

Gapping and Constituent Order in Apurinã

Preliminary Analysis

By
Wilbur N. Pickering

Associação Internacional de Linguística, SIL – Brasil
Anápolis – GO
2009*

* This paper is essentially the same as the 1973 original edition (No. 005 of the Arquivo Linguístico, Summer Institute of Linguistics, Brasília, DF); only the page size and numbering have been changed, and a few minor typographical errors corrected.

Gapping and Constituent Order in Apurinã¹

1. Gapping

Such is the name given by John Ross to a rule that "operates to delete indefinitely many occurrences of a repeated main verb in a conjoined structure."² Using Joseph Greenberg's classification of language based on 'dominant' order of constituents³ (SVO, SOV, ETC.), he proposes the following hypothesis: "The order in which GAPPING operates depends on the order of elements at the time that the rule applies; if the identical elements are on left branches, GAPPING operates forward; if they are on right branches, it operates backward."⁴ This hypothesis is important because it is crucial to Ross' purpose to use GAPPING to discover the deep structure order of subject, verb, and object. Ross goes on to affirm, "I know of no language which exhibits no gapping behavior of any kind, but even if such languages should prove to exist, hypothesis (8) (given above) would not be refuted."⁵ To be sure, the hypothesis would not be refuted, but what of the claim of universality for GAPPING and what of its dependability as a device for determining deep structure order of constituents?

Apurinã appears not to exhibit any gapping behavior.

- (1) (a) ximaky nota nika kimi Pedro nika kema Jaimy nika.⁶
fish I ate corn Peter ate tapir Jim ate
I ate fish, Peter ate corn, Jim ate tapir.
- (b) *ximaky nota kimi Pedro kema Jaimy nika.
- (c) *ximaky nota nika kimi Pedro kema Jaimy.
- (2) (a) Pedromony xamynaky awa Jaimymony saasara awa
Peter-at gun is Jim-at bush-knife is
Peter has a gun; Jim has a bush-knife.
- (b) *Pedromony xamynaky Jaimymony saasara awa.
- (c) *Pedromony xamynaky awa Jaimymony saasara.
- (3) (a) netamakiniry nymōpotorykiniry nijōkatsopatiniry sākiretxi
My-try-ing-it my-begin-ing-it my-write-ing-it discourse
nota nyrakake, oētemakiniry omōpotorykiniry
I want her-try-ing-it her-begin-ing-it
ojōkatsopatiniry xikaretxi Maria nyrekaka.
her-write-ing-it lyrics Mary wants
I want to try to begin to write discourse, Mary wants to try to begin to write lyrics.

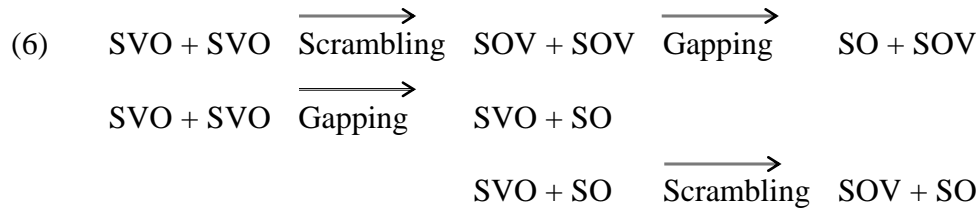
O V s-#, V s-#-o O (s-V-o and s-V-o S are also possible). Capital O and S indicate free forms; s- and -o indicate bound personal pronouns, # indicates a pro or dummy verb. If context is excluded from the model, then we have ten possible surface structures to represent a single underlying structure. What is the underlying constituent order? Because the presence or absence of the bound pronouns is absolutely predictable in terms of the free forms, O and S, their evidence must be regarded as secondary. The only surface order that has no bound pronouns is OSV. Of special interest is the surface order for di-transitive sentences, O S V-o, which is obligatory--e.g.

- (5) anana nota syka-i / anana nota syka-i pite
 pineapple I give-you pineapple I give-you you
 I give you pineapple.

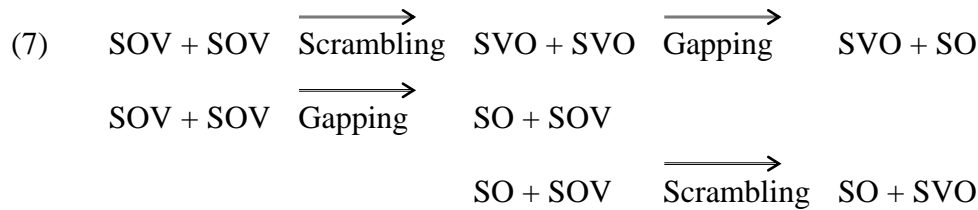
It contrasts with all ten orders above in that the two objects are not coreferential--the bound object pronoun (and the optionally following coreferential free form) represents the indirect object. Thus, the di-transitive structure furnishes strong evidence that OSV is the basic order. Apurinã is postpositional and an inflected auxiliary (my 'dummy' verb) always follows the main verb which in terms of Greenberg's universals 4 and 16 may be construed as support for an OSV order (SOV is clearly out).¹³ Were gapping to occur a resulting SO would be intolerable, which weighs against SVO (and SOV and VSO) as the underlying order. If both free forms either precede or follow low V their order must be OS. Any order having S before O is opposed by the evidence just given. Further, SOV has no surface support at all, while only the surface order V s-#-o O could possibly be adduced in favor of VSO. Surface orders s-V-o O S and s-V-o S might be said to support VOS, while O s-V S and O V s-# might be said to support OVS, but the fact that the Subject (in some form) almost invariably precedes V argues against these possibilities. Besides, there is no motivation for positing them. Both motivation and evidence point to OSV.¹⁴

3. Gapping and Constituent Order

This section is concerned with the validity of using GAPPING to determine underlying order. Ross states frankly at the outset that his paper "is devoted to discovering the deep structure order of subject, verb, and object"¹⁵ and that his aim is to provide a way of deciding when a language is SVO or SOV.¹⁶ It is "with this goal in mind" that he introduces GAPPING even though the formulation of the rule presents problems which he sidesteps.¹⁷ He notes that GAPPING operates forward in English (an SVO language) but backward in Japanese (an SOV language) which facts give rise to the directionality constraint already quoted on page one. He adduces Russian as further support for this metarule since sentences of both the forms SVO + SO and SO + SOV occur in that language. Assuming that Russian has only one order in deep structure, Ross now introduces SCRAMBLING, a rule which optionally permutes major elements of a clause (subject to conditions which Ross does not state or discuss¹⁹ --but without conditions such a rule has power bordering on magic), which if ordered before GAPPING will account for the Russian data and save the metarule. But Russian also has sentences of the form SOV + SO, so in order to save the directionality constraint he assumes that GAPPING is an 'anywhere rule'. However, that is not enough--he must also posit SVO as the underlying order. E.g.



(With this scheme SO + SVO cannot be derived and Ross claims that no language in the world has sentences of this form.²¹) But:



In this case, SOV + SO cannot be derived, but the "impossible" SO + SVO can. Consequently Ross argues that Russian must be underlying SVO and infers that GAPPING, as an 'anywhere rule' and with its directionality metarule, is a necessary part of the chain of inference.²² It seems to me that Ross has begged the question—we could just as well say that we don't know what the underlying order of Russian may be (if it has only one) and that GAPPING is no help. In any case, Dingwall argues that sentences of the form SO + SVO are marginally acceptable to Russian informants and are freely acceptable in Polish (along with the other 3 forms so far discussed).²³ The existence of sentences of the form SO + SVO destroys Ross' chain of argumentation--i.e., it is impossible to derive both SO + SVO and SOV + SO with Ross' scheme using a single underlying order, no matter which one is used (one would need both SVO and SOV as basic orders to account for Polish). Even without the existence of SO + SVO type sentences, Ross' scheme is demonstrably untenable. The directionality metarule is necessary to account for both English and Japanese; to account for Russian as well, GAPPING must be an 'anywhere rule'; but to account for Quechua, GAPPING cannot be an 'anywhere rule'.²⁴ So, Ross' scheme cannot account for both Russian and Quechua. According to Pulte, Quechua exhibits all the following: SOV + SO or SOV + OS, SVO + SO or SVO + OS, VOS + SO or VOS + OS, VSO + SO or VSO + OS, OSV + SO or OSV + OS, OVS + SO or OVS + OS.²⁵ A single SCRAMBLING rule, which could derive any of the foregoing forms from a single base, can only be described as magical. Ross' scheme produced the following typology:

- (8) Type A. SVO + SO
 Type B. SOV + SO
 Type C. SO + SOV
 Type D. SO + SVO

wherein type D was impossible. Based on this typology he gave the following classification of languages:²⁶

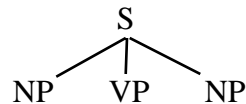
(9)	<u>None</u> (none)	<u>Only A</u> English French	<u>Only B</u> (none)	<u>Only C</u> Japanese Siouan
	<u>Only AB</u> (none)	<u>Only AC</u> (none)	<u>BC</u> Hindi Turkish	<u>ABC</u> Russian Latin

Ross' scheme appeared to account for the facts on such a chart very nicely, but the chart must be revised:

(10)	<u>None</u> Apurinã Hausa	<u>Only A</u> English French	<u>Only B</u> (none)	<u>Only C</u> Siouan	<u>Only D</u> (none)
	<u>Only AB</u> Quechua	<u>Only AC</u> (none)	<u>BC</u> Japanese Hindi Turkish	<u>ABC</u> Russian Latin	<u>ABCD</u> Polish (Russian)

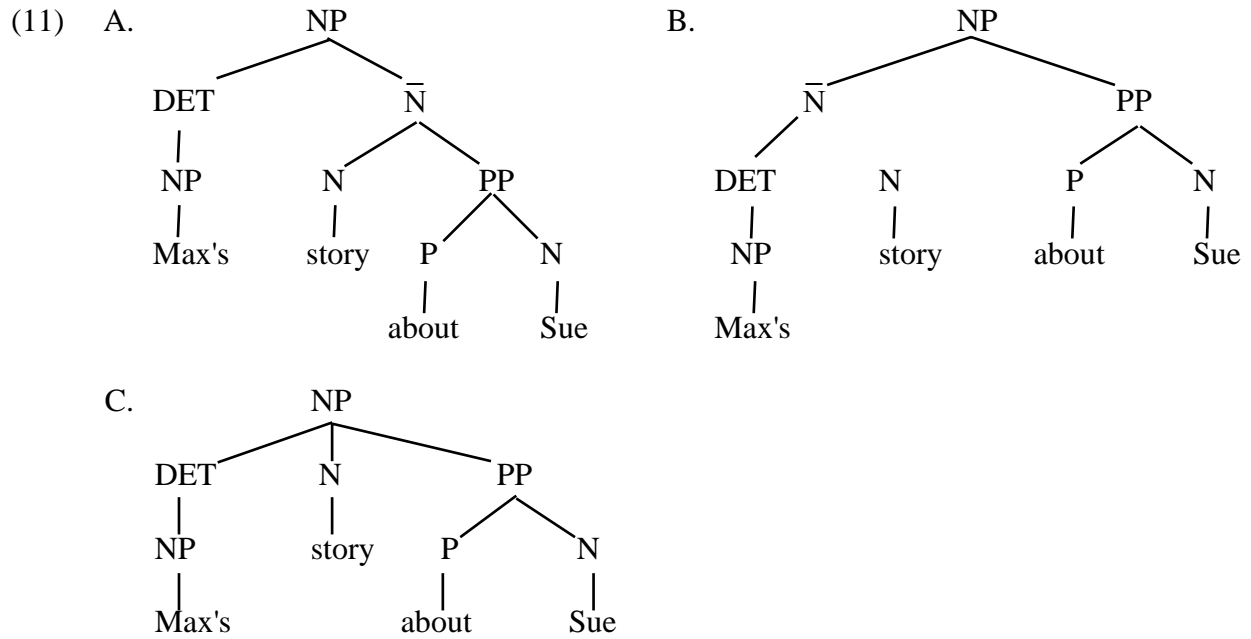
The evidence from Quechua is damaging to Ross' scheme and that from Polish is fatal.

I believe there is another problem with Ross' treatment of GAPPING. The directionality constraint presupposes only bifurcating nodes (immediate constituent analysis), and yet Greenberg's classification of languages based on 'dominant' order of major constituents and rules such as SCRAMBLING seem to presuppose that the major constituents are essentially of equal rank. In that case, they should be treated as sister nodes. E.g. SVO should have the following tree:



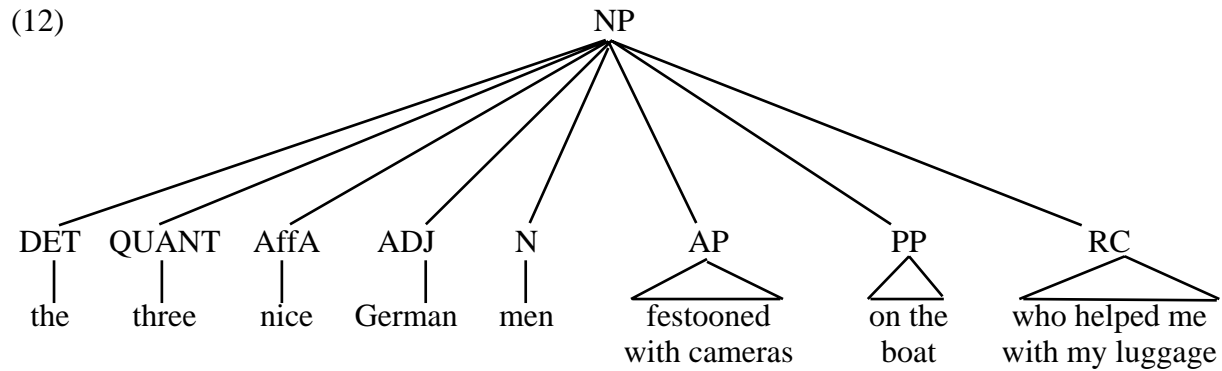
I take it that VSO or OSV must be diagrammed as three sister nodes and if this fact reflects the reality of their relative status then mere changes in the order of their arrangement should not alter that reality. To judge from semantic treatments of transitive verbs, which speak of them as being two-place predicates or predicates with two arguments, at least O and S should be viewed as equal in rank. Of course, one can always call on AUX or MOD or something else to guarantee that V will never be alone in the middle, and presumably the tree can be drawn so that V is on the desired side, but how does one quell the suspicion that the tail may be wagging the dog?

Ray Jackendoff's suggestion that GAPPING be applied to \bar{N} constructions²⁷ seems to me to involve the same problem. What is the proper tree for a noun phrase like 'Max's story about Sue'? I will compare three:



It seems to me that A and B are equally ad hoc and counter-intuitive--in neither case is \bar{N} a constituent within the dominating NP. Only C reflects the fact that the pre-modifier and post-modifier are equal in rank--DET modifies everything to its right and PP modifies everything to its left.

Perhaps a more challenging example will be more instructive, like "the three nice German men festooned with cameras on the boat who helped me with my luggage." I suggest:



Each of the modifying constituents modifies the head noun independently and in some sense modifies everything else in the NP. They are equal in rank, therefore. To introduce \bar{N} into the tree or to impose an IC analysis on the NP would be entirely ad hoc and would obscure, perhaps seriously, the real structure and the relationships obtaining among the constituents. But how might one phrase a directionality constraint for the above?

To conclude, although it may be useful to include GAPPING in the grammars of individual languages, it has yet to be demonstrated that such a rule either should or can be used to determine underlying constituent order.

FOOTNOTES

1. Apurinã (Ipuriná) is generally assigned to the Pre-Andine branch of the Arawak language family. There are around 1,000 speakers of Apurinã who are presently scattered along 1,500 kilometers of the Purus River in the state of Amazonas, Brazil.

2. Ross, John Robert. "Gapping and the Order of Constituents", Manfred Bierwisch and Karl Erich Heidolph eds. Progress in Linguistics. Mouton: 1970. p.250.

3. Greenberg, Joseph H. "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements," Joseph H. Greenberg ed. Universals of Language. M.I.T. Press: 1966. pp.73-113.

4. Ross, p.251.

5. Ibid., p.256.

6. Or, (a) ximaky ny-nika etc.
fish I-ate
- (b) ximaky ny-nika nota etc.
fish I-ate I
- (c) nota nika-ry ximaky etc.
I ate-it fish
- (d) ny-nika-ry ximaky etc.
I-ate-it fish
- (e) ny-nika-ry ximaky nota etc.
I-ate-it fish I
- (f) ximaky nika ny-txa etc.
fish ate I-it
- (g) nika ny-txa-ry ximaky etc.
ate I-#-it fish

All of the above would have essentially the same gloss although each has a context in which it is the unmarked form. Different ones of the above forms could be used with 'I', 'Peter' and 'Jim' in a single conjoined sequence, but each would be a complete clause in each case.

7. "Command in Apurinã" pp. 2-4

8. See my paper, "Relativization in Apurinã".

9. See my paper, "Interrogatives in Apurinã". "Command in Apurinã" p.5.

10. Ross, p. 259
11. Pulte, William. "Gapping and Word Order in Quechua," CLS 7 (1971).
12. Dingwall, William O. "Secondary Conjunction and Universal Grammar," PIL 1:2 (1969).
13. Greenberg, pp.110-11.
14. I wonder how useful the notion of any underlying constituent order will prove to be, once all the evidence is in. The fact that translation between two completely diverse (including surface constituent order) languages is possible, suggests that that which is universal has no linear order. Evidence from Aguaruna, a Jivaro language of Peru, wherein in narrative the verb normally follows the other constituents while in legend it normally precedes the other constituents (Mildred Larson, "A Method of Checking Discourse Structure in Bible Translation," Notes on Translation. 17 (1965)) gives rise to the hypothesis that differing constituent order may be a function of discourse type or genre. In such an event it could be misleading to speak of "the underlying order" for Aguaruna.
15. Ross, p. 249.
16. Ibid., p. 250.
17. Ibid.
18. Why should it be assumed that Russian, or any other language, has only one underlying order for major constituents?
19. Ross, p. 252.
20. Ibid.
21. Ibid.
22. Ibid., p. 253.
23. Dingwall, pp. 216, 227.
24. Pulte, p.194.
25. Ibid. Ross' scheme makes no provision at all for this sort of thing, nor for the existence of languages with a basic order other than SVO or SOV.
26. Ross, p. 256.
27. Jackendoff, Ray S. "Gapping and Related Rules," LI 2:1 (1971).