On Perceptual and Grammatical Constraints

Alexander Grosu
On Perceptual and Grammatical Constraints

Alexander Grosu

0.0. In "The Cognitive Basis for Linguistic Structure," T. G. Bever suggests that "some so-called 'output conditions', 'derivational constraints' and/or 'interpretive rules' are formalizations of behavioral constraints." He further remarks that by removing such powerful devices from the grammatical apparatus proper and by viewing them as perceptual in nature, "grammatical simplicity and purity would be purchased at the expense of descriptive power granted to other aspects of sentence description," and chiefly, the study of grammar would not reduce "to an unrevealing taxonomy". It is the purpose of this paper to follow up Bever's remarks in a few concrete cases involving derivational constraints, as well as devices of even greater power which had not yet been proposed at the time when Bever's article was written, namely, trans-derivational constraints.

0.1. A few preliminary remarks are in order at this point. First, it would be unreasonable to expect that there be a fool-proof test which enabled us to decide whether any given constraint is grammatical or perceptual. I believe, however, that there are signs which may help us make a decision. For one thing, some, although not all, native speakers are capable of distinguishing between unacceptability and ungrammaticality on intuitive grounds. Furthermore, the unacceptability resulting from the violation of a perceptual constraint seems to increase in severity as the value of some parameter, like the length of a constituent, is varied, while instances of violation of grammatical constraints usually have a yes/no character. Notice that this phenomenon is different from those considered by Chomsky and other writers in their discussions on "degrees of grammaticality," since they were concerned with comparing the inadmissibility of constructions in which different kinds of requirements had been violated, while I am concerned with varying degrees of admissibility in cases where the same kind of constraint has been violated.

0.2. Second, we should notice that Bever makes his claim in relation to some constraints. Indeed, there is no reason for assuming a priori that all constraints should be reducible to cognitive limitations, since there is a large body of purely syntactic cases, outside of the phenomena considered here, and it is perfectly possible that some derivational and transderivational constraints may not involve perception. As an example of a phenomenon that I regard as purely syntactic, consider the fact that some languages have pre-nominal, while others have post-nominal, modifiers; I do not see here what the distinction could have to do with perception.
Another important word in Bever's formulation is formalizations. Indeed, although behavioral constraints reflect limitations of the perceptual apparatus and must therefore be assumed to be universal, it does not follow that they must be manifested as merely perceptual in all languages. The reason is that some languages may have "grammatized" specific constraints, with the result that constructions of increasing complexity no longer lead to increasing unacceptability, but rather exhibit a partition into grammatical and ungrammatical constructions, with the cut-off point often arbitrarily located way below the limit of processability. As a case in point, consider the decreasing intelligibility of increasingly complex pre-nominal modifiers, which I think may be safely regarded as a universal phenomenon. Now, German tolerates highly complex pre-nominal modifiers, particularly in writing, while English sets a fairly low cut-off point. The German sentence in (1a) is not at all an extreme example: however its perceptual difficulty can be appreciated by examining (1b), a literal English rendering, which is extremely hard to interpret, in addition to being ungrammatical. It appears that German, unlike English, has not grammatized the constraint, since complex adjectival constructions can be used up to (and even beyond!) the limit of intelligibility.

(1a) Die während der Nacht mit dem aus Frankreich kommenden Zug eingetroffenen Touristen fanden keine Unterkunft.

(1b) *The during the night with the from France coming train arrived tourists did not find any lodgings.

That the cut-off point is fairly low in English can be seen in (2a), which is ungrammatical but easily interpretable. That the constraint is in fact grammatized can be seen by comparing (2a) and (2b), only the former of which is out, although they contain modifiers of equal complexity.

(2a) *The captain wants a ready to die volunteer.

(2b) Mary buys ready-to-wear clothes.

0.3. Third, a perceptual constraint may not be overtly manifested in specific languages if it is overridden by structural factors. In other words, a constraint will be observable only in the absence of features of language structure that may guide the processing device. Consider, for example, Bever's proposed strategy that NP-V-NP sequences are interpreted as ACTOR-ACTION-OBJECT (in fact, the interpretation must be formulated more generally, possibly as LOGICAL SUBJECT-PREDICATE-LOGICAL OBJECT, since not all verbs occurring in this framework allow Agents). We would certainly not expect this strategy to apply to languages with overt case markers, where the position of NPs is not crucial in determining their roles.

1.0. Let us begin our substantive discussion by considering Postal's cross-over phenomena, which were initially formulated as a constraint on movement transformations, but which I understand Postal would
want reformulated as a derivational constraint, presumably involving underlying structure and some intermediate derivational stage or stages. I shall confine my attention to the so-called "A-movements" (which do not make crucial use of variables), as represented by the rules of PASSIVE, FLIP, ABOUT-MOVEMENT, SUBJECT-RAISING and TOUGH-MOVEMENT, as I do not understand very well what goes on in the B-movements. The inadmissible structures resulting from the application of the five rules mentioned above in cases where some NP crosses over another with which it is coreferential are exemplified in (3) - (7).

(3a) John stabbed himself.
(3b) *John was stabbed by himself.
(4a) I am amazed at myself.
(4b) *I amaze myself.
(5a) 'I talked to John about himself.
(5b) *I talked about John to himself.
(6a) It seems to me that I am stupid.
(6b) *I seem to myself to be stupid.
(7a) It is difficult for me to shave myself.
(7b) *I am difficult for myself to shave.

If the cross-over constraint exhibited in (3) - (7) is perceptual, it has certainly been grammaticalized in Postal's speech, as well as in a large number of dialects of English and in some other languages. Postal notes, however, that the phenomenon is by no means universal. On the other hand, it might not be too easy to detect the constraint in those dialects and/or languages in which it does not have grammatical status, as the b-sentences in (3) - (7) are intelligible to everybody; this suggests that the purely perceptual problems these sentences might create cannot be too serious. I believe, however, that this line of investigation is worth pursuing, for I have found native speakers of English who accept the sentences in question, but nevertheless find them mildly disturbing or slightly odd.

1.1. Let us now ask what perceptual problems could be involved in (3b) - (7b). We may begin by noticing that Bever's strategy concerning the interpretation of NP-V-NP sequences is applicable (in the more general form that I suggested above) to (3a) and (4a) since the singly understood NPs are perceived as Agent and Experiencer respectively. Bever offered some experimental evidence that although a violation of this strategy is possible, in the presence of some overt marker (like the passive form of the verb) the resulting structures are perceptually more complex, or, to put it somewhat differently, "marked." In (6a) and (7a), the superficial segmentation is NP-V-NP-NP, the last NPs being a full clause and a subject-less infinitival phrase respectively. However, given the lexical data that seem and difficult are two-place predicates and that their logical Subject is an Experiencer, the first NP, it, is interpreted as a dummy and disregarded in role-assignment. Moreover, given the additional lexical information that these predicates cannot be preceded by their Experiencer in surface structure, (6a) and (7a) are perceived as "unmarked." If we can also argue that there are
strategies based on the "natural" order of NPs (given predicates with more than two arguments) so that the natural position of the about-phrase is after the indirect object, it will be the case that all the sentences in (3a) - (7a) are perceived as unmarked, while all the (b) sentences are perceived as marked.

1.2. Notice that in the unmarked structures considered above, the left-to-right ordering of arguments overtly expresses the intuitive notion of "hierarchy of roles," since every argument both commands and precedes the arguments to its right. In (3b), (4b), and (5b), two arguments have switched places, with the result that a hierarchically lower argument bears both primacy relations to a higher one; the effect is that the referent of the lower argument is thrown into focus. In (6b) and (7b), the focusing effect is even stronger, since an NP from a lower sentence is both upgraded to the rank of clause-mate of the Experiencer and given both primacy relations over the latter. We are now ready to formulate a tentative explanation of the cross-over phenomena considered above:

(A) A hearer assumes that a marked form was used in order to focus on the referent of some NP, and concludes that this referent would not have been mentioned in that position, had the unmarked form been used.

Put differently, a marked form carries the presupposition that an NP in positional focus is not coreferential with any NP with respect to which it is focused.

1.3. If we grant (A), we are in a position to explain several problems which required fairly elaborate treatment in Postal's monograph.

1.3.1. Most informants I have consulted agree that the violation in (4b) is much less severe than in the other (b) sentences. We can explain this by pointing out that a principle like (A) depends on a hearer's ability to recognize marked forms. But flipped sentences look very much like ordinary unmarked ones, and it is only through a semantic interpretation of the arguments that a hearer can realize that the subject of the sentence is not an Experiencer. Moreover, considering that the relation between adjectives and verbs is much less regular than that obtaining between active and passive verbal forms, some informants may be unaware of a relation between (4a) and (4b), and thus fail to recognize the latter as marked.

1.3.2. Another problem arises in connection with the paradigm in (8), where both sentences exhibit violations of the cross-over constraint, but where (8a) only is unacceptable.

(8a) I talked about Bill to himself.
(8b) I talked about myself to myself.
It seems reasonable to assume that the referent of some NP_1 can be thrown into focus by appearing earlier than in the unmarked order only if there is no still earlier mention NP_1 of that referent in the structure considered, such that NP_1 commands and precedes NP_1. If there is such an NP_1, it grabs the limelight to itself, as it were, and although the structure constraint is technically marked, no focusing effect arises. Therefore, as the NP of the about-phrase in (8b) is both commanded and preceded by a coreferential Experiencer, (8b) does not in fact count as marked for the purposes of (A).

1.3.3. Consider now the fact that the (b) sentences (3) - (7) are no longer perceived as deviant if the singly underscored NPs bear contrastive stress, and especially if particles like even are added. For example, (9b) is much better than (9a), and (9c) is better than both.

(9a) "John was stabbed by himself.
(9b) John was stabbed by himself.
(9c) John was stabbed even by himself.

Now, one of the functions of contrastive stress is to point out some error in connection with a previously mentioned constituent, as in (10).

(10) Speaker A: I believe John gave Bill a present.
Speaker B: No, John gave Mary a present.

If the presupposition of a sentence is information assumed in advance to be true, and if the use of the passive in (9b) carries the presupposition that the second NP is not coreferential with the first, the contrastive stress serves the purpose of explicitly contradicting this assumption. (9c) is more immediately acceptable than (9b), for even suggests that John was an unlikely candidate for the role of Agent, and thus makes even clearer the presupposition of non-coreferentiality, which is ultimately contradicted by contrastive stress.

2.0. We shall now ask whether derivational constraints that do not involve surface or underlying structure can have anything to do with perception. The answer seems to be no, given Bever's assumption that intermediate stages in a derivation have no psychological reality. I do not, however, believe that Bever's assumption can be maintained in this strong form, although it seems that a weaker version can be defended. The weak and strong formulations are given below as (B) and (C) respectively.

(B) Some intermediate stages in derivations have no perceptual reality.
(C) No intermediate stages in derivations have any perceptual reality.

2.1.1. In support of (B), Bever claims that derivational and perceptual complexity do not always correlate, and offers some
experimental evidence involving agentless passives, which are easier to perceive than full passives, although an additional transformation has allegedly applied. Unfortunately, agentless passives are ambiguous between a reading on which the Agent was deleted and one on which it never existed, as can be seen in the following paradigm, discovered by Fillmore (1968):

(11a) The boy was killed with an axe.
(11b) The boy was killed by an axe.
(11c) The boy was killed.

In (11a), there is an "understood" Agent; in (11b), there is none, while (11c) is ambiguous between the readings of (11a) and (11b). Therefore, it is perfectly possible that Baer's subjects who were presented with a sentence like (11c) interpreted it as synonymous with (11b), which, given no background information, seems to be the more likely reading.

2.1.2. A more convincing argument in support of (B) could be made on the basis of Ross' discovery that violation of his constraints results in more acceptable constructions if the island crossed by some constituent is in turn "sluiced." The point is exemplified in (12) with the Coordinate Structure Constraint.

(12a) *Jill and someone were having tea, but Tom doesn't know who Jill and were having tea.
(12b) ?Jill and someone were having tea, but Tom doesn't know who.

The greater acceptability of (12b) suggests that hearers do not interpret it by reconstructing an intermediate stage like (12a), for then (12b) would be at least as difficult, possibly more difficult to process than (12a).

2.1.3. Another observation made by a number of writers which supports (B) is that multiply center-embedded constructions increase in acceptability if various transformations, e.g. Passive or Extraposition, apply to them. The effect of these two transformations can be seen in (13) and (14) respectively, where the (b) sentences are much easier to perceive than the (a) ones.

(13a) The girl the teacher the school fired flunked hit the boy.
(13b) The boy was hit by the girl who was flunked by the teacher who was fired by the school.
(14a) The rumor that the report which the advisory committee submitted was suppressed is true is preposterous.
(14b) The rumor is preposterous that it is true that the report which the advisory committee submitted was suppressed.

The relevance of (13) -- (14) to (B) could be challenged by a claim that passive and extraposed structures are in fact derivationally
less complex than their active and non-extended counterparts respectively. I do not know of anyone who made this claim in connection with Passive; however, such a claim has been made by J. Emonds in relation to Extraposition. Notice, however, that even if Emonds' claim is accepted, there will still be instances of inverse perceptual/derivational correlation regarding complexity, because of verbs like prove, which may take a clausal subject and a clausal object. The paradigm in (15) shows that Extraposition (as traditionally formulated) is optional with prove, but results in highly unacceptable structures when the object is a that-clause.

(15a) That John told everyone he is staying proves my point.
(15b) It proves my point that John told everyone he is staying.
(15c) That Tom told everyone he is staying proves that he thinks it would be a good idea to stay.
(15d) \(\text{It proves my point that Tom thinks it would be a good idea to stay.} \)

Thus, Extraposition supports (B) regardless of the direction of movement in grammatical description, for its application sometimes increases and sometimes decreases acceptability.

2.2.1. Whatever the force of the examples offered in sections 2.1.2. and 2.1.3., it should be clear that they support (B), but not necessarily (C). In fact, there are good grounds for believing that (C) is in general incorrect, and that it can only be upheld with respect to very simple sentences, but not sentences like (16).

(16) John likes everybody, even George, and so does Mary, but Bill doesn't believe it.

In (16), the antecedent of so does is \(\text{1 likes everybody, 1 even likes George, while the antecedent of it is John likes everybody, he even likes George, and Mary likes everybody, she even likes George. It is not at all obvious how the perceptual strategies could recover the meaning of (16) in one swoop. It is highly probable that they have to operate in three stages, first reconstructing the full form of the constituent even George, then the antecedent of so does, and finally the antecedent of it.}

2.2.2. We are now in a position to answer the question raised at the beginning of section 2.0., with respect to the possibility that a derivational constraint which mentions neither surface nor underlying structure be based in perception. Given the untenability of (C), the answer is clearly Yes. All that is needed is that one of the intermediate stages constructed in the process of decoding be perceptually complex. I shall use one of Lakoff's derivational constraints to argue that a sentence whose deep and surface structures are well-formed can be perceived as inadmissible if an
ina.ible ihtenri.Mio.te stn!se is reconsi:a·u:cted _by the perceptual strategies. Lakoff mentions the following four successive stages in a derivation:

*(17a)* I know the kings of England and Sam knows the kings of Spain.

*(17b)* I know the kings of England and Sam knows the ones of Spain.

*(17c)* I know the English kings and Sam knows the Spanish ones. (Where the Spanish ones means the kings of Spain, not the kings from Spain)

*(17d)* I know the kings of England and Sam knows those of Spain.

*(17b)* is ungrammatical because it contains the ill-formed surface sequence *one of NP*. The interesting features of this derivation are that *(17e)* is ill-formed only if it is derived from *(17b)*, and that *(17d)* is well-formed, although it is derived from the ill-formed *(17b)* and *(17c)*.

From a perceptual point of view, the well-formedness of *(17d)* follows in a straightforward way from the well-formedness of its surface and deep structures. However, the ill-formedness of *(17c)* cannot be explained by a theory of language that assumes *(C)*, for its surface structure is not ill-formed (since it has a perfectly acceptable reading), and neither is its underlying representation. On the other hand, the situation can be handled by a theory that does not incorporate *(C)*, as it is possible to argue that an intermediate stage in the decoding of *(17c)* is the ungrammatical *(17b)*, while no such stage is recaptured when *(17d)* is processed.

3.0. In sections 3. and 4., I shall discuss two proposed trans-derivational constraints, and will argue that the formulation in purely formal terms is inadequate.

3.1. *(18a)* A woman who was pregnant hit a girl.

*(18b)* A woman hit a girl who was pregnant.

*(18c)* A woman hit the curb who was pregnant.

Perlmutter noticed that *(18b)* is not ambiguous, although one would expect it to have a reading synonymous with *(18a)*, due to the existence of the optional rule of EXTRAPosition-FROM-NP. *(18b)* has only a continuous reading, although a discontinuous reading is in principle possible, as shown by *(18c)*. To prevent the discontinuous reading in *(18a)*, Perlmutter formulated a trans-derivational constraint roughly as follows:

*(D)* A derivation D is ill-formed, if EXTRAPosition-FROM-NP applies in D, and if the last line of D thereby becomes string-wise identical with the last line of D', where D and D' differ in semantic representation.

It should be pointed out that the discontinuous reading of *(18b)* is not universally ill-formed; thus, Arnold Zwicky reports that
this sentence is ambiguous in his dialect, although the continuous reading is far more likely. This state of affairs suggests that (D) should be reformulated as a perceptual strategy something like (E):

(E). In the absence of any clues to the contrary, an interpretation that involves discontinuities is rejected, if one that does not is possible.2

If the purpose of a transderivational constraint like (D) is to prevent ambiguities of a certain kind, then it is not clear why it should be restricted to EXTRAPosition-FROM-NP, and why the string-wise identity of the surface structures should be a significant condition for its applicability. That is, it is not clear why the configurational distinctions between (19a) and (19b), which exhibit the surface structures of the continuous and discontinuous readings of (18b) respectively, should not be sufficient to keep them apart.

(19a)

(19b)

3.1.1. With respect to the first objection, it should be pointed out that (D) is not in fact restricted to EXTRAPosition-FROM-NP, since it is applicable to structures involving PARTICLE-MOVEMENT. To see this, notice that the discontinuous reading is possible, with some awkwardness, in (20a), although it is not possible in (20c) (both sentences should be read without a pause before off).

(20a) John pushed the girl who was married to his brother off.
(20b) John pushed off the girl who had fallen.
(20c) John pushed the girl who had fallen off.

This observation does not, in itself, invalidate (D), as the latter could easily be generalized, perhaps by referring to movement transformations in general rather than to EXTRAPosition-FROM-NP: It does, however, create some real problems for (D), if it is taken in conjunction with my second objection.
3.1.2. Let us begin by noticing that the discontinuous reading is imperative in (20c). If we pause before the extraposed constituent off, but it is still impossible in (18b), even if we pause before who was pregnant. Why should pause make a difference between structures involving PARTICLE-MOVEMENT and EXTRAPOSITION-FROM-NP? The answer is that pause is impossible between a verb and its immediately following particle, while it is perfectly possible between a head noun and its relative clause modifier; in fact, it signals that the latter is non-restrictive. Given the possibility of choosing between an extraposed relative clause and a non-restrictive one, the perceptual strategy (E) forces the latter choice. We can see that differences in the surface parsing of sentences do prevent ambiguity, so long as they can be unambiguously signaled, and that the sufficiency of the string-wise identity condition in (10b) was merely a lucky accident.

It is not clear how (D) could be reformulated in a non-ad hoc way to account for the differences between (18b) and (20c), unless one would want to claim that sentences with and without pause are string-wise distinct. Such a way out would not, however, save (D), since the latter asserts that string-wise identity arises because of the application of EXTRAPOSITION-FROM-NP, and pause is not at all an automatic consequence of this transformation.

4.0. The next case that I wish to discuss concerns sentences like (21a), whose unacceptability was claimed by Lakoff to require a transderivational constraint:

(21a) *John and Bill entered the room, and he took off his coat.

(21b) John and Mary entered the room, and he took off his coat.

(21c) John told Bill that he had won the sweepstakes.

The reason invoked by Lakoff was the ambiguity is a property of more than one derivation. I shall propose the perceptual principle (P) for handling (21a), and will show in section 4.4. below that there are cases which a transderivational constraint cannot handle.

(F) Sentences containing pronouns are incomprehensible, if the antecedent of a pronoun cannot be discovered by any means whatsoever, or if there are two or more equivalent candidates for the position of antecedent.

With respect to the means that may enable a hearer to discover the antecedent of a pronoun, I can see at least semantic, deep-configurational, surface configuration and extralinguistic factors.

4.1. The way in which semantic features effect disambiguation can be seen in (21b). Indeed, John, but not Mary, carries the feature [+MASCULINE], and is therefore the only possible antecedent of he. In (21a), however, John and Bill are equivalent candidates with respect to the semantic features in terms of which they must agree with he.
4.2. The part played by deep-structure configurations is demonstrated in (21c). Indeed, John and Bill appear in different underlying configurations, and are therefore non-equivalent candidates. Furthermore, as there is no principle which could force a choice between the two, both qualify as antecedents. In (21a), on the other hand, they are configurationally non-distinct in underlying structure, for John and Bill = Bill and John semantically.

4.3. With respect to surface structure, John and Bill are differently ordered. However, a rather curious feature of English is that, although surface order is in general important, it does not count for purposes of antecedent-recovery. I have heard that there are languages in which the question "Would you like tea or coffee?" can be answered "Yes," with the unambiguous meaning of "I would like coffee," as coffee is the last alternative offered. In such languages, the antecedent of he in (21a) might well be Bill only. It is clear that English is not such a language, for, if it were, there would be no need for items like the former, the latter, respectively, and (21c) would be unambiguous.

4.4. We shall now consider disambiguation through extralinguistic knowledge, which I claim cannot be handled formally. Indeed, I do not see how a transderivational constraint can be formulated to exclude (21a), but not (22).

\[(22) \{\text{John and Napoleon} \} \text{ entered the room, and he said he was going to make war on Russia.}\]

The only way out that I can see is to require that whatever is asserted of a pronoun must be represented in the semantic characterization of its antecedent. This would, however, require the ascription to Napoleon of a feature like [+MADE WAR ON RUSSIA], which anyone would recognize as intolerably ad hoc. The correct generalization seems to be that any piece of information that the speaker and the hearer share can be used in antecedent-recovery.

4.4.2. I shall conclude this talk with a little story that supports the point just made even more strongly, and that cannot, as far as I know, be handled by any existing theory of pronominalization.

[The dashing one-eyed Israeli Minister of Defense, general Moshe Dayan, is equally well-known in his country for his military exploits and for his appreciation of the fair sex.] Upon being asked by a foreign correspondent what she would do if a scandal similar to the Profumo case were to break out in Israel, the Prime Minister replied: "I would take out his other eye."

Although no one can have any doubts as to the referent of his, it cannot seriously be maintained that a constituent like Moshe Dayan is somehow represented in the underlying structure of the
correspondent's question, or that such a constituent can be recovered by purely formal interpretive rules.

Footnotes

1The paradigm in (17) does not, unfortunately, constitute a perfect example of a behavioral constraint on intermediate reconstructed stages, for the unacceptability of *one of NP is probably not due to perceptual factors. The paradigm does, however, suggest that such intermediate stages are important in decoding sentences, and I see no reason at this point for ruling out the possibility of sentences whose unacceptability is due solely to a perceptually complex intermediate representation.

2I have introduced the proviso in the absence of any clues to the contrary in (E), as it is important for my subsequent line of argumentation, but since it is applicable to all perceptual strategies, it should be formulated as a metacondition on perceptual strategies.

3Rosenbaum (1967) thought that distance measured along tree-branches is significant in English for recovering the antecedent of a deleted complement subject, but it was later shown by Postal (1968) that such a "minimal distance principle" cannot be correct since antecedent-uniqueness is only found in a subset of cases, where it appears to be determined by semantic factors. The problem is discussed extensively in my M.A. thesis.

Bibliography
