when the sentence they belong to were said,
one could guess what should have been there. (The criterion is Kuno's
own.) Otherwise they are not predictable at all. Cf. the possible an-
swers of Speaker B: 'John does, and Bill does, too, but who is the
third guy?', 'Why are you asking me?'. In fact, it is even worse. It
seems impossible even for Kuno to remember what he means with his
vague claims. In his latest version [Kuno 1976], he repeats that 'He
likes' of the above written sentence 'represents old, predictable in-
formation because even if that part of sentence is garbled, it is re-
coverable from the preceding context'. Then he claims again that 'John',
'Bill', and 'Tom' represent 'unpredictable information' according to
the same criterion'. (Kuno 1976 p. 120-121). But suddenly (Kuno 1976,
p. 181) we read:

"Speaker A: Who do you like better, John or Mary?
Speaker B: Between these two, I like him better."

'...the subject of like is him (= John), which is unstressed because
it does not represent unpredictable information' (emphasis by the au-
thor of this paper). If we use Kuno's criterion again ('... if that
part of sentence is garbled ...' etc.) we shall find that 'him'
does represent unpredictable information - the answer could have been
'... I like her better'.

It seems to me - as far as it is possible to say something about a
language one has only a 'second-hand knowledge' about - that the fa-
mous Japanese particles are multifunctional and that’s why their total identification with a sort of FSP markers (which is probably what Kuno does) causes such a chaos. Kuno also uses some examples of very dubious value:

"Speaker A: Who killed Mary?
Speaker B: "With an accomplice that John hired, he killed Mary."
(Kuno 1976, p. 123). The sentence is, of course, at least very strange but it has nothing to do with pronominalization. The following versions would not be better:

Speaker B: ?? With an accomplice that he hired, John killed Mary.
or Speaker B: ?? With an accomplice that Jane hired, John killed Mary.

All these answers do not suit as answers to the question, irrespective of where the pronoun is placed or not.

Kuno (1976) presents ten rules of pronominalization (one of them is the traditional Constraint II), which are largely unrelated and whose validity is impossible to test because of his vague terminology and numerous definitions and ‘counter-definitions’. Some of the rules are quite mysterious and remain unexplained (e.g., "If a given noun phrase in a sentence has a discourse anaphora, pronominalize all but one occurrence of that noun phrase in the sentence." - Kuno (1975, p. 280).), others are unnecessary (e.g., "Do not pronominalize the noun phrase of the exhaustive listing interpretation." - Kuno (1975, p. 280) - which is completed (p. 281) with the possibility to have stressed pronouns of ‘exhaustive listing’. As Kuno’s ‘exhaustive listing’ translated into a more general FSP term is ‘rheme proper’, it is obvious that it cannot be an unstressed pronoun and there is no need to formulate a special rule of ‘exhaustive listing’).
The only statement of Kuno’s which is correct (after having been “translated from ‘Kunoese’ into English” is: All pronouns must have some ante cedents, either mentioned before or ‘given’ in the situation; i.e. so-called backward pronominalization is possible only when the referent is already ‘known’ ( = The speaker supposes that the referent is ‘known’ for the listener. ) Therefore:

(21) Before I could talk to him, a policeman turned away from me.

The indefinite article indicates here that ‘a policeman’ is ‘unknown’ and for the first time under discussion. ‘him’ cannot have as its ‘anteecedent’ somebody who is introduced on the scene first after the pronoun has been used.

Derek Bickerton treatment of ‘pronominalization’ (Bickerton 1975) is radically different from all the above-mentioned works. He has completely rejected the old syntactic rules of pronominalization and tries to use notions similar to those of FSP: "Pronominalization flows bidirectionally, and across sentence boundaries, from presupposed to asserted NP, and between presupposed NP, except where one NP has been presupposed throughout its derivational history and the other has not, in the latter case, pronominalization shall be from the more-consistently to the less-consistently presupposed.” (Bickerton 1975, p. 32-33). However, his instrument, i.e., the sentence bi-partition into ‘presupposition’ and ‘assertion’ (in FSP terms: the context dependent part and the context independent part) is insufficient.

As Bickerton’s hypothesis is based on various constructions which were extremely difficult to explain within the traditional framework (above all the constructions where rhematicity is expressed syntactically – for example the types:
It was + rheme + who + theme (It was PETER who killed his wife.)
What + theme + was + rheme (What annoyed Bill was my punching him.)
he can accommodate these constructions which consist of sharply delimited context-dependent and context-independent parts. However, there are many sentences where such delimitation is not possible, e.g., sentences which consist of context-independent elements only, Bickerton’s rule cannot account for the following examples either:
(What happened with John?)
(22) *He was killed when John tried to escape.
(23) The cops killed him when John tried to escape.
It is impossible to see any difference in ‘presuppositional consistency’ between ‘he’ in the first and ‘him’ in the second sentence.
Bickerton seems to have completely missed the fact that there do exist sentence-stressed ‘asserted antecedents’ (in FSP terminology: antecedents which are the most rhematic elements). For example, our sentence (12) (It was PETER who killed his wife.). There are some dubious cases in his reasoning, too. It is not clear why, e.g., ‘Rhoda’ in ‘Rhoda’s appartment’ is claimed to be asserted and ‘appartment’ is presupposed – Bickerton claims that ‘Rhoda’s appartment’ is derived from ‘The appartment belongs to Rhoda.’, while the more usual interpretation ‘Rhoda has an appartment’ where ‘Rhoda’ is ‘presupposed’ and ‘appartment’ is ‘asserted’ seems as plausible as Bickerton’s claim.

Coreference Rules Stated in Terms of FSP3

A) All pronouns must have antecedents in the preceding linguistic context or the situation.8

B) Coreference (‘pronominalization’), both ‘forward’ and ‘backward’, is possible only when the degree of Communicative Dynamism (CD)
carried by the 'antecedent' is not substantially higher than that carried by the pronoun(s). This rule is subject to the following restrictions:

C) The 'antecedent' can be a rheme proper in sentences of second-instance level (see Appendix A) if coreference is possible in the first-instance level sentence from which the second-instance level sentence is derived.

D) No pronominal rheme proper is allowed before the 'antecedent'.

E) The 'antecedent' is obligatory as the theme proper.

Let's exemplify these rules:

Rule A

A has been already demonstrated in (21), which is repeated here for convenience:

(21) Before I could talk to him, a policeman turned away from me.

Some more examples:

(24) He is stupid.

(25) (Who do you suppose I calmed when I saw him getting mad?) I calmed him before Harry did something else.

(24) would be without meaning if the speaker did not refer to a male person. It is always the speaker/writer who determines whether there is an antecedent, but if the listener/reader cannot find an antecedent in the context/situation, he reacts: (A: He is stupid. B: Who?).

As (25) is preceded by a rhetorical question, the speaker cannot presume that the listener knows the antecedent. As for the following dialogue, it is possible only when the speaker B presumes that A closely connects John with Mary, so that A immediately identifies the referent of the pronoun.

A: Tell me about John. B: (26) Since I despise her, I have not seen
John and Mary for ages.

Rule 8

(3) John left town after he robbed the bank.
(4) *He left town after John robbed the bank.
(5) After John robbed the bank, he left town.
(6) After he robbed the bank, John left town.

(3), (5), and (6) allow coreference as the antecedents/postcedent carry a low degree of CD. It is the linearity of (3) and (5) that determines the low degree of CD. As for (6), even though 'John' belongs to the theme (because of Rule A - 'John' must be context dependent and therefore thematic. If 'John' were context independent, then it could not be coreferent with 'he' which demands the existence of an antecedent). (4) does not allow coreference. The pronominal subject of (4) is definitely the theme proper, while 'John' in the subordinate temporal clause placed after the main clause (= a rhematic subclause) belongs to the rhyme. It is not enough that 'John' carries a low degree of CD within the rhematic part. The difference in degrees of CD is too great, so coreference is impossible.

(27) *It disturbs her that Mary is pregnant.
(28) It disturbs Mary that she is pregnant.
(29) That Mary is pregnant disturbs her.
(30) That she is pregnant disturbs Mary.

As has been shown in Svoboda (1968), there are two possible FSP interpretations of sentences with extraposition. Either the subclause is thematic and VP of the main clause is rhematic, or they are both rhematic and the subclause is the rhyme proper. In (27) the former possibility is out of question because the unstressed 'her' cannot belong to the rhyme, and even the latter interpretation is modified: The subclause is rhematic and 'her' is thematic for the same reason as in the
first case. The thematic pronoun (the theme proper in this case) cannot be coreferent with ‘Mary’ which belongs to the rhyme. As for (28), the second one of the above-mentioned interpretations allows coreference. In (29), pronounced in the ‘normal’, ‘unmarked’ way (with sentence stress on ‘disturbs’), both the subclause (and therefore even ‘Mary’) and the pronoun are thematic; coreference is possible. In (30) either ‘Mary’ is unstressed (and bears a low degree of CD) — then coreference is possible — or it carries sentence stress, then the sentence in question comes from a sort of ‘presupposition’ (a first-instance level sentence) like ‘That she is pregnant disturbs somebody,’ and ‘Mary’ identifies the ‘somebody’ — coreference is possible, too (Rule C and B). (Of course, the first-instance level sentence must have an antecedent for ‘she’ (Rule A), and therefore it demands a context where the antecedent does exist, for example, we have to know that a certain girl/some girls of a limited number is/are disturbed by her/their pregnancy.)

Rule C

Naturally, all the previous sentences may get a lot of various contrastive stresses. Such sentences belong from FSP’s point of view to the second-instance level. The heavily stressed word/phrase is rhe-
ematic and the rest of the sentence is thematic. But even then it is possible to speak about different degrees of CD within the first-in-
stance level sentence from which the second-instance level sentence originates. (And the same is true even for sentences in real contexts.)

(31) John DID leave town after he robbed the bank.

(32) John left town after he robbed the BANK.

(33) JOHN left town after he robbed the bank.
(34) It was JOHN who left town after he robbed the bank.

(35) *HE left town after John robbed the bank.

(36) *He left town AFTER John robbed the bank, etc.

Coreference is possible in (31), (32), (33), and (34) (Rule B and Rule C - for example, (33) and (34) come from a first-instance level sentence which is something like ‘Somebody left town after he robbed the bank.’, where coreference is possible (Rule B), and ‘JOHN’ identifies ‘Somebody.’) In (35), it is not possible to get coreference reading in the first-instance level sentence (‘Somebody left town after John robbed the bank.’), therefore coreference is not possible in the second-instance level sentence, either. (Besides, (35) would have clashed with Rule D, too.) (36) does not allow coreference as there is a great difference in CD between the pronoun and ‘John’ even within the first-instance level sentence (Rule B), therefore coreference is impossible even within the second-instance level sentence (Rule C). The same rules (Rule B and Rule C) determine whether coreference is possible or not in (37) – (40):

(37) *It disturbs HER that Mary is pregnant. [Cf. (35)]

(38) *It disturbs her that Mary IS pregnant. [Cf. (36)]

(39) It disturbs MARY that she is pregnant. [Cf. (33)]

(40) That Mary is pregnant disturbs HER (only).

Coreference is possible in (40), (Rule B)

As it has been mentioned, ‘backward pronominalization into main clause’ is possible at least for some speakers when ‘the resulting pronoun’ is not the subject of the main clause (Lakoff 1968). But this is true in certain cases only, e.g. in (23). The difference in the degree of CD between the pronoun and the NP which is supposed to be coreferent with it, is not so great and that’s why coreference is possible, at least
for those speakers who are more sensitive to the various 'shades' of CD. (According to Firbas' theory, an unstressed pronominal object is thematic but it carries higher degree of CD than a pronominal subject.)

On the other hand, coreference is impossible in (41). 'John' may be stressed or unstressed but it belongs to the part of (41) which carries the highest degree of CD. Coreference is ruled out by Rule B.

(23) The cops killed him when John tried to escape.

(41) *The cops killed him when they tried to arrest John.

Naturally, coreference is out of question in (22) since 'He' is the theme proper and the degree of CD carried by 'John' is in relation to it too high (Rule B).

(22) *He was killed when John tried to escape.

Even (42) - (47) confirm our Rule B:

(42) They are not very clever, these politicians.6

(43) He would have been like a son to both of us if we could have kept

John away from the influence of his friends.6

(44) John said that he was safe.

(45) *He said that John was safe.

(46) He was safe, John said / said John.

(47) *John was safe, he said / said he.

'these politicians' of (42) is a mere repetition of the thematic 'they' regardless of its final position. Of course, one could keep the usual 'precede and command' rule if one claimed that the surface structure of (42) is such that 'they' does not command 'these politicians', but nothing similar can help in (43) and (47). (43) cannot be explained with any other proposed rules, but it is okay according to our Rule B.

'John' in (43) is definitely not too rhematic - the subclause can be interpreted as a sort of parenthetical clause which follows the highly
rhematc 'a son to both of us' and even in the subordinate clause some elements carry higher degree of CD than ‘John’. The ‘strange’ (47) does not allow coreference as ‘he’ is the theme proper both in S45) and (47). The distribution of CD which is determined by the semantics of these sentences, is not changed by the different word orders. ‘he said / said he’ in (47) is a sort of ‘parenthetical theme’ and the difference in the degree of CD between ‘he’ and the rheumatic ‘John’ is too great.

Some more examples:

(48) *He believes that Hubert will win big in ’72.
(49) *Mary told him that Hubert would take it all in ’72.
(50) *We gave him a picture of Hubert.
(51) His mother believes that Hubert will win big in ’72.
(52) Mary told his mother that Hubert would take it all in ’72.
(53) We gave his mother a picture of Hubert.

In (48) – (50) coreference is impossible as ‘he’ resp. ‘him’ carries too low a degree of CD. A little increase of CD carried by the pronoun (sentences (51) – (53)) makes coreference possible. For example, the pronominal subject of (48) is the theme proper; besides, it is (being a pronoun) the theme proper with the lowest possible degree of CD.

‘His’ of (51) belongs to the NP which carries the lowest degree of CD in the sentence, however, ‘His’ carries higher degree of CD than "He" of (48).

As (53) shows, the ‘antecedent’ can carry a relatively high degree of CD if the pronoun carries a high degree of CD, too. (Our Rule B again.)

In (50) coreference is impossible as the difference in CD between the pronominal object (the element with the next lowest degree of CD in (50)) and the ‘antecedent’ which belongs to the rhyme is too great. The
pronoun of (53) carries a little higher degree of CD and coreference becomes possible.

Our rules are valid not only for whole sentences but also for their parts:

(54) *Mary's telling him that Hubert would win was a major mistake.
(55) Mary's telling his mother that Hubert would win was a major mistake.

In relation to the whole sentence, 'Mary's telling him that Hubert would win' is thematic but when coreference possibility is judged within the cycle in question, the result is the same as in (49). Naturally, coreference is possible in (55) as it has been in (52).

The rules which has been exemplified so far are A), B), and C). What about D) and E]? Rule D prohibits coreference in sentences like (56) and (57) (and (35), (37)):

(56) *After HE robbed the bank, John left town.
(57) *The cops killed HIM when John tried to escape.

Rule E is, in fact, just a special case of Rule B. All the previous sentences that begin with pronominal subjects of the main clause exemplify not only Rule B but also Rule E. The pronominal subject of the main clause placed as the first word of the sentence cannot be anything else than the theme proper, but this is prohibited by our Rule E, which says that the difference in the degree of CD between a pronominal which is the theme proper and its 'antecedent' is too great. That's why only (58) is acceptable in the following dialogue and (59) is not:

A: What will John do tomorrow?
B: (58) If he can, John will go to see a movie.
(59) *If John can, he will go to see a movie.
In another context is (59) okay:

A: Will John go to see a movie if he can?
B: (58b) If he can, John will certainly go.

(59b) If John can, he will certainly go.

(The semantics has been slightly changed to get more natural answers.)

Both (58b) and (59b) are okay since 'if X can' and 'X' are equally context dependent (and therefore thematic). In such case we can choose any of the Xs as the 'antecedent'.

As for (43), it is not a counterexample to Rule E. We are talking about 'us' (about our relations to John) and about 'John'. 'John' is not the whole theme proper - it is just a part of it. That's why it can be 'pronominalized' in accordance with Rule B.

Rule E can also explain the 'strange' sentences with topicalized adverbials:

(60) In her apartment, Mary smokes pot.
(61) "In Mary's apartment, she smokes pot.
(62) In Mary's apartment, her brother smokes pot.
(63) ?? In Mary's brand new apartment in Detroit, she smokes pot.
(64) Near her, Mary found a snake.
(65) "Near Mary, she found a snake.

(60) is okay since 'Mary' is the theme proper. (61) does not allow coreference as it is impossible to get some other theme proper than 'she'. Neither 'Mary' which belongs to the adverbial that carries higher CD than 'she' nor 'apartment' can be the theme proper. As for (62), any of the two genitives suits as the theme proper. (Even 'In her apartment, Mary's brother smokes pot.' is okay.) (63) is acceptable for some people because the structure of the long adverbial phrase strongly resembles that of a clause, where 'Mary' can become
the theme proper of the whole sentence. The explanation of (60) and (61) is valid even for (64) and (65). The adverbial carries higher degree of CD than the subject, which is the theme proper. (It is obvious that 'Near' does not suit as the theme proper, either.)

Some 'Bad Cases of Pronominalization'
The above-mentioned 'direct discourse analysis' of (7) (Kuno 1972a) was refuted by Hinds (1975a). The invalidity of Kuno's claim was proved by (8), (9), and (10). According to Hinds, 'pronominalization' is impossible when passivation had been used to indicate that a NP is 'rhematic'. Quotation marks are used because, as it has been mentioned already, Hinds' concept of theme and rheme has only a certain superficial similarity with the terminology of Prague school. However, Hinds is wrong even within his own theory, as (66) shows. What is wrong with (7) when (8), (9), (10), (65), (67), and even (60) are okay?

(7) That John was the best boxer in the world was claimed by him.
(8) That John was the best boxer in the world was never claimed by him.
(9) That John was the best boxer in the world was loudly and repeatedly claimed by him.
(10) That John was the best boxer in the world was claimed by him but nobody would believe such nonsense.

(65) The women who rejected Peter is hated by him.
(67) That John was the best boxer of the world was claimed by HIM.
(68) That John was the best boxer of the world was denied by him.

(7) is bad because the theme-rheme structure it indicates does not make sense. The subordinate clause is thematic and the unstressed pronoun in the by-phrase is thematic, too. The only possible thematic element is the remaining participle and copula. But it seems difficult (without contrastive stresses, i.e. without reading the sentence in question as
belonging to the second-instance level - for example: That John was the best boxer in the world was claimed by him.) to find a sensible context where the copule or the participle could be the rheme proper. This has nothing to do with coreference. (69) sounds as bad as (7). (69) ?? That John was the best boxer in the world was claimed by her
(with 'her' unstressed).

(8), (9), (10), (66), (67), and (68) are okay since the VPs are made naturally rhematic with the rhematic adverb 'never', with the modification of the verb with 'loudly and repeatedly' etc. (67) has the rhematic 'Him'. (66) and (68) contain participles which can be rhematic in a natural way – e.g. ‘deny’ implies negation and negation is usually the most rhematic part of a sentence.

Our rules can also explain sentences (70) and (71). It is usually said that coreference is possible in (70) and impossible in (71). It is, in fact, true only for the most usual ways of reading of these two sentences, i.e. (20) resp. (77). (70) can be pronounced, for example, as (20), (72), (73). (80) can be pronounced as (74) – (77).

(70) His brother visited John.
(71) His brother was visited by John.
(20) His BROTHER visited John.
(72) *HIS brother visited John.
(73) His brother visited JOHN.
(74) His BROTHER was visited by John.
(75) *HIS brother was visited by John.
(76) His brother WAS visited by John.
(77) His brother was visited by JOHN.

Everything is as our rules predict. The ‘antecedent’ carrying a relatively low degree of CD makes coreference possible in (20), (74), (75),
and coreference is impossible in (72) and (75), where our rule D prohibits it. As for (73) and (77), which are not acceptable 'normally', they become acceptable in a context in which the issue is whether John visited his brother or his brother him. In such a context (i.e. as sentences of second instance level), (73) and (77) allow coreference (Rule C). Finally there are two types of sentences for which Postal claims the existence of his Global Constraint on Pronominalization and the restriction prohibiting 'backward pronominalization' across a copular verb of referential identity.

(78) Who hates his wife?
(79) *Who does his wife hate?
(80) Who is hated by his wife?
(81) It was Peter who killed his wife.
(82) *It was Peter who his wife killed.
(83) It was Peter who was killed by his wife.
(84) His wife hates who?
(85) It was my punching Bill that annoyed Bill.
(86) What annoyed him was my punching him.
(78), (80), (81), and (83) allow coreference. These sentences are second-instance level sentences coming from 'presuppositions' like 'Somebody hates his wife', 'Somebody is hated by his wife' etc. (Rule C and Rule B). In (79) and (82), coreference is impossible since the first instance level sentences, which are of following types: 'His wife hates somebody' and 'His wife killed somebody', do not allow coreference. These 'presuppositions' are wrong as the given 'wife' is falsely 'specified' as belonging to an unspecified person (Rule A).
(87) does not allow coreference, either. The first-instance level sentence would be 'He was bitten by somebody's dog,' which cannot contain any intresentential coreference. 'He' must have some earlier antecedent (Rule A) and the known antecedent cannot be coreferent with the unspecified 'somebody'.

(87) *He was bitten by PETER's dog.

One should be conscious of the difference between the semantic rules of coreference and the possibilities given by the semantics of a given sentence and our knowledge of the world. For example, (77) is okay, while (88), which has the same theme-rheme structure (as well as deep structure and surface structure) does not allow coreference since it is practically impossible to imagine a context (a first-instance level sentence) where the question is who was eaten by whom. If we succeed in imagining a suitable context, coreference becomes possible.

(88) ? Their keepers were eaten by the TIGERS.

As for the possible reading of (78) - (80) and even (84) as incredulity questions, the coreference interpretation is possible, probably for the majority of speakers, because such questions do not correspond with their distribution of CD to the statements they are 'echoes' to, 'Who' in the incredulity questions does not demand an answer, such a 'who' stands for an already known NP and the speaker just shows that he is surprised that the statement which has been made is valid for the NP in question. Naturally, there may be speakers who evaluate the theme-rheme structure of such sentences according to some more 'formal' criteria ('who' in all questions is rhematic for them), and for them (79) is wrong as an incredulity question, too, not to mention (84) which is 'completely wrong', the obligatory placing of 'who' at the beginning of the sentence and Do-Insertion not having been accomplished. (In fact,
there is another sentence type which is wrong for some speakers: 'This is the man who the fact that he had cancer surprised.' The possibility or impossibility of coreference depends again on the speaker's evaluation of the degree of CO carried by the relative 'who' - Rule B again.)

In (16), the syntactic construction guarantees that 'Bill' is thematic and coreference is possible (Rule B), except for (16) with sentence stress on 'him' (Rule D). In (85), coreference is possible if sentence stress is placed on 'my' or 'punching', (Rule B), but if it is placed on 'Bill', coreference is impossible (Rule C). (15) prohibits coreference as our Rule B predicts. (86) allows coreference as 'Bill' belongs to the theme (Rule B).

Conclusion

The rules that have been presented and exemplified in this paper can do without the whole heterogenous collection of rules one was forced to accept otherwise. The new rules are simpler, fewer and homogenous. They are also interesting because they show that the theory of FSP, which is ignored by the majority of linguists (or - if accepted - which is 'man-handled' in unbelievable ways), is an instrument necessary for explanation of various linguistic phenomena. Even the validity of the analytical procedures used in FSP is indirectly confirmed by their use which enables us to make correct predictions concerning coreference relations.

Just to name some other cases where FSP can explain certain phenomena which cannot be understood otherwise, or which are usually 'explained' by placing a 'name label' on them: Fillmore (1970) gives the following sentences and cannot do more than note that the last one is ungrammatical.
(89) An oak developed out of every acorn.

(90) Every acorn developed into an oak.

(91) Every oak developed out of an acorn.

(92) *An acorn developed into every oak.

The indefinite article of a thematic subject means either 'one of the', or 'a single', or it can signal a generic noun. None of these meanings fits in (92); it is impossible (because of the semantics of (92)) to interpret 'An acorn as generic, or 'one of the acorns', or 'a single acorn'. It is quite unnecessary to try to speak about different 'scopes' of the quantifiers.

A similar problem was brought up by Lakoff (1970)

(93) Johan and Max saw an explosion.

(94) An explosion was seen by John and Max.

(93) means either 'John and Max saw an explosion and it was the same explosion', or 'John saw an explosion and Max saw another one'. (94) is said to mean only 'John and Max saw an explosion and it was the same one.' The difference in possible meanings of (93) and (94) is usually explained with different 'scopes' of the indefinite articles in the active and passive sentence. But this is not true. If (94) is pronounced with sentence stress on 'explosion', it can be interpreted in both ways. Then, those who believe in 'scopes' would have to formulate an additional rule for certain sentence stresses which can change 'scopes'. The indefinite article of the thematic subject of (94) cannot obviously mean that the 'explosion' was generic and the other two meanings ('a single' and 'one of the') limit the number of explosions to one. However, if 'explosion' of (94) carries sentence stress, it cannot be thematic and the indefinite article with a rhe- matic noun means 'some', 'an unspecified'. That's why both interpre-
tations are then possible. The same is true even for (93). The rhematic NP with the indefinite article allows both interpretations. And vice versa - (95) can mean only 'John and Max saw an explosion and it was the same one,' as the thematic 'explosion' with the indefinite article must mean 'a single explosion' or 'one of the explosions'.

(95) JOHN and MAX saw an explosion.

To sum up: there may be reasons to formulate rules of the semantic interpretation of the sentences above in terms of 'scopes', but only if these are based on theme-rheme structure, not on sentence linearity only (as the generally accepted mechanical 'scopes' are).

Appendix A

"The concept of communicative dynamism is based on the fact that linguistic communication is not a static, but a dynamic phenomenon. By CD I understand a property of communication, displayed in the course of the development of the information to be conveyed and consisting in advancing this development. By the degree or amount of CD carried by a linguistic element, I understand the relative extent to which the element contributes to the development of the communication, to which, as it were, it 'pushes the communication forward'. Thus, if examined in its unmarked use, the sentence He was cross could be interpreted in regard to the degrees of CD as follows. The lowest degree of CD is carried by He, the highest by cross, the degree carried by was ranking between them."

(Firbas 1971, p. 135-136)

"We hold that between the comparatively least important element, the theme proper, and the comparatively most important element, the rheme proper, one can observe a gamut of degrees of varying importance, of varying communicative value, of varying CD. It is, of course, not always easy to draw an exact dividing line between the transition and the theme on the one hand, and the transition and the rheme on the
other. In such cases it is necessary to attempt at least a correct estimate of the relative importance of the elements composing the analysed structure."
(Firbas 1959, p. 42)

"In this connection it is worth notice that in some cases even a thematic element may contribute considerably to the development of discourse. Thus in a sentence of the type A girl broke a vase, the thematic subject carries a comparatively high degree of CD (the non-generic indefinite article marking out a new idea). Yet as the other elements are more dynamic still, the subject is felt to be thematic owing to the pressure exerted by the basic distribution of CD. This means that the fact of the theme carrying the least amount of CD does not preclude the possibility of its carrying a new piece of information."
(Firbas 1959, p. 42-43)

"I believe that much valuable light can be thrown on the function of language in the very act of communication by a consistent inquiry into the laws determining the DISTRIBUTION of degrees of CD over linguistic elements capable of carrying them.

The following note will be relevant here. It will answer the question of what linguistic elements can become bearers of degrees of CD. I agree with D.S. Worth that a linguistic element - sentence, noun phrase, word, morpheme, submorphemic segment, etc. - may be singled out in order to establish a sharp ad hoc opposition (contrast): John WAS winning. Jenda Vitézil. The fact that Vitézil may appear in sharp ad hoc contrast to vitézil ('is winning') shows that the element -il may become the actual bearer of the contrast. Under the circumstances it would be the only element conveying new information (and therefore be contextually independent), whereas all the other elements would convey known information (and therefore be contextually dependent). Normally, the element -il would not carry the highest degree of CD as it does in the highly marked example under discussion; but even then, on account of its semantic content, it would have to be regarded as a carrier of CD. All linguistic elements, including morphemes and submorphemic exponents, are capable of carrying degrees of CD, as long as they convey some meaning. It may have been gathered from the above note that elements become con-
textually dependent and in consequence carriers of the lowest degrees of CD owing to the operation of the context. They assume this function irrespective of the positions they occupy within the linear arrangement. (I avoid the term 'word order' here, because words are not the only elements concerned.) Strictly speaking, contextual dependence or independence is determined by what I have called the narrow scene, i.e. in fact the very purpose of the communication. Thus in the sentence

John has gone up to the window, the window may be well known from the preceding context, but the purpose of the communication being the expression of the direction of the movement, the window necessarily appears contextually independent. Under the circumstances, it is - to use Halliday’s appropriate terms - non-deriveable, non-recoverable from the preceding context.

Let me now turn to contextually independent elements. In determining their degrees of CD, two further factors are in play: (i) the semantic structure, (ii) positions of the elements within the linear arrangement. By the semantic structure of a sentence I understand the semantic contents of the sentence elements and the semantic relations into which they enter.

I will first illustrate the operation of the semantic structure. An object expresses the goal (outcome) of an action conveyed by the accompanying verb. Provided it is contextually independent, it will carry a higher degree of CD than the verb. This is because from the point of view of communication, an unknown goal (outcome) of an action appears to be more important than the action itself. A contextually independent object will carry a higher degree of CD than the verb irrespective of the positions occupied within the linear arrangement. (I have read a fine book, Ich habe ein schönes Buch gelesen. Četl jsem pěknou knihu.)

Similarly, a contextually independent adverbial element of place expressing the direction or destination of a motion will exceed in CD a verb expressing the motion. This is because, communicatively speaking, an unknown direction or destination of a motion is more important than the motion itself. (I flew to London, Ich flog nach London. Ich bin nach London geflogen. Letěl jsem do Londýna.)

As to the subjects of the two structures discussed in the two preceding paragraphs, each will carry the lowest degree of CD within the sem-
tence provided at least one of the remaining two elements is contextually independent. This is understandable, for a known or unknown agent expressed by the subject appears to be communicatively less important than an unknown action expressed by the verb and/or an unknown goal (expressed by the object or the adverbial element of place) at or towards which the action is directed. (Cf. the examples offered above and also A girl was reading an interesting book, A girl was travelling to an unknown town, Ein Mädchen hat ein interessantes Buch gelesen, Ein Mädchen ist nach einer unbekannten Stadt gereist.)

The situation would be different if the subject expressed a person or thing 'existing' or 'appearing on the scene' and if it were accompanied by a verb expressing the notion of 'appearance' or 'existence on the scene'. If under these circumstances it is contextually independent, the subject will carry a higher degree of CD than the verb. This is because, communicatively speaking, an unknown person or thing appearing on the scene is found to be more important then the fact of existence or act of appearing itself. This holds good irrespective of the positions occupied by the respective elements within the sentence. (A girl came into the room, Ein Mädchen kam ins Zimmer, Ins Zimmer kam ein Mädchen, Ins Zimmer ist ein Mädchen gekommen, Do pokoj vešla dívka.)

Not all semantic contents and relations, however, are capable of signalling degrees of CD in the way indicated above. There are evidently also such types of semantic content as let the linear arrangement itself determine the degrees of CD. Thus a contextually independent infinitive of purpose carries a lower degree of CD when occurring initially than when occurring finally (In order to see him, he went to Prague, He went to Prague in order to see him). Similarly, with the indirect and direct object, provided they are contextually independent, the one coming later within the linear arrangement carries a higher degree of CD. (He gave a boy an apple, He gave an apple to a boy.)

The notes that have just been offered have shown that the distribution of degrees of CD over the sentence elements (the signalling of the degrees) is an outcome of an interplay of three factors: context, semantic structure, linear arrangement. The notes have also indicated that according to the contextual situation, in other words, the contextual dependence, the distribution of degrees of CD over a sentence struc-
ture may vary. All the possible variations (realizations of the distribution) constitute the contextual applicability of the sentence structure. Before proceeding further, let me insert a note on the linear arrangement.

Sentence linearity is an indisputable fact. It makes the speaker/writer arrange the linguistic elements in a linear sequence, in a line, and develop the discourse step by step. I believe to be right in assuming that the most natural way of such gradual development is to begin at the beginning and proceed in steady progression, by degrees, towards the fulfilment of the communicative purpose of the discourse. If this assumption is correct, then a sequence showing a gradual rise in degrees of CD (i.e., starting with the lowest degree and gradually passing on to the highest degree) can be regarded as displaying the basic distribution of CD. I also believe to be right in assuming that this conclusion is quite in harmony with the character of human apprehension.

On the other hand, it seems to be equally in accordance with the character of human apprehension that in a discourse made up of a longer string of verbal sentences, the basic distribution of CD in the fullest sense of the word (i.e., one throughout which, gradually, every element becomes a carrier of a higher degree of CD than its predecessor) can practically never be accomplished. Within such a distribution every element would convey new information. But this is not the way the discourse is structured. In order not to jeopardize comprehension, the discourse is continually interspersed with elements conveying information known (derivable, recoverable) from the previous context, i.e., by elements that have been designated here as contextually cendent. It is through these elements that relief from the steady flow of new information is constantly provided.

The field within which the distribution of CD takes place is naturally the entire discourse. This field, which may often become very extensive, is subdivided into fields of lower rank, provided, e.g., by chapters, paragraphs, sentences, subordinate clauses. Remaining within the sphere of complex sentences and structures ranking below them, I subscribe to A. Svoboda’s view that distributional fields are provided by grammatical structures that convey either explicit (open) or implicit (hidden) predication. (Under the latter heading come structures formed by headwords and their accompanying attributive words or phrases.) In provid-
ing distributional (communicative) fields, grammatical structure cuts, as it were, longer or shorter sections out of the linear flow of the discourse. As may have been gathered from what has been said before, according to their contextual dependence distributional fields may function in different perspectives. As the sentence is not the only type of structure providing a distributional field, it is - strictly speaking - possible to speak of other kinds of functional perspective than that of a sentence, e.g., that of a subordinate clause or that of an attributive construction.

The mutual relations between context, semantic structure and linear arrangement may be summed up as follows. Determining the contextual dependence of the distributional field, the context overrules the semantic structure and the linear arrangement in the interplay of means signalling the degrees of CD. It will be remembered that the semantic structure and the linear arrangement (in other words, the basic distribution of CD) can effectively perform the signalling function only within that section of the distributional field which has remained unaffected by the context, i.e. that section which has remained contextually independent."

(Firbas 1971, p. 136-139)

"Semantic structure ceases to operate in what we have termed after D.L. Bolinger second instance sentences, i.e. in such as contain one element singled out for special attention (usually for the sake of heavy contrast) and functioning as one-element rheme proper, all other elements forming an extensive theme proper. Any element can become rheme proper within second instance [HE wrote an interesting book]."

(Firbas 1966, p. 241)

"... we maintain that the function of the sentence in the act of communication can be successfully interpreted if three levels are kept separate: those of the semantic and the grammatical structure of the sentence and that of FSP. As we see it, the theory of FSP makes it possible to understand how the semantic and the grammatical structure of the sentence function at the very act of communication, i.e., at the moment they are called upon to convey some extra-linguistic reality reflected by thought and are to appear in an adequate kind of perspective."

(Firbas 1966, p. 241)
It is important to stress the great difference between the Prague school's FSP and the use of the terms 'theme' and 'rheme' by Halliday (1967-68) or Enkvist (1974), whose 'themes' and 'rhemes' are synonymous with initial resp. final sentence elements. The Prague concept of CD is a generalization and an abstraction from analyses of concrete texts, where it is possible to divide every sentence into the known ('given', context dependent) part and the unknown ('new', context independent) part, a generalization and an abstraction which is valid even for the initial sentence of a discourse, where all information may be 'new' ("Once upon a time there was a king."), and isolated sentences without any known context.

Appendix B (Two Remarks on Reflexives)

It is rather difficult to discuss the earlier attempts to explain English reflexives since nearly all crucial examples in Jackendoff (1972) or Postal (1971) are semantically dubious - e.g. 'I sold the slave myself,' or 'I bought the slave for himself.' Let's take just one reasonable 'sample':

*John was shaved by himself, (with unstressed 'himself')

Postal (1971) proposes for sentences like this a constraint called the Cross-Over Principle, which says roughly that a transformation cannot move an NP over another NP with which it is coreferent. However, this cannot explain why the same sentence is okay with stressed 'himself' and Postal has to claim that the above-mentioned sentence with stressed 'himself' has the following 'fancy' Deep Structure:

[the one [one shaved John]S1 was John]S2

Of course, it is impossible to prove that a theory which uses such powerful instruments as global constraints and which can postulate a new
abstract deep structure every time it meets a problem, is wrong, but it would be nice to do without such things.

Jackendoff's solution is based on his so-called Thematic Hierarchy Condition on Reflexives (this has nothing to do with the notion of 'theme' of FSP) which operates with semantic roles vaguely reminiscent of Fillmore's 'cases'. As another version of Jackendoff's Thematic Hierarchy Condition (on the passive transformation) is wrong (as Gae (1974) has shown) it would be rather optimistic to believe that this one does work, and Jackendoff does not make any attempt to explain why there should be any difference between the effect of his condition on the sentence with an unstressed and a stressed 'himself', either.

I believe that even the behaviour of reflexives can be explained within FSP. Our sentence with unstressed 'himself' is ungrammatical since it makes no sense from the point of view of FSP: The deep structure 'Himself shaved John' (which is, of course, wrong and which can be 'saved' by a passive transformation 'triggered' by the needs of FSP) cannot become the above-mentioned sentence with unstressed (and therefore thematic) 'himself' which would annul the result of the passive transformation. On the other hand, the stressed (and therefore thematic) 'himself' does not clash with the needs of FSP, which triggered the passivization, and the sentence with stressed 'himself' is okay.

Harada - Saito 1971 has shown that there exist sentences which cannot get the right reflexive interpretation in the cyclical way Jackendoff 1972 proposes. In 'John believes himself to be hard for Bill to understand.', 'himself' must get coreferent with 'Bill' in the lower S it comes from and the correct coreferent interpretation (John + coref. himself) is blocked. But it is the same thing with all NP's which should be interpreted as coreferent but cannot be in Jackendoff's way, as Jackendoff
reasons about the cases of 'interpretatively asserted coreference' only and ignores the cases of 'sententially asserted coreference' (e.g. The morning-star and Venus are the same celestial body.) and 'factual coreference' (coreference based on our knowledge of the world, which makes it possible to 'mark as coreferent' e.g. the following NP's: 'the infamous American president at the beginning of 1970's' and 'Richard M. Nixon').

The solution may be in running the interpretation in cycles parallel to the syntactic transformational cycles with the possibility of 'loops' - e.g., reflexives would be first interpreted on their respective cycles and then, if they are raised, on the next cycle. If the reading on a higher cycle clashes with that obtained on the lower cycle the latter is annulled. This solves the problem presented in Harada - Saito 1971. Some kind of 'reinterpreted rule' is necessary even for the cases of 'sententially asserted coreference' and 'factual coreference' where the NPs in question become first marked as noncoreferent by the rules of 'interpretatively asserted coreference'.
Notes

This paper is an enlarged and revised version of ‘Pronominalization Rules’ published in Papers from the Third Scandinavian Conference of Linguistics, ed. by F. Karlsson, Turku 1976.

1 Wasow (1975) argues against the proposal of Harman (1972) who answers in Harman (1976). However, Harman (1976) avoids the argument presented above and his reasoning is far from convincing. For example, one of Wasow’s arguments is: It is well known that the operation of Dative Movement depends on whether the direct object is pronominal. (‘He gave me it. vs. He gave me the book.’) So Dative Movement must follow NP Placement. On the other hand, Passive can follow Dative Movement; and, according to Harman’s proposal, passivization must precede the substitution of variables (NP Placement) — cf. the example with ‘burglars’. These assumptions are, of course, jointly inconsistent. To save his theory, Harman (1976) is forced to ‘cheat’ in the following way:

a) \( x \) gave \( y \) to \( z \)

b) \( x \) gave NP to \( z \) (= NP Placement of \textit{one of the variables only})

c) \( x \) gave NP to \( z \) (= \textit{Dative Movement})

d) \( z \) was given NP by \( x \) (= \textit{Passive which moves only variables})

e) NP was given NP by NP (= NP Placement of the remaining variables).

This is no honest solution and it does not work either, it can save the crucial examples with ‘burglars’ only if the idea of transformation cycle is given up and there is no sensible way to generate sentences like ‘The woman he loved deceived John,’ as the pivotal point of Harman’s proposal is that the leftmost occurrence of a variable is substituted by a NP while all other occurrences of that variable get pronominalized. However, Harman’s proposal was quite revolutionary, as it was the first attempt to get rid of the ‘precede and command’ rule. For example, NP Placement applies to underlying structure of (3) to yield (3):

f) \([x\text{ left town } after x\text{ robbed the bank}] S1\) \( S2, [x; John] \)

(3) John left town after he robbed the bank.

Adverbial Clause Movement applies to (3) to yield (6):

(6) After he robbed the bank, John left town.

Alternatively, Adverbial Clause Movement applies to (6) yielding
g) [After x robbed the bank, x left town], [x:John]
and NP Placement follows:
(S) After John robbed the bank, he left town.
There is, however, no way to get (4):
(54) He left town after John robbed the bank.

2 There is a way to avoid the problems with NPs containing quantifiers. One can claim that only a N is pronounized by another N instead of a NP by another NP. Then 'his' from Dougherty's example comes about from 'son + genitive'. Unfortunately for the transformationalists, stronger arguments against their hypothesis appeared during the seventies. For example, Kayne (1971) describes the behaviour of clitic pronouns in French. The argument is as follows: Clitic movement operates only on pronouns; therefore, if there is a transformational rule of pronounization, it must precede cliticization. However, there are cases where clitic movement should then be able to move a pronoun to the left of its antecedent, but the resulting sentence is ungrammatical. It means that Constraint TT demands that pronounization follows cliticization. These incompatible demands can be avoided only by assuming that pronouns are present underlingly and get interpreted later. Those sentences where there is no possible interpretation are ungrammatical.

3 These rules are based on what can be called 'a common denominator' of the categories previously used to limit the possibilities of 'pronounization', which include a specification of sentence linearity, a specification of stress level, a definition of subject and nonsubject, a definition of main clause and subordinate clause, a specification of the sentence type (complex sentences with indirect speech versus other complex sentences), etc., etc.

4 Rule E is just a special case of Rule B. It specifies one typical case where the degree of Communicative Dynamism carried by a full NP would be too much higher than that of a pronominal theme proper. The restriction as it is formulated in Rule E is, in fact, to severe. As we shall see later, sentences (42) and (43) show that it is possible to have pronominal themes proper if their respective antecedents are placed in some kind of parenthetical expression which carries a very low degree of CD.
There is a usual sort of dialogue which has been observed by everybody: Person A is sitting alone in an otherwise empty room. Person B pokes his head in and asks: 'Is he in here?'. A replies: 'Who?' regardless of the logical truth that the answer should be an instant 'No.' for any possible antecedent of 'he'.

See Note 4.

The interpretative theory has been used as it is not as discredited as the more usual transformational hypothesis. However, it must be emphasized that the conclusions are in no way dependent upon this theory. It would be possible to construct, for example, a model where pronouns are generated with indices determining coreference relations and the resulting sentences would be grammatical only when the coreference markers would not clash with the FSP interpretation of the transformations the sentences in question had undergone and their semantics.

and 9 follow after the bibliography.

Bibliography


Bickerton, D. (1975), "Some Assertions about Presuppositions about Pronominalization", in Papers from the Parasession on Function- alism, 24-35

Bresnan, J. (1971), "A Note on the Notion "Identity of Sense Anaphora"", Linguistic Inquiry 2:4, 589-597

Chomsky, N. (1957), Syntactic Structures

(1965), Aspects of the Theory of Syntax

(1970), "Remarks on Nominalizations", in Jacobs and
Chomsky, N. (1971), "Deep Structure, Surface Structure, and Semantic Interpretation", in Steinberg and Jakobovits (eds), Semantics, 183-216


Daneš, F. (1964), "A Three-Level Approach to Syntax", Travaux linguistiques de Prague 1, 225-240


Firbas, J. (1959), "Thoughts on the Communicative Function of the Verb in English, German and Czech", Brno Studies in English 1, 39-68

(1964a), "From Comparative Word-Order Studies", Brno Studies in English 4, 111-128

(1964b), "On Defining the Theme in Functional Sentence Analysis", Travaux linguistiques de Prague 1, 257-260

(1966), "Non-Themetic Subjects in Contemporary English", Travaux linguistiques de Prague 2, 239-256


Harman, G. (1972), "Noun Phrases Derived from Variable Binding Operators", To appear in Forms of Representation, ed. by Freed, Marras and Maynard

(1976), "Anaphoric Pronouns as Bound Variables: Syntax or Semantics?", Language 52:1, 78-81


(1975b), "Backward Pronominalization across Coordinate Structures", Linguistic Inquiry 6:2, 330-335

Jackendoff, R.S. (1968), "An Interpretative Theory of Pronouns and Reflexives", (Mimeographed by Indiana University Linguistics Club, Bloomington)

(1972), Semantic Interpretation in Generative Grammar


(1972b), "Functional Sentence Perspective", Linguistic Inquiry 3:3, 269-320

(1975), "Three Perspectives in the Functional Approach to Syntax", in Papers from the Parasession on Functionalism, 276-335

(1976), "Three Perspectives in the Functional Approach to Syntax", in Sound, Sign and Meaning, Quinquagenary of the Prague Linguistic Circle, edited by L. Matejka, p. 119-190


Lakoff, G. (1968), "Pronouns and Reference". (Mimeographed by Indiana University Linguistics Club, Bloomington)

(1970), "Repartee, or a Reply to "Negation, Conjunction and Quantifiers”", Foundations of Language 6, 389-422


McCawley, J.D. (1968), "Lexical Insertions in a Transformational Grammar without Deep Structure", in Darden et al., Papers from the Fourth Meeting of the Chicago Linguistic Society, 71-80

(1970), "Where Do Noun Phrases come From?", in Jacobs and Rosenbaum (eds), Readings in Transformational Grammar, 166-183

Postal, P.M. (1971), *Cross-Over Phenomena*
(1972a), "A Global Constraint on Pronominalization", *Linguistic Inquiry* 3:1, 35-59


Svoboda, A. (1968), "The Hierarchy of Communicative Units and Field as Illustrated by English Attributive Constructions", *Brno Studies in English* 2, 49-102
(1974), "On Two Communicative Dynamisms", in Dnnels 1974, 38-42


Note 8 After having finished this paper, I could read the doctoral dissertation of T. Wasow ("Anaphoric Relations in English", unpublished, Massachusetts Institute of Technology, 1972), which contains some counter-examples to Rule A, mostly sentences with generic NPs. ("When they are hungry, tigers can be dangerous.", 'If he has an ugly wife, a man should find a mistress.' etc. - The male chauvinist pig is Postal ("On Coreferential Complement Subject Deletion", *Linguistic Inquiry* 1:4, 439-500), not me.) Such 'generic pronouns' can be perhaps explained as having 'antecedents in the situation' in the sense of our knowledge of the words used in generic
sense. As for Wesow's other examples with indefinite but specific 'postcendants;' ('After Bill kissed her, a certain young lady blushed repeatedly.', 'That he was not elected upset a certain leading politician.' etc.), these are examples of the speaker's power to present a part of a sentence as 'known to the listener', even when it is probable that the knowledge is not shared by the listener. The speaker talks in indeterminate terms, but he has a definite lady/politician in his mind.

Note 9 In the paper mentioned in Note 8, Wesow has also shown that some pronominal rhemes proper standing before their 'antecedents', are possible in special contexts, e.g.: 'Was it after you robbed the bank, that Mary left town? No. After HE robbed the bank, John left town.' It is quite impossible in sentences blocked by Rule B and C (cf. (35) and (37)). Some more examples where Rule D cannot be cancelled by any context:

* It was HE who killed Peter's wife.
* It was HIM Peter's wife killed.
* It was HIS wife who killed Peter. (Compare with 'It was his WIFE who killed Peter.')
* It was HIM Peter's wife killed, etc.

(See also sentences (78) - (86) in the main text.)