SOME THOUGHTS ON TRANSDERIVATIONAL CONSTRAINTS

George Lakoff

Recent work by Grinder, Postal and Perlmutter has shown that the theory of derivational constraints is too weak to account for the facts of such natural languages as English, Spanish, French and Slovenian. The basic point of the Grinder-Postal examples is that in certain structural configurations, ambiguities of a certain sort are not permitted, though they are permitted in other structural configurations. Consider (1) and (2).

(1) John and Mary entered the room, and he took off his coat.

(2) *John and Bill entered the room, and he took off his coat.

In (1), John is the antecedent of he. In (2), where there are two possible antecedents, the sentence is ill-formed on both readings. This is not generally true, as (3) shows.

(3) John told Bill that he had won the sweepstakes.

In (3) the sentence is acceptable on both readings, rather than unacceptable in both readings. Grinder and Postal observed that (2) cannot be blocked within the theory of derivational constraints, since ambiguity is not a property of a single derivation. Since blocking occurs just in the case where ambiguity would arise, the only way a rule of grammar could block either of the two derivations of (2) was if it knew about the other possible derivation. Thus, there must be rules that apply not to individual derivations, but to classes of derivations. In short, transderivational constraints are required, since there are cases where the well-formedness of one derivation depends on certain properties of other, related derivations. Perlmutter's

cases are even more interesting, since there the application of an otherwise obligatory transformation rule is blocked if it would yield an ambiguity. Discussion of such cases is, however, beyond the scope of this note.

I would like to discuss a number of cases which I suspect will require transderivational constraints, though I have no proof at present. I am bringing these cases up merely to stimulate further discussion of the matter. Consider the following:

(4) John scratched his arm and so did Bill. (scratch his arm)

(5) The boys scratched their arms and so did the girls. (scratch their arms)

(6) John scratched his arm and so did Mary. (scratch her arm or scratch his arm)

(7) I scratched my arm and so did Bill. (scratch his arm or scratch my arm)

There are two principal dialects corresponding to these sentences.

Dialect I: All of 4, 5, 6, and 7 are ambiguous.

Dialect II: 4 and 5 are ambiguous. 6 and 7 have only the scratch his arm and scratch my arm readings respectively.

The differences between the dialects appears to involve the definition of constituent identity. In dialect II, the pronouns must have the same phonological form in order for the two verb phrases to be considered identical.

Let us assume that the dialect difference is based on the difference in identity requirements, as I have suggested, and that dialect II requires the pronouns to have identical phonological forms. Then there exists one (and I think only one) general hypothesis under which the facts of dialect II can be handled without transderivational constraints, namely, that there is a rule of verb phrase deletion and that the phonological forms of the pronouns

appear in the derivation before the application of that rule. Then, one can define constituent identity in dialect I so that the phonological forms of pronouns were irrelevant, but define it in dialect II so that the phonological forms of pronouns had to be identical. No transderivation constraint would be required.

However, there are two other possible hypotheses which would require transderivational constraints.

Hypothesis A: There is a rule of verb phrase deletion, but the phonological forms of pronouns are not introduced until after it has applied.

Hypothesis B: There is no rule of verb phrase deletion. Instead, the meanings of missing verb phrases (and other anaphoric expressions) is to be determined by a surface structure interpretation rule.

On Hypothesis A, deletion would take place before the phonological forms of pronouns were introduced, so that the deletion rule could not be sensitive to a difference in phonological forms. The only way that dialect II could be described would be to say that <u>if</u> the verb phrase had not been deleted, then there <u>would have been</u> a difference in the phonological forms of the pronouns in (6) and (7). In other words, there is a corresponding derivation in which the deletion does not take place, and in which there <u>is</u> a difference in the phonological forms of the pronouns. That is, a transderivational constraint would be required.

On Hypothesis B, the right hand verb phrases in (4) - (7) would not be present in underlying structure. Instead, a surface structure interpretation

rule would specify their meanings. Presumably, the output of a surface structure interpretation rule is a semantic representation, and therefore would not contain any phonological forms. In any event, it would not contain any phonological form which was not present in the surface structure. Thus, a surface structure interpretation rule would have no way of referring to any difference in the phonological forms of the pronouns in (6) and (7), and thus could not handle the facts of dialect II. Under Hypothesis B, one would be required to say that <u>if</u> the corresponding VP had been generated in the underlying structure, there <u>would have been</u> a difference in the phonological forms of the pronouns. Again this is tantamount to saying that there is a corresponding derivation where the phonological forms of the pronouns <u>are</u> different. In other words, a transderivational constraint.

In Lakoff (1969b) and Lakoff (in press), I suggested that there were a pair of global derivational constraints on the occurrence of quantifiers and negatives that held throughout all derivations, down to shallow structure. The reasons why I proposed that the constraints held throughout derivations, rather than simply relating P_1 and shallow structure, had to do with cases like (8).

(8) Sue likes many boys, but Sheila doesn't.

Here the relative scope of the negative and quantifier is understood on the right-hand side of the sentence in the same way as it would be if the verb phrase had been present. Since the verb phrase would presumably be deleted by shallow structure, stating the constraint simply on P_1 and shallow structure would not account for cases like (8). On the other hand, having the constraint apply throughout the derivation down to shallow structure would guarantee the correct interpretation of (8). However, Ray Jackendoff (personal communication)

has observed that if adverb preposing is a post-cyclic rules, and if quantifier-lowering is a cyclic rule, then the constraint cannot hold throughout the derivation. Sentences (9) and (10) illustrate the point.

(9) Some men have wives in many cities.

(10) In many cities, some men have wives.

If adverb preposing is postcyclic, then (10) would have to be derived via the intermediate stage of (9). But if the constraints held throughout the derivation, the reading of (10) would be blocked at the point (9) was formed, and so (10) could not be derived. Thus, on the assumption that adverb-preposing is postcyclic (which may or may not be a reasonable assumption), the quantifier constraints would have to be stated on P_1 and shallow structure and not on the entire derivation. But then, what about sentences like (8)? What one would want to say about such cases is that sentences where verb phrases have been deleted are subject to the same restrictions as the corresponding sentences where the verb phrase has not been deleted. To say this, one needs the notion of corresponding derivations and, hence, transderivational constraints.

I would guess that the same would be true of Postal's treatment of Equi-NP-deletion using doomed pronouns. Postal observed that pronouns "deleted" by Equi-NP-deletion in the cycle acted as though they were still there with respect to post-cyclic output conditions. He proposed that instead of being deleted in the cycle, they be marked [+DOOM] at that point, and then deleted after the pronominalization constraints. In Lakoff (1969) and Lakoff (in press), I proposed that such facts could better be handled by global derivational constraints. In such a treatment, the pronoun would be deleted late in the derivation depending on whether or not the SD of Equi-NP-deletion were met at the appropriate earlier point in the derivation.

The arbitrary marker [DOOM] would thus be unnecessary. This solution, though an improvement, still involves such possible difficulties as the following: (i) the pronoun to be deleted must still be kept around throughout the derivation, and (ii) since the pronominalization constraints involve considerations of length (see Lakoff (1968)), they must occur very late, if not at the level of surface structure itself, and so it may turn out not to be possible to have even later deletion transformations. I do not know whether these possible difficulties will turn out to be real difficulties. If they do, then I think a solution in the direction of transderivational constraints would be indicated. In such a proposal, Equi-NP-deletion would delete the coreferential pronoun cyclically, as previously assumed. But deleted pronouns would be subject to the same constraints as if they had not been deleted, that is, the constraints that <u>do</u> apply in corresponding derivations where the pronouns are in fact not deleted.

It would be nice if this were true in general, that is, if derivations containing deleted elements were subject to the same constraints as the corresponding derivations in which the corresponding elements are not deleted. Unfortunately, this is not the case, as was shown in Ross (1969). Ross showed there that the rule of sluicing produced grammatical sentences which would have been ungrammatical had the deleted material not been deleted. The question arises as to whether there is a general characterization of what kind of rules such a proposal will work for and what kind it will not work for. If such a general characterization is possible, then it would be possible to build an account of the above phenomena into the theory of grammar, rather than having to state language-particular transderivational constraints to account for them.

It may be thought that transderivational constraints are a new and radical idea. But actually, the basic idea of a "corresponding derivation" has been inherent in transformational generative grammar since its beginnings. For example, consider (11).

(11) Sam believes that Harry is a fink.

Suppose one were to ask whether the clause that Harry is a fink is a noun phrase in (11). What kind of evidence would bear on the question? Sentences (12) and (13) would usually be taken to be adequate evidence, since they show that the clause can be passivized and clefted, and presumably only noun phrases can undergo such rules.

(12) That Harry is a fink is believed by Sam.

(13) What Sam believes is that Harry is a fink.

But notice that one is using evidence from the derivations of different sentences, (12) and (13), to prove something about the constituent structure of (11). Somehow, the derivations of (12) and (13) are taken to be "corresponding derivations" of (11), in some still unexplicated sense of the term. Clearly, (14) and (15) could not be used in this argument in place of (12) and (13), since they are not "corresponding derivations".

(14) False teeth are made in Bayonne, N. J.

(15) What Sam saw was a star-nosed mole.

Such an intuitive notion of "corresponding derivation" is central to all argumentation in generative grammar. It is not clear to what extent this intuitive notion of "corresponding derivations" overlaps with the concept that is needed to characterize transderivational constraints. But it should not be too surprising that a concept akin to a notion that is inherent in

just about all syntactic arguments should turn out to be necessary for the statement of rules of grammar.

So far we have discussed two possible types of derivational constraints: those involving ambiguity and those involving constraints on deleted elements. We will now turn to a third possible type. Again, I have no strong evidence that what I will propose is the correct way to handle the following phenomena, though I know of no other alternative at present. Let us begin by considering sentences like (16), some of whose peculiarities have been pointed out by Wilson (to appear).

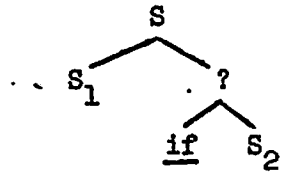
(16) (Only) three girls left, if that.

Apparently, (16) is to be derived by a series of reductions, as indicated in (17).

(17) (Only) three girls left, { if even that many girls left. / if even that many. / if that many. / if that. }

The <u>if</u>-clause in (17) is not an ordinary <u>if</u>-clause. It does not have an <u>if-then</u> meaning and it cannot be preposed. (17) is also somewhat odd semantically, since the <u>if</u>-clause seems to "qualify" the main clause. Unfortunately, it is not clear how to express such qualifications semantically without deriving contradictions between the main and subordinate clauses. Be that as it may, it would appear that at some point in the derivation, there is a structure with a main and subordinate clause, as in (18).

(18)



If we restrict ourselves to cases like (17), it would appear that S_1 and S_2 are constrained so that the parts of them that do not contain the contrasted quantifiers must be identical. Thus, we do not get:

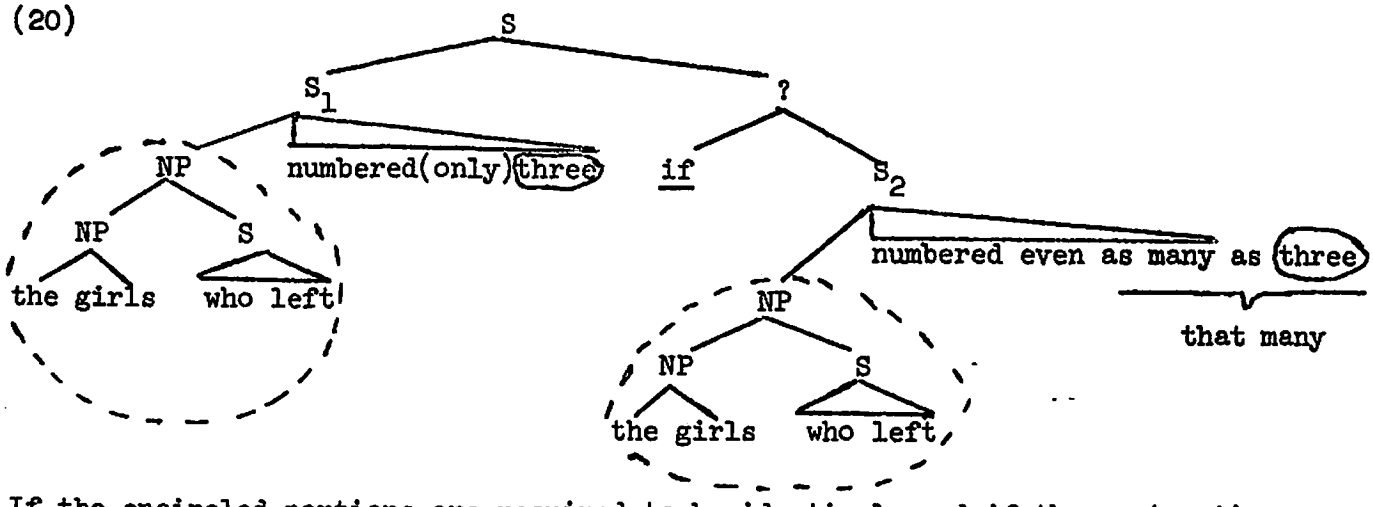
(19) a. *(Only) three girls left, if even that many chinchillas are shedding.

b. *(Only) three girls left, if even Nixon has hemorrhoids.

c. *(Only) three girls left, if even that many bombers were shot down today.

Assuming the (in many respects inadequate) treatment of quantifiers as higher predicates with relative clause subjects given in Lakoff, 1965, and Carden, 1967, we might state a first approximation to this condition as in (20).

(20)



If the encircled portions are required to be identical, and if the contrasting quantifiers are the highest predicates in S_1 and S_2, then we can account for the facts of (19). In addition, such a constraint provides a basis for the explanation of the remarkable fact discovered by Perlmutter and Ross (personal communication) that sentences like (17) obey Ross' constraints on movement transformations. Consider the following:

Complex NP Constraint (Relative Clause Case)

(21) *The man who knew (only) three girls left,
- if the man who knew even that many girls left.
- if even that many.
- if that many.
- if that.

Sentential Subject Constraint

(22) *That (only) three girls will leave is likely,
- if that even that many girls will leave is likely.
- if even that many.
- if that many.
- if that.

Compare (22) with (23), where the subject clause is extraposed and the sentential subject constraint does not hold.

(23) It is likely that (only) three girls will leave,
- if it is likely that even that many girls will leave.
- if even that many.
- if that many.
- if that.

Coordinate Structure Constraint

(24) *John and three girls left,
- if John and even that many girls left.
- if even that many.
- if that many.
- if that.

Complex NP Constraint (Complement-with-head-noun Case)

(25) *Sue believe Max's claim that only three girls had left, if she believed his claim that even that many girls had left.

Compare (25) with (26), which has a complement without a head noun.

(26) Sue believed that only three girls had left, if she believed that even that many girls had left.

Note, incidentally, that although sentences like (27) are grammatical, the if that is interpreted only as originating inside the complement S following claim, and hence supports the Perlmutter-Ross claim.

(27) Sue believed Max's claim that only three girls left, if that.

The facts of (21) - (26) follow from an analysis of the sort given in (20), since quantifier-lowering is subject to Ross' constraints (see Lakoff, in press b) and since all of (21)-(26) would involve quantifier-lowering. If these were all the facts to be accounted for, we would have a rather neat account of them. However, the facts are somewhat more complicated. Jerry Morgan and Georgia Green have pointed out the following cases.

(28) a. Harry is (only) a white liberal, if that.

b. Today is (only) Thursday, if that.

It is clear that for such cases the identity condition given in (20) cannot hold. Yet there can be little doubt that the same construction is involved, and moreover, Ross' constraints also hold for such cases.

(29) *The man who said that Harry was (only) a white liberal left, if that.

(30) *Today is (only) Thursday and I'm already tired, if that.

(31) *That Harry is (only) a white liberal surprised Sam, if that.

Clearly (29) - (31) should be accounted for on the same basis as (21)-(26). The question is how.

If one attempts to expand the sentences of (28) to full form, one finds that there are various alternatives, among which are the following.

(32) a. Harry is (only) a white liberal, if he's even {that far left. / as far left as a white liberal}

b. Today is (only) Thursday, if it's even {that late in the week. / as late in the week as Thursday.}

Of course, there are other possibilities, such as:

(33) a. Harry is (only) a white liberal, if he's even {that masochistic. / as masochistic as a white liberal.}

b. Today is only Thursday, if it's even {that early. / as early as Thursday.}

Further examples can be produced easily. The problem here is why sentences such as (32) and (33) are grammatical, even though they don't meet the identity condition of (20), while the sentences of (19) are ungrammatical. Moreover, if the <u>if that</u> construction is to be derived via a rule which deletes material in the right-hand clause under identity, how can the sentences of (28) be derived at all? If the material in the right-hand clause of (33) is to be subject to deletion under identity, what is there for it to be identical to?

The answer, I think, resides in the observation made in Lakoff (1969a) to the effect that there are rules of grammar which require identity conditions between a constituent in a sentence S_1 and a corresponding constituent in a sentence S_2 that can be deduced from the speaker's presuppositions concerning S_1 and the meaning of S_1. Note that (32a) requires the presupposition that white liberals are to the left of the political spectrum in the views, and that (32b) requires the presupposition that Thursday is late in the week. Thus, it would seem that the deletions in the derivations of (28) would occur under identity with a constituent in sentences which can be deduced from the speaker's presuppositions plus the content of the left-hand clauses of the sentences of (28). Since there are many such possible presuppositions and deductions, (28) will mean different things to different people, depending on what they most naturally assume about the nature of the world.

In Lakoff (1969a), I assumed that deductions from the presuppositions and semantic content of a sentence could be considered extensions of the derivation of the sentence itself. This is a somewhat odd way of looking at things, but there was no alternative at the time. I now think it would be more natural to think of deductions as involving sequences of the semantic representations of derivations. Identity conditions between a derivation and what can be deduced from the semantic representation, including the presuppositions, of that derivation should, I would guess, be stated by transderivational constraints, now that we know that such mechanisms must be available. By such a mechanism, it would be possible, if the theory of transderivational constraints were set up appropriately, to account naturally for the ungrammaticality of (29)-(31). Recall that although a sentence like Harry is a white liberal does not contain a quantifier corresponding to the one in Three girls left, the sentence Harry is a white liberal, if that makes sense only if it is presupposed that white liberal is some point on a scale, presumably political or sociological. Given such a presupposition, what one can deduce from Harry is a white liberal is that Harry is at some point on some sort of scale, in other words, one can deduce a sentence containing a quantifier expression. Although the sentence Harry is a white liberal, if that does not contain a quantifier expression, what is communicated by that sentence, given certain presuppositions, does contain a quantifier expression. What is remarkable is that if white liberal is in a position in the sentence where the corresponding quantifier expression would produce a violation of Ross' constraints due to quantifier-lowering, then white liberal produces a violation of Ross' constraints, even though the sentence itself does not contain a quantifier. In short, sentences like (29)-(31) violate Ross' constraints because there are

there are "corresponding" sentences that violate Ross' constraints. "Corresponding" sentences in this case are defined as sentences whose P_1's are deduceable from the PR and P_1 of the given sentence, and which meet condition (20):

Such an account of the facts of (29) - (31) should not be at all surprising, given that what one communicates (and means to communicate) by saying a sentence is very often not the content of the sentence itself, but what can be deduced given the sentence and those presuppositions shared by speaker and hearer. It should not be too surprising that grammatical constraints on the sentence communicated (via deduction from common presuppositions) should also apply to the sentence uttered.

Another possible application of transderivational constraints might come in the study of idioms. Any theory of idioms must incorporate somehow Weinreich's observation that, by and large, the class of idioms is drawn from the class of well-formed surface structures of the language in question. The outstanding exception is "by and large". In many cases there will be well-formed derivations corresponding to these surface structures, and in many cases there will not. One way of approaching the problem of idioms through transderivational constraints would be to assume that well-formed derivations containing idioms are "secondary" in the sense that they are formed via transderivational constraints on the basis of other well-formed or partly well-formed derivations. So the surface structure of an idiom derivation would be identical to the well-formed surface structure of some non-idiom derivation, which itself would be fully or partly well-formed. Similarly for the semantic representations of idiom-derivations. Thus the class of semantic representations and surface structures of idiom-derivations

would be drawn from the class of well-formed semantic representations and surface structures in primary derivations. The tough question is how to match up the right semantic representation and right surface structure with the right intermediate structures in forming such derivations, so that all the odd transformational properties of idioms can be accounted for. I have nothing of substance to say on this matter beyond the very little that has already been said. I merely throw out the suggestion as something to keep in mind while formulating a theory of transderivational constraints.

One final thing to worry about: It is conceivable that certain theories of transderivational constraints could do some of the work of transformational rules. For example, primary derivations might contain no passives. Passive sentence might be formed transderivationally on the basis of the corresponding well-formed or partly well-formed actives. Of course this would involve such madness as ordering of transderivational constraints, cyclical transderivational constraints, exceptions to transderivational constraints, and perhaps the elimination of transformations altogether. This sounds wrong to me, so far as pure gut reaction is concerned, and certainly no theory of transderivational constraints is sufficiently close at hand so that one could reasonably see what would be involved. However, it is nonetheless important to try to imagine what kinds of empirical considerations could choose between such alternatives when one goes about constructing such a theory.

University of Michigan
March, 1970

REFERENCES

Carden, Guy. "English Quantifiers." In Report NSF-20, Mathematical Linguistics and Automatic Translation, Harvard University Computation Laboratory, 1968.

Lakoff, George. "On the Nature of Syntactic Irregularity", NSF-16, Mathematical Linguistics and Automatic Translation, Harvard University Computation Laboratory, 1965.

Lakoff, George. "Pronouns and Reference", Harvard mimeo, 1968. Available from Indiana University Linguistics Club.

Lakoff, George. "The Role of Deduction in a Grammar", presented at the First Annual Ohio State Semantics Conference, 1969a.

Lakoff, George. "Baker's Conjecture and Natural Logic." In Phonetics Lab Notes, Volume 4, October, 1969.

Lakoff, George. "On Generative Semantics." In Jakobovits and Steinberg, Semantics, Cambridge University Press, in press.

Ross, J. R. "Guess Who." In Binnick et al., papers from the Fifth Regional Meeting of the Chicago Linguistics Society, 1969.