# CLS 28

## Papers from the 28th Regional Meeting of the Chicago Linguistic Society 1992

### Volume 2: The Parasession
### The Cycle in Linguistic Theory

compiled and edited by

Jeannette Marshall Denton
Grace P. Chan
Costas P. Canakis

heory," in *Alternative Conceptions of*
al., The University of Chicago

'. 93-121.
celandic," Syntax and Semantics

inding Category in Danish,"
23, Department of Linguistics,

LA 18. 210-234.
h and Icelandic", LA 15.
al Grammar," ms. Seoul National



# THE ORIGINS OF THE CYCLIC PRINCIPLE*

Geoffrey K. Pullum
University of California, Santa Cruz

## 1 Introduction

This paper is about the origins of the cyclic principle of rule application in syntax — origins in two senses: (i) the origins of the cyclic principle constraint in the literature of transformational-generative (TG) syntax, and (ii) the origins of the property of grammars and/or languages that the syntactic constraint seeks to encapsulate.

The outlines of the history of the cyclic principle in the literature have an all too familiar character. The initial characterizations of the principle failed to accomplish the intent of those propounding it; the attempts at explicitly motivating it mostly involved invalid arguments; and when a potentially valid form of argument for it was developed it was rapidly undercut by theoretical revisions; and its ultimate abandonment by most TG linguists had a lot more to do with intellectual fashion and political orthodoxy than factual evidence.

Yet the cyclic principle contains a hint of something real, an imperfect reflection of a true generalization about the character of the syntactic systems humans use. It is that hint of verisimulitude that I will be grasping at in this paper; in this sense the critical review of argumentation below is only a preliminary. I argue in the concluding sections that the grain of truth makes the principle even more general than most linguists have imagined, general enough that it goes beyond the domain of linguistic phenomena.

## 2 Formulations of the cyclic principle

The cyclic principle is part of the definition of the way in which a set of transformations defines a derivation and thus a structural description for a sentence. Like many 'definitions' in generative syntax throughout its history, it has few of the characteristics of definitions in formal theories.

The cyclic principle originates in phonology, with the discovery reported in Chomsky, Halle, & Lukoff (1956) that certain rules of English suprasegmental phonology appear to operate "cyclically to successively more dominant constituents of the surface structure" as Chomsky (1965:29) puts it. Formulations of the analogous principle in syntax are generally quite similar to the one given by Chomsky (1965:143):

> Given a generalized Phrase-marker, we construct a transformational derivation by applying the sequence of transformational rules sequentially, "from the bottom up" — that is, applying the sequence of rules to a given configuration only if we have already applied it to all base Phrase-markers embedded in this configuration.

Many restatements of the principle follow Chomsky almost slavishly (Akmajian & Heny (1975:362), expounding the principle of the cycle in their own words, use the phrase "from the bottom up" and put it in needless scare quotes exactly the way Chomsky does). Statements of the principle may be found in numerous textbooks, monographs, and articles on TG syntax, for example, Jacobs & Rosenbaum (1968:235–6), Grinder & Elgin (1973:145), Bach (1974a:110), Keyser & Postal (1976:382), Baker (1978:183), Radford (1981:201), and so on through the years. Nearly all of them (however generously interpreted) are inadequate, lacking the consequences that the cyclic principle was originally intended to have.

An argument intended to illustrate this is presented by Perlmutter & Soames (1979:154–171). Very briefly, and without going into details of the data (like word order and case marking) that would lengthen the exposition, the argument is based on subject-to-subject copy raising and passivization in Modern Greek. We can get a rough sense of the argument from analogous English data if we assume (for now) that the English *look like* construction is derived by copy raising. That is, assume that strings of the form (1a) yield strings like (1b) under a copy raising rule. Notice that if passivization were free to apply after that on the embedded clause it would yield strings like (1c).

(1) a. [$_S$ it looks like [$_S$ the thieves are killing the store owner]]
b. [$_S$ the thieves look like [$_S$ they are killing the store owner]]
c. *[$_S$ the thieves look like [$_S$ the store owner is being killed by them]]

This is not in fact very convincing as an argument about English; notice that sentences like *This room looks like an army of occupation has moved in*, with no pronoun in the complement clause, suggest that there will be a source for (1c) anyway, and thus that (1c) may be bad simply because intrasentential anaphoric pronouns often sound bad in a passive *BY*-phrase, e.g. **I swear you are loved by me* – an oft-noted constraint that is rather hard to pin down precisely (see Pullum 1979:301–309). I have not studied the Greek case directly myself; I get the impression from Moore (1992) that it too may have problems, but it will do to illustrate the type of possibility I am talking about

The trouble with the standard formulations of the cyclic principle is that they do not rule out such possibilities; they merely require that "applying the sequence of rules to a given configuration only if we have already applied it to all base Phrase-markers embedded in this configuration." But this condition would be *met* when passivization was applied to the embedded clause in (1b) to derive (1c): there is no further embedded clause on which transformations have not applied. The formulation fails to exclude *returning* to an embedded clause during the cycle on a higher clause and applying to it another cyclic rule that was not selected for application during the cycle on the embedded clause.

What is called for (as Perlmutter & Soames argue) is an additional clause with the kind of content that Chomsky (1973:243) expresses in his Strict Cycle Condition:



> The Strict Cycle Condition
> No rule can apply to a domain dominated by a cyclic node *A* in such a way as to affect solely a proper subdomain of *A* dominated by a node *B* which is also a cyclic node.

This is still fairly loose, of course. For example, 'dominated by' seems to be intended to convey 'rooted in': *B* is the root of the proper subdomain, not merely some arbitrary higher node that dominates it. But it can be tightened. Thompson (1975) makes a partially successful effort to elucidate the cyclic principle and the Strict Cyclicity Condition more clearly within a formalized theory of generative grammar (a version of Lakoff's 'correspondence grammar' closely allied to classical TG and relational grammar):

> If it is possible to construct a derivation in such a fashion that some [transformational] rule could apply to some tree in that derivation in two different ways, such that some variable in the rule instantiated to two different nodes in the tree *of different depth* in the two applications, then the rule is not a well-formed rule (for cyclic application). (Thompson 1975:601)

This is focussed on excluding a *rule* type rather than a class of derivations, but brings out something like the same intuition: what the Strict Cyclicity Condition aims to forbid is application in a manner that is promiscuous regarding depth of embedding of the material operated on. Such a constraint would prevent the application of subordinate passivization after raising on a superordinate cycle (though of course, that does not establish that the Strict Cyclicity Condition is *necessitated*; see Moore (1992) for an argument that under current TG assumptions certain conditions on chains remove the need for the Strict Cyclicity Condition in the Greek case).

## 3 Linear order and the Reversible and Sandwich arguments

The classic original arguments for the cyclic principle are not arguments for the Strict Cyclicity Condition, but only for cyclic application *simpliciter*. They are based entirely on the assumption that the set of transformations in a grammar is associated with an ordering (total or partial) stipulated as part of the grammar. But this assumption was never justified by any stretch of the imagination in syntax (and in phonology it is subject to a good measure of doubt).

Part of what kept linguists confused about this issue for some years was that there seemed to be good evidence for allowing rules to feed one another and thus apply in sequence. The notion that rules could apply to other rules' outputs and provide other rules' inputs somehow became confounded with the notion that some statement in a grammar should stipulate *when* they could do this. And then the latter notion was equated unthinkingly with the claim that the only way to do this was to stipulate a single linear order defined on the entire rule set to define the invariant sequence of rule applications. In perhaps the most explicit work in the literature of generative grammar, Chomsky & Halle (1968), we find this statement justifying linear ordering of rules (p. 18):

(13) It is always possible to order the rules in a sequence and to adhere strictly to this ordering in constructing derivations without any loss of generality as compared to an unordered set of rules or a set ordered on a different principle.

(14) Such linear ordering makes it possible to formulate grammatical processes that would otherwise not be expressible with comparable generality.

Notice that the claims made are first that it is *possible* to impose a linear order to the rule set (regardless of whether it is either necessary or desirable), and second that there are improvements in the capturing of generalizations. But of course, Chomsky & Halle had not tested linear ordering against all possible alternatives; even where they gave arguments against other ways of ordering rules (see pp. 19, fn. 5, and p. 342), the arguments were based on rules for imaginary languages. What was missing was an attempt to show that rule ordering was *necessary* in some situation in some language, and a recognition that it was not whether rules were applied in a strict sequence that was important for linguistic theory, nor whether the ordering was imposed "extrinsically" or "intrinsically" (in terms of the thoroughly obscure distinction drawn by Chomsky (1965:223, n. 6) and critiqued by Pullum (1979:11–15)), but whether grammars had to stipulate rule orderings on a parochial (nonuniversal) basis.

For a number of years (roughly 1969 – 1976), Andreas Koutsoudas and his students and colleagues developed and elaborated arguments that it was never necessary to stipulate rule orderings parochially. Their work was widely published but almost completely ignored by linguists elsewhere. Then, suddenly, Chomsky & Lasnik (1977:431) announced without warning or argument — and without citing Koutsoudas or anybody else — that syntactic rules were unordered, and from that day on there was no more talk of rule ordering in TG. Parochial ordering was abandoned as mysteriously as it was introduced — yet bafflingly, the cycle was retained (see Chomsky & Lasnik 1977:429).

Chomsky and Lasnik did not note that all the original arguments that had been given for the cycle had depended entirely on the principle of linear ordering as expressed by Chomsky & Halle, but this is certainly the case. The idea of cyclic application of transformations is in essence due to Fillmore (1963), who illustrates the motivation for "re-cycling through" the transformations that apply to simple clauses by pointing out that in the derivation of (2) there is an application of the passive transformation, then an application of the embedding transformation that embeds an infinitival clause after *BELIEVE*, and then another application of the passive transformation.

(2) The butler is believed to have been murdered.

The implicit argument is that since we can say neither that passivization always precedes embedding nor that embedding always precedes passivization, we have to have a "re-cycling" mechanism.



The straw-man theory implicitly assumed here (in order to be rejected) is that each transformation on an ordered list is applied at all points in the tree where it can be and then discarded permanently, the derivation proceeding until the last transformation on the list has been applied to the tree. The only alternative to this straw man that was ever seriously defended was the "linear" or "iterative cyclic" theory advocated by John Robert Ross in unpublished work *circa* 1967 and discussed in formal terms by Kimball (1967, 1972), and this still assumed parochial ordering; briefly, iterative cyclic application meant application of each individual rule in the ordered list from bottom to top — transformations are still ordered, but each transformation has its own bottom-up cycle.

One type of argument in favor of the cyclic principle was called the "Reversible" argument by Pullum (1979:105–110). An argument of this type involves a crucial but usually suppressed premise: (i) rules apply in a strict order; from there, the argument says simply: (ii) here is a derivation with $T_1$ applying before $T_2$; (iii) here is a derivation with $T_2$ applying before $T_1$; therefore, (iv) there must be a cycle. This is, of course, not convincing even if rules *are* ordered (as Grinder (1972) tried to show), but it is flagrantly invalid without (i), for which good motivation was completely lacking. Yet textbook after textbook repeated the various Reversible arguments. Jacobs & Rosenbaum (1968, ch. 28) note a sentence in which subject-to-object raising applied before reflexivization and a sentence in which reflexivization applied before subject-to-object raising, and conclude that the cycle is necessary; Bach (1974a:120–124) does the same thing; Akmajian & Heny (1975) use equi-NP deletion and subject-to-subject raising to the same end; and so on.

Much the same is true for what Pullum (1979) calls "Sandwich" arguments. Lakoff (1968) noted essentially the same facts that Fillmore had pointed to (though Lakoff wrote after embedding transformations had been replaced in Chomsky (1965) by base generation of subordinate clauses), and noted that not only can passive both precede and follow subject-to-object raising in a single derivation, but another application of raising can follow a raising + passive sequence, and this alternation can be extended indefinitely, so that no possible finite listing of rules could match all the combinations; Grinder & Elgin (1973:144–145) repeat this argument, and so do many other works.

But if transformations can just apply whenever their structural descriptions are met, there is nothing necessarily problematic about Reversible or Sandwich situations; whether the right orders can be achieved in particular cases, whether by allowing free application or by employing universal constraints on ordering that have different effects in different circumstances, was a matter for investigation. Pullum (1979) undertook some of that investigation, and found that none of the standardly cited alternations were difficult to predict from plausible assumptions or universal principles even *without* the cyclic principle. (Paradoxically, assumption of the cycle makes it yet easier to remove all support for parochial rule ordering and thus to undercut the Reversible and Sandwich arguments for the necessity of the cycle; as argued by George Lakoff in 1971

lectures at the University of California, Santa Cruz (see Grinder (1972:110), and illustrated further by McCawley (1984, 1992), many potential cases for parochial rule ordering dissolve once the role of the cycle is properly recognized.)

## 4 Antecedent Removal arguments and their subversion

In search of what Grinder (1972) calls "primary motivation" for the cycle (motivation that makes the cycle necessary rather than just acceptable or workable), new types of argument were developed. The most important one was called the "Antecedent Removal" argument type by Pullum (1979:110ff). A clear presentation is set out by Grinder (1972:95–97). It is based on the observation that an obligatory rule of anaphora that needs an antecedent NP in a certain position will, in some permitted derivations, be deprived of that antecedent when a rule applying in a superordinate domain moves that NP away to a position where it is not available as an antecedent. This type of situation is seen in examples like (3a), which has a deep structure of the general form shown in (3b). (To highlight bracketing contrasts, in this section I will show a clause as tensed if it has a subject and infinitival otherwise.)

(3) a. We believe ourselves to have shown ourselves to be competent.
b. [$_S$ we believe [$_S$ we have shown [$_S$ we are competent ] ] ]

Under the assumption that reflexive pronouns in clausal argument positions must be bound by antecedents in their own clauses, neither of the NPs in (3b) that end up as reflexive pronouns can be licensed as reflexive pronouns until they have been raised into their superjacent clauses by subject-to-object raising. But without the cyclic principle dictating that subordinate clauses must be transformed before superordinate ones, raising could apply in the topmost clause first, with the result that there would never be a way to get the second reflexive pronoun licensed, even when a second application of raising occurred; the offending derivation would go as seen in (4).

(4) a. [$_S$ we believe [$_S$ we have shown [$_S$ we are competent ] ] ]
b. [$_S$ we believe we [$_S$ to have shown [$_S$ we are competent ] ] ]
c. [$_S$ we believe ourselves [$_S$ to have shown [$_S$ we are competent ] ] ]
d. [$_S$ we believe ourselves [$_S$ to have shown we [$_S$ to be competent ] ] ]
e. [$_S$ we believe ourselves [$_S$ to have shown us [$_S$ to be competent ] ] ]
f. *We believe ourselves to have shown us to be competent.

This does not depend on rule ordering. For example, it would make no difference if steps *c* and *d* were reversed; what is crucial is that it is possible to raise an antecedent too early (in step *b*), and the result will be an obligatorily bound anaphor without an antecedent.

However, as pointed out by Grinder (1972:97f), this argument would be completely undercut if one assumed that raising rules leave phonetically null (or subsequently deleted) copy pronouns in the pre-raising position of the raised NP.



MIT-influenced work proceeded almost immediately to adopt not only this claim but a more general one: that *all* movement rules leave a coindexed element in the pre-movement site (see Wasow (1979:159–161 and references cited there). Perlmutter (1972) took these elements to be anaphoric pronouns ("shadow pronouns"); most current work calls them "traces" and has concluded that they are *not* pronouns (but have syntactic behavior more like that of nonpronominal definite NPs); but whatever they are, if they can be antecedents for reflexives etc., Antecedent Removal arguments are completely undercut. For example, if subject-to-object raising leaves traces, there is no problem in guaranteeing the correct binding in (3a), with every reflexive bound to a subject in its own clause:

(5) [$_S$ we$_i$ believe ourselves$_i$ [$_S$ $t_i$ to have shown ourselves$_i$ [$_S$ $t_i$ to be competent]]]

But under Chomsky's assumptions the undermining of the Antecedent Removal argument for the cycle is quite independent of traces, because of his rejection of two other crucial claims: (i) that there is a rule of raising into object position, and (ii) that reflexive binding respects clause boundaries.

It has been relatively little remarked in the literature that the clause-boundedness assumption about English reflexives seems to be untenable anyway, regardless of raising. Although I find no direct mention of it in Chomsky (1973) or Postal (1974), the clausemate condition is falsified by reflexive pronoun subjects of *for-to* clauses bound by a superordinate NP:

(6) a. They arranged for themselves to be arrested.
b. The deranged president called for himself to be impeached.
c. She opted for herself to be removed from the list of candidates.

If we take these facts to indicate ordinary (Principle A) binding of reflexives across a boundary, they invalidate the Antecedent Removal argument even if there are no traces and there is raising to object (on raising, see Lasnik & Saito (1992) for a recent even-handed review of the issues).

Thus under the assumptions increasingly current throughout the MIT-influenced part of the field of transformational syntax from 1973 onward, there has been no shred of valid rule-interaction arguments for the cyclic principle in syntax since the early 1970s.

## 5 Strict cyclicity and successive cyclic movement

Despite the lack of direct arguments for it, MIT linguists continued to assume the cyclic principle. Indeed, they regarded more rules as cyclic than other syntacticians did. Specifically, they regarded *wh*-movement rules as cyclic at a time when they were widely regarded as postcyclic.

If *wh*-movement is assumed to be cyclic and at the same time the conditions on its landing site are relaxed by collapsing it with all other movement transformations as "Move α" (the endpoint of the program begun in Chomsky

(1973)), there is a real problem with ruling out "improper movement" derivations in which *wh*-movement feeds NP movements like passivization, e.g. deriving from the string underlying (7a) the extremely ungrammatical (7b).

(7) a. It was never asked what effects this would have
b. *What effects were never asked this would have.

The once-widespread assumption of *wh*-movement postcyclicity would at least have the virtue of completely ruling out this sort of monstrosity. In current theories it is most unclear how the ungrammaticality of such strings can be explained. A story can be told about them being Principle C violations (because the *wh*-trace in the embedded CP specifier position is not free, being coindexed with the main clause subject), but it seems most unconvincing to say that (7b) violates nothing more than the pragmatically overridable bias against *Einstein knew that Einstein, they would believe*.

Since cyclic *wh*-movement bears a large share of the responsibility for making strings like (7b) generable, what are the arguments for it? The claim that *wh*-movement is cyclic is bound up with a very confusing part of the history of recent linguistics, namely Chomsky's (1973) argument that *wh*-movement is *successive cyclic*, i.e. that a *wh*-phrase moved to the front of sentence from a position $n$ levels of embedding below the main clause would move $n+1$ times to get to their surface positions, once on each clausal cycle. What is remarkable about the wide acceptance gained by this claim is that Chomsky's theoretical justification for it was completely mistaken, and his factual illustration of it used highly misleading and unrepresentative data.

Chomsky's theoretical justification for postulating successive cyclic *wh*-movement is that from the Strict Cyclicity Condition "it follows that *wh*-Movement must be a cyclic rule, since it applies in indirect questions and relatives." But this does not follow at all. It does not in any way dispose of the position that *wh*-movement might be postcyclic. *Any* postcyclic transformation $T$ must violate the Strict Cyclicity Condition, because $T$ must apply in some domain $D$, and by the definition of postcyclicity, if $T$ is postcyclic the cycle on $D$ will be over before $T$ applies. If the Strict Cyclicity Condition is to apply to postcyclic rules, then it applies trivially, because no postcyclic rule can ever apply in any domain. But Chomsky (1973) accepts that there are some postcyclic rules.

In any case, perhaps the most plausible position within classical TG about *wh*-movement is not that it is postcyclic but that it is *trigger-cyclic*, the claim defended by Pullum (1979:235). It would apply on the cycle most immediately containing the trigger that defines its landing site — the specifier or COMP position in an interrogative clause, or whatever. It would apply just once, from the argument position in which previous cyclic rules had placed it, direct to its nonargument position in surface structure. That would be completely compatible with the Strict Cyclicity Condition.



No direct factual justification of successive cyclic *wh*-movement was attempted in Chomsky (1973), and the introductory illustration of its operation was extraordinarily misleading. It consisted in showing two steps in the derivation of (8).

(8) What did you tell me that Bill saw?

Chomsky exhibits the deep structure (9a), to which he says "we first assign *wh* and apply *wh*-Movement on the innermost cycle," and then exhibits the resultant stage (9b), which is the first warning in Chomsky's paper (or anywhere in the literature of linguistics) that COMP-to-COMP hopping *wh*-movement is to be assumed.

(9) a. COMP you told me [ COMP Bill saw something ]
b. COMP you told me [ [$_{COMP}$ what] Bill saw ]

Now, it just so happens that (9b) looks like (10), which is grammatical — *but not related at all to the structure of (8)*.

(10) You told me what Bill saw.

It happens that *TELL* has a dual subcategorization: it takes either an indicative declarative *that*-clause or an interrogative *wh*-clause. If Chomsky had picked less misleading matrix material like *hope* instead of *tell me*, he would have been exhibiting this derivation:

(11) a. COMP you hoped [ COMP Bill saw something ]
b. COMP you hoped [ [$_{COMP}$ what] Bill saw ]
c. What did you hope that Bill saw.

This would have encouraged readers to notice that the stage (11b) corresponds to the glaringly ungrammatical (12).

(12) *You hoped what Bill saw.

The structure (11b) constitutes a stark counterexample to the otherwise plausible constraint on derivations suggested by Bach in a number of places (Bach 1974a; Bach 1974b; Bach & Horn 1976:297). It is hard to resist the conclusion that successive cyclic *wh*-movement was slipped by the reader in a way designed to prevent anyone noticing what a strange and radical innovation it was.

I return in section 7 to consider some of the positive evidence for it that was subsequently presented.

## 6 Residual arguments for the cycle: cyclic structure constraints

There is a class of overlooked potential arguments for the cycle that depend not on showing that pairs of rules have to interact in a certain way, but on exhibiting constraints stated on a significant level of structure that is only definable if the cycle is assumed: the level of *cyclic structure*, the output of the last

transformation to apply to a given domain of application. Pullum (1979:155–224) surveys twenty constraints that appear to make reference to cyclic structure. There is insufficient space here to examine and re-evaluate each of them in detail, but in (12), as a starting point and reference checklist for future research and re-examination, I provide a list of them with a brief description and a brief example or reference. Some refer to the term *quasi-clausemates*; the definition is as follows: *A* and *B* are quasi-clausemates iff every clause that includes *B* but not *A* has lost its cyclic structure subject by the end of the cycle on the minimal clause containing *A*. (Note the crucial dependence of this definition on the notion of cyclic structure.) Page references in the following list are to Pullum 1979.

(13) 1. *Extraction from subjects.* Constituents properly contained in cyclic subjects cannot be extracted. The reference cannot be to deep structure (because cyclically derived subjects are islands) or surface structure (because extraposition of a sentential subject restores extraction possibilities) or application point of the extraction rule (because raising-derived objects are islands if they were cyclic subjects); see Jacobson & Neubauer (1974). [155–9]

2. *Contraction of TO. Want + to* can contract to *wanna* if the two elements are quasi-clausemates (this is probably not the right formulation, but it is a lot closer to adequacy than most of its rivals have been; see Postal and Pullum 1978, 1982). [159–61]

3. *Auxiliary reduction.* If an account of auxiliary reduction anything like that of Selkirk (1972) is adopted, the point in the derivation at which it is appropriate to insert the phonological boundary markers whose interaction with subsequent movements determines reducibility is not (*pace* Selkirk) deep structure, but rather the cyclic structure of the auxiliary verb's clause. [162–8]

4. *Participial inflection.* The English participial inflections perfect *–en* and progressive *–ing* appear on verbs which are immediately c-commanded by perfect *HAVE* and progressive *BE*, respectively, in the cyclic structure of the minimal clause containing the *HAVE* or *BE*. [168–72]

5. *Predicate agreement.* Predicates agree with their cyclic-structure subjects (this covers those cases of agreement with cyclically moved subjects in Ancient Greek for which Avery Andrews argued for a global account). [172–5]

6. *Agreement with expletives.* If the cyclic structure of a clause has an expletive as subject of a verb, that verb either takes third person singular form or agrees with the NP to which the dummy is linked (its brother-in-law in relational grammar terminology). [175–6]

7. *Sluicing identity.* The identity condition on the ellipsis construction known as Sluicing (Ross 1969) compares not surface structures (because the bracketed portion can be elided in *Kim has been dating someone, but who [has Kim been dating]?* despite not being identical to the surface form of the first clause) but cyclic structures. [176–9]

8. *Seem*-class subjects. A *seem*-class verb must have a predicative or clausal complement in cyclic structure (note **That things happen seems (to me)*, contrasting with **It seems (to me)* that things happen: *That things happen seems (to me)* to be indubitable). [180–81]

9. *The Doubl-ing Constraint.* (Ross 1972) Two *–ing*-inflected verbs cannot be adjacent in surface structure if they are quasi-clausemates. [182–4]

10. *Conjoining dissimilar illocutions.* It is not possible to conjoin mixtures of declarative, imperative, and interrogative sentences. Some classical TG analyses involving deep structure coordinate sources are incompatible with stating the constraint on deep structure, and at surface structure there may be too little information to identify things like *wh*-in-situ questions as interrogative; the cyclic structures of the conjoined clauses contain the right information. [184–6]

11. *The Inclusion Constraint.* Subjects cannot overlap with objects in reference (*??I heard us*). Antecedent Removal arguments exclude a deep structure statement, and certain Gapping outputs (*You think he saw her, and we think me*) may exclude surface structure; cyclic structure is the right level to access. [186–9]

12. *Strong crossover and the structural constraint on pronominal coreference.* A pronoun must not c-command its antecedent in cyclic structure of the smallest domain that includes the pronoun and the antecedent (see Jackendoff 1972). Strong crossover facts, for example (e.g. **Who$_i$ did he say Mary kissed $t_i$*) follow readily under this approach. [189–94]

13. *Weak crossover.* Postal's (1972) global constraint on pronominalization, now known under the heading of weak crossover, can be improved in formulation if it makes reference to cyclic structure. [194–8]

14. *Linearization of constituents.* The cyclic structure of a clause appears to be the earliest derivational stage at which left-right order of constituents is critically relevant (for describing the Intervention Constraint of Jacobson & Neubauer (1976) and a variety of other phenomena); order in underlying structures could be ignored, ordering constraints being imposed successively, domain by domain, on cyclic structures. [199–210]

15. *Subject-oriented adverbs.* Adverbs like the *cleverly* of *John cleverly has been examined by doctors* have to be associated semantically with whatever is the cyclic structure subject of the clause they are in. [210–12]

16. *The Like Subject Constraint.* The constraint on obligatory control verbs like *try* is not, *pace* Perlmutter (1971) and others, that they must have a deep structure subject coreferent with the subject of their complement, and it is not a surface constraint either (note: **Who$_i$ did you try ___$_i$ to fix your car?*); the identity of matrix and complement subjects is required at the cyclic structure of the matrix clause. [212–14]
17. *Imperative subjects.* The constraint on second-person imperatives is not that they must have a second person subject in deep structure (note *Be warned!*), and clearly cannot refer to surface structure, where ordinary imperatives have no subject at all; the constraint is on the main clause cyclic structure. [214–15]
18. *Root modals.* Animate subjects are strongly favored for root sense modals. The level at which the subject must be animate is not deep structure (note that *The grass won't be cut by John* does not mean 'John refuses to cut the grass') or surface structure (where the subject may have been moved or deleted), but the cyclic structure of the clause most immediately containing the modal. [216–17]
19. *Subject persistence.* Languages like English and French that have to have subjects in finite clauses do not necessarily have them in deep structure (if the Unaccusative Hypothesis is true for any predicates) and do not necessarily have them in surface structure (because of casual speech rules, imperatives, etc.); the level at which every clause has a subject is cyclic structure. [217–19]
20. *Instrumental adverbial clauses.* The instrumental adverbial clauses used by Fodor (1970) in arguing against syntactic lexical decomposition (note the contrast between *Ian$_i$ caused Jay$_j$ to die by PRO$_j$ taking poison* and **Ian$_i$ killed Jay$_j$ by PRO$_j$ taking poison*) are not constrained in the way Fodor argues, to have deep structure subjects identical to the deep structure subject of the matrix clause; they have to have cyclic structure subjects nondistinct from the cyclic subject of the matrix clause (either unspecified or controlled by the matrix clause cyclic subject). [219–24]

Many of these suggested constraints may be subject to reanalysis or elimination under current assumptions; I know some of them are, but I also know that some of them appear still to merit attention; there has been no systematic demonstration that all the above constraints can be reassigned to other levels such as LF or an 's-structure' more abstract than classical surface structure.

Bach's proposed "local grammaticality" condition mentioned above in section 5, adds to the theoretical significance of the level of cyclic structure — a level that cannot be defined if the cyclic principle is not assumed.



## 7 Picture nouns, successive cyclicity, and reconstruction

I argued in section 5 that successive cyclic *wh*-movement was introduced in a way that must be judged to lie somewhere between confusion and subterfuge. But after it had languished for some years without a shred of support, some remarkable evidence began to turn up that at least looked as if potentially it could offer support for what Chomsky had posited. I am aware of potentially relevant evidence from Chamorro, French, English, Icelandic, Irish, Palauan, and certain Bantu languages. References to the literature on most of the languages can be found in Zaenen (1983), an interesting attempt to account for some of the phenomena involved *without* using successive cyclic movement. *Wh*-agreement in Chamorro and Palauan, described by Chung (1982, 1992), Georgopoulos (1985), and Chung and Georgopoulos (1988), is not dealt with by Zaenen; and whether such facts would submit to analysis in similar terms is a matter for debate (see Goldberg (1985) for an initial contribution).

Various arguments *against* successive cyclic movement were also published during the 1970s; Postal (1972) and Bach (1975) were early dissenters discussing English; Epée (1976) presented an argument from a West African language, Duala. Pullum (1977) argued that evidence from French and Swahili looked just as likely to disconfirm successive cyclicity as to confirm it.

I have no space to discuss here all the interesting sets of facts involved. Here I just want to point out an example from English of the way that what looks at first like subtle and interesting confirmation for successive cyclicity can dissolve away and leave nothing of the argument remaining.

When a picture-noun phrase containing a reflexive or reciprocal anaphor is fronted by *wh*-movement, it can be bound not only by (i) the subject of the matrix clause immediately superjacent to it or (ii) the subject of its d-structure clause, but also, more surprisingly, by the subject of any intermediate clause it is extracted from. Thus in (14), no less than three different reflexive pronouns are grammatical in the same NP position.

(14) a. He$_i$ knew which photos of himself$_i$ they claimed she had.
b. He knew which photos of themselves$_i$ they$_i$ claimed she had.
c. He knew which photos of herself$_i$ they claimed she$_i$ had.

Moreover, as seen in (15), if the reflexives are replaced by plain personal pronouns, the results are much worse (the structures violate Condition B).

(15) a. *He$_i$ knew which photos of him$_i$ they claimed she had.
b. *He knew which photos of them$_i$ they$_i$ claimed she had.
c. *He knew which photos of her$_i$ they claimed she$_i$ had.

This does not mean that picture-noun reflexives can be assigned antecedents arbitrarily. Examples like (16a), in which the only available antecedent for the reflexive is a subject two clauses above it, seem ungrammatical, the alternative in (16b) with a plain personal pronoun being more acceptable.

(16) a. *We$_i$ said he knew which photos of ourselves$_i$ she had.
b. We$_i$ said he knew which photos of us$_i$ she had.
c. *Which photos of himself$_i$ convinced you that he$_i$ thought she had been spying on him?
d. *They claimed he$_i$ thought she had taken some photos of himself$_i$.

(I say "seem ungrammatical" because the judgment on such sentences will be revised below.)

Under the hypothesis that *wh*-movement (assuming one trace per extracted phrase, left in the immediately pre-extraction position) operates in a single long-distance movement, the grammaticality of sentences like (14a) is something of a mystery. Neither the s-structure nor the d-structure position of the extracted phrase has the reflexive in a position where it could be bound, except under assumptions that would mean sentences like those in (16) would be predicted grammatical.

Suppose, however, that we made the following assumptions:

(i) *Wh*-NP movement involves successive cyclic movement.

(ii) All domains — or at least, certainly CP and IP — define cycles.

(iii) Binding of anaphors contained in moved phrases is done optionally in cyclic structures.

This would appear to offer a way to achieve the necessary bindings. The derivation of, e.g., (14b) would be as indicated in (17).

(17) He knew $\boxed{\text{[which photos of themselves}_i\text{]}_j}$ they$_i$ claimed [$_{CP}$ $t_j$ she had $t_j$]

After the first movement, the *wh*-phrase containing the reflexive would be in a position where it could be bound to a c-commanding (in fact IDC-commanding) subject in the immediately superjacent clause in a structure with no intervening potential antecedent.

Huang (1993) is concerned to provide an explanation for a further interesting fact. The binding possibilities are not the same with anaphors contained in fronted "predicative" phrases. Thus of the cases in (18), only (18c) is grammatical.

(18) a. *How proud of themselves$_i$ did they$_i$ claim he thought she had been?
b. *How proud of himself$_i$ did they claim he$_i$ thought she had been?
c. How proud of herself$_i$ did they claim he thought she$_i$ had been?

(Similar but somewhat more awkward examples can be constructed with fronted VPs, as in *Criticize ourselves, they agreed that we did* as opposed to **Criticize themselves, they agreed that we did*, and also with predicative NPs.) Huang's account of these facts involves an extension of the 'VP-internal subject' hypothesis — that subjects are generated inside predicative categories and moved



out to higher positions. He tries to show that although the structures of sentences in which predicative categories are extracted are much the same as where nonpredicative constituents have been extracted, with traces in intermediate positions, the subject trace in an extracted predicative category will not be bound in its minimal complete functional complex except with regard to its original d-structure position, so only for that position do we get a well-formed output.

Huang does not propose that binding is done during the cycle as I have suggested it might be; he assumes "reconstruction" at the LF level based on the traces left by *wh*-movement. He does not go into detail about what "reconstruction" means or how it solves the problem of getting the right bindings, and nor will I. It suffices to note that his account is quite compatible with a cyclic binding account — something like what Barss (1986) proposes — and crucially presupposes not just traces of the form [$_{NP}$ *e*] as seen in the abbreviated (17) but full copies of the extracted constituents at all trace positions, which makes the LF structures involved highly complex.

Huang's "reconstruction" looks to me like a metaphorical reference to some as-yet undeveloped procedure for deducing from s-structures what the cyclic structures of all the subordinate clauses looked like, so although it is more obscure than a Barss-style account, it is apparently not going to be different in its predictions. Hence one might conclude that there was a novel argument for successive cyclic *wh*-movement here, since only successive cyclic movement of the phrase containing the picture-noun phrase could place it in the right position to be bound in cyclic structure (or leave traces in just the right positions for the reconstruction account to work on).

The potential argument from these facts collapses, however, when faced with a wider range of data than considered so far. Observations made briefly by Pollard & Sag (1983) and Hukari (1989:219–223) and pursued in more detail by Pollard & Sag (1992) show this rather clearly. Pollard & Sag characterize a class of "exempt anaphors" — roughly, anaphors cannot be coindexed with a coargument (argument of the same predicate) that outranks them on the grammatical relations hierarchy. Consider here just a few of the sentences Pollard & Sag cite (from their pp. 264, 279, and 296) to show that "the antecedent of an exempt anaphor may appear (modulo nongrammatical factors) lower in the structure, higher in the structure, or even in the prior discourse context" (p. 296):

(19) a. John$_i$'s intentionally misleading testimony was sufficient to ensure that there would be pictures of himself$_i$ all over the morning papers.
b. Which pictures of herself$_i$ did it seem to Mary$_i$ that they were going to publish?
c. Which picture of herself$_i$ did John say was most to Mary$_i$'s taste?
d. John$_i$ was furious. The picture of himself$_i$ in the museum had been mutilated.
e. John$_i$ was quite agitated. It was uncertain which picture of himself$_i$ would appear in the *Times*.

f. The picture of herself$_i$ on the front page of the *Times* made Mary$_i$'s claims seem somewhat ridiculous.
g. The pictures of herself$_i$ on the front page of the *Times* confirmed the allegations Mary$_i$ had been making over the years.
h. Bill$_i$ finally realized that if the *Times* was going to print that picture of himself$_i$ with Gorbachev in the Sunday edition, there might be some backlash.

What, then, is the explanation for the apparent ungrammaticality of such sentences as (16a), **He$_i$ said we didn't know which photos of himself$_i$ they claimed she had taken*, cited above? The answer Pollard & Sag provide is that there is a constraint that such sentences violate, but it is "a processing-based factor that interacts with grammatical constraints in such a way as to render unacceptable a family of sentences that are otherwise grammatical" (p. 273).

Pollard & Sag offer not only counterexamples to the kind of assumptions about structural constraints on anaphora that previous studies have made, but also a full analysis of their own, one which not only allows the cases just cited but additionally makes the right predictions about cases like those in (18), with anaphors contained in fronted predicate-headed phrases (this is not noted by Pollard & Sag, but it is easy enough to determine from the exposition they provide). I conclude that no argument for successive cyclic movement has yet been constructed on the basis of picture-noun reflexives.[1]

## 8 The non-formal character of the cyclic principle

If, counter to the trend in recent work, the cyclic principle *is* assumed to govern rule interaction in a transformational grammar, it becomes relevant that, I remarked above (in section 2), it is unfortunately very easy to make arbitrary rule interactions compatible with the cyclic principle. This can be readily illustrated by reference to McCawley's discussions (McCawley 1984; 1988:24ff, 154ff; 1992) of how the cyclic principle can predict the sequencing of rule applications, making parochial stipulations about word order unnecessary. McCawley notes that although a structure like (20a), in which IMP is a clause-initial marker of imperative mood, would meet both the structural description for Imperative Subject Deletion in (20b) and the structural description for Reflexivization in (20c), potentially allowing for the former to override the latter and yield (21a) instead of (21b), changing the structure to (22) would mean that this was not so; Reflexivization would apply on the inner cycle and Imperative Subject Deletion only on the outer cycle.

(20) a. $[_S$ IMP you$_i$ defend you$_i$ ]
b. ... IMP ... you ...
c. ... NP$_i$ ... NP$_i$ ...



(21) a. *Defend you!
b. Defend yourself!

(22) a. $[_S$ IMP $[_S$ you$_i$ defend you$_i$ ] ]

But the cycle could easily be made to predict the reverse of the correct orderings. We postulate a marker of clause reflexivity, REFL, and assign it a position in phrase structure above the root node of the domain including all the material that appears in the clause. In the case of the Imperative Subject Deletion and Reflexivization example above, we would postulate the deep structure (23a), and restate the structural description of Reflexivization as in (23b).

(23) a. $[_S$ REFL $[_S$ IMP $[_S$ you$_i$ defend you$_i$ ] ]
b. REFL NP$_i$ ... NP$_i$ ...

Now the cyclic principle, far from blocking the application of Imperative Subject Deletion before Reflexivization, enforces it (if both rules are cyclic, which I assume here, with McCawley, for the sake of argument). And many other assumptions would yield the same result — for example, that imperative clauses are identified by a feature marking on the verb or verb phrase but (as in earlier work of McCawley's) the content of NPs originates outside their clauses:

(24) a. $[_S$ $[_{NP}$ $x$ = you ] $[_S$ $[_{NP}$ $x$ ] $[_{V[+IMP]}$ defend ] $[_{NP}$ $x$ ] ] ]

Again, the cycle would force deletion of the subject of the imperative verb before the cycle on which the content *you* was lowered in.

Similar points could be made about other orderings that McCawley predicts from the cyclic principle, for example the interaction of Passivization and Quantifier Floating seen in (25), where Passivization of a clause is not permitted after a quantifier has been floated off its subject NP.

(25) a. Both the secretaries have praised the manager.
b. The secretaries have both praised the manager.
c. The manager has been praised by both the secretaries.
d. *The manager has both been praised by the secretaries.

The explanation, according to McCawley, is that quantifiers originate in higher clauses and are transformationally lowered (by a rule that is an exact inverse of the Quantifier Raising rule of LF familiar in current work and historically prior to it). The relevant deep structure would be roughly as in (26).

(26) $[_{S_1}$ $[_{NP_x}$ both $[_{NP}$ the secretaries]] $[_{S_2}$ $[_{NP}$ $x$] have praised the manager]]

Passivization is applicable to $S_2$, but the quantifier *both* is not even in that clause: it gets optionally floated, and the rest of the lexical content of the subject NP lowered in, on the $S_1$ cycle, and if Passivization has applied this will be blocked (since Quantifier Lowering needs a bound variable in the subject NP position): "there is no well-formed derivation in which both Passive and Quantifier-float

apply," McCawley observes, and this, unlike a rule-ordering stipulation, provides "insight into **why** Passive and Quantifier-float would interact in this way" (1988:598). But this depends on, for example, passivization not being triggered by a higher-clause marker of passive voice, call it PASS, to which passivization must make reference.

Of course, one could argue about whether elements like the REFL and PASS markers could possibly be motivated. But regardless of whether they might (which is not inconceivable; recall the work of Langacker and Munro (1975) on special properties shared by passive and reflexive but not active transitive clauses), arguing that way would be missing my point, which is about the sensitivity of the cycle to rule formulations and details of phrase structure. While McCawley prefers (reasonably enough) to see the situation as providing a research strategy in which structures and rules are designed to permit the cycle to do as much as possible of the work of guiding derivations and adjudicating rule interactions, the situation can also be seen in a different light: even if the cyclic principle is stated quite strictly, the notion that specific facts about languages follow from it, in the way that facts about the absence of center embedding *do* follow from restricting grammars to right-linear phrase structure rules, is thoroughly misguided.

The point I am making has nothing specifically to do with the use of abstract markers like IMP, REFL, or PASS, either (though such devices have always been admitted in TG, and are more in use now, with the fashionability of functional projections, than ever before). Because we have no way to individuate transformational rules (a fact which makes it much less surprising that one can collapse all movement transformations into the cryptic imperative "Move $\alpha$" by removing all context details) we can never rule out the possibility of obtaining countercyclic applications that should have been illicit by combining the structural description of the rule to apply in the embedded domain with aspects of the structural description for a perfectly legitimate transformational operation in a higher domain.

For example, since IMP is proposed as a marker appearing only on root clauses (imperatives being unembeddable), on the one hand the cycle could be evaded in imperatives by a rule that instead of having the structural description $\ldots X \ldots Y \ldots$ had the structural description forcing the rule to apply counter to the Strict Cyclicity Condition; and on the other hand, assuming IMP were somehow eliminated, and the non-embeddability of imperatives were captured instead by direct reference to rootness (recall the "where $\bar{S}$ is a root sentence" condition of filter (180) in Chomsky & Lasnik (1977:486)), that device (whatever it was) could be used for the same illicit purpose.

This point is not affected at all if we consider more precise formulations of the cyclic principle such as that developed by Thompson (1975); the marker-triggered rules discussed above would be fully legal under his formal and fairly strict statement of the cyclic principle.



Any suggestion that languages simply do not permit such special behavior in different construction types (the kind of rule-free-grammar rhetoric that is very common in recent linguistics) would face much embarrassment if confronted with even quite small fragments of the details of a well-known language like English. For example, the distribution of positive and negative *DO* in emphatic and nonemphatic declaratives, interrogatives, subjectless imperatives, and imperatives with subjects is startlingly erratic (italicization indicates emphasis):

(27) a. We *do* like you.
b. *We do *like* you.
c. We *don't* like you.
d. We don't *like* you.

(28) a. *Do* you like me?
b. D'you *like* me?
c. *Don't* you like me?
d. Don't you *like* me?

(29) a. *Do* touch me!
b. *Do *touch* me!
c. *Don't* touch me!
d. Don't *touch* me!

(30) a. **Do* you touch me!
b. *D'you *touch* me!
c. **Don't* you touch me!
d. Don't you *touch* me!

It just is not true that we have available detailed transformational (or any other) descriptions of English that will permit us to demonstrate that we can account for facts like these without special ad hoc rules mentioning IMP or NEG in ways that if abused could vitiate the cyclic principle.

As in so many other domains, generative grammarians have neither a substantial enough record of descriptive accomplishments to provide a well-understood testbed for theoretical hypotheses nor rigorous enough theories to make testable predictions in factual domains.

## 9 The grain of truth: what the cycle says about languages

The history of the cyclic principle in the literature of generative linguistics is somewhat bizarre, and that literature leaves unresolved the question of whether it has been reliably established that any such principle governs the grammars of human languages. Nonetheless, it seems to me that there is a real grain of truth about the cyclic principle. I want to try and encapsulate it in a way that the linguistic literature has not. It has to do with a claim that complex structures in language are assembled from well-formed parts which may be modified in the process of being concatenated and embedded but retain much of their own structural integrity.

Consider, in terms as near as we can get to being pretheoretic, what the cyclic principle actually guarantees. The effects of a rule, principle, or constraint in syntax may be either obligatory or optional in some clause. Let us look first at obligatory effects.

If some rule, principle, or constraint has an obligatory effect in any clause, then what the cyclic principle guarantees is that it will have that effect in a subordinate clause too, regardless of changes consequent on the embedding of

that clause under other material. To put this another way, what the cyclic principle excludes is embedding a constituent in a form that lacks some property that would and could have been true of it if it were not embedded.

Of course, embedding can require specific changes; for example, the embedding of a complement clause under a certain verb might call for the verb of the complement to be in the subjunctive mood, a morphological shape that is not used in unembedded contexts. But the cyclic principle requires that at least the deafult will be that properties of clauses will also be properties of complement clauses.

So, for example, if an English clause has a meaning involving logical argument-sharing (stipulated coreference) between its subject and direct object, the direct object must not have the form of an ordinary personal pronoun:

(31) a. us seeing someone
b. us seeing ourselves
c. *us seeing us

So when a clause with this semantic property is embedded under the verb *HOPE*, which imposes requirements like base form of the verb and presence of *TO*, and allows lack of complementizer and complement subject, the cyclic principle demands that the same be true:

(32) a. We hoped to see someone.
b. We hoped to see ourselves.
c. *We hoped to see us.

Countercyclic application of rules (which here would mean removal of the complement subject prior to checking binding constraints) would make complement clauses under *HOPE* less similar to other clauses than they could have been: it would be as if we had taken the ill-formed (31d) and combined it with *HOPE* (making just the required verb-form changes) and the result turned out well-formed. Random noncyclic application would have the same effect: (32c) would be predicted grammatical.

Now consider optional effects in embedded domains. An optional rule effect yields alternative possibilities of form. What the cyclic principle guarantees is that those extra possibilities will be seen to the extent possible in subordinate clauses as well. Take passivization as an optional rule for the sake of argument. Passivization allows (33b) alongside (34a) as an English clause.

(33) a. People saw us.
b. We were seen by people.

What the cyclic principle guarantees is that the two options will survive in embedded domains. When we embed these same clauses under a subject-to-subject raising verb like *tend* we will have the following:



(34) a. People tended to see us.
b. We tended to be seen by people.

Countercyclic application in this case would mean raising the embedded subject position without allowing for the passivization option in the embedded clause, so (34a) would be defined as well-formed but (34b) would not. An option available in a matrix clause would have been closed off, rendered unavailable in one type of subordinate clause. And random noncyclic application would have the same effect as cyclic application in this case.

In sum, with obligatory rule effects in embedded domains, the cyclic principle maximizes the fit between embedded constituents and unembedded ones, and with optional rule effects in embedded domains the principle maximizes the extent of the availability of the alternatives of form that the rule allows for.

One direction we can take as we look for a way to express this more clearly is the direction taken by monostratal theories of grammar. Now consider what an attempt to institute countercyclic syntax would mean for a monostratal theory. In a monostratal theory, a phrase like *to see us* would be represented as semantically incompatible with a first person plural subject, and this would be true whether it was immediately combined with a subject or combined with other material to make a larger phrase like *hoped to see us*; the compatibility property is inherited, so that both **for us to see us* and **We hoped to see us* involve attempts to combine a nonreflexive first person singular VP with a first person singular subject, and both are bad for that reason.

Again, *to see us* is a VP (in the terms of Gazdar et al. 1985, henceforth GKPS, it is a VP[INF]), and *to be seen by people* is another, so given a class of verbs like *tend* which are identified as compatible with a verb node that has a VP[PAS] as sister, both *tended to see us* and *tended to be seen by people* are predicted to be grammatical; there is no way to distinguish between them when we are considering what complement is compatible with the verb *tended*. (Note that although *seen by people* is a VP[PAS] under the GKPS account, *be seen by people* is not, and nor is *to be seen by people*.)

The effects of the cyclic principle essentially fall out from modes of combination of complements with heads that Pollard (1984) describes as *generalized context-free*. Variety available in independent clauses survives when they are employed as dependent clauses, and generalizations about their structure that hold when they are not embedded or concatenated continue to hold when they are embedded or concatenated.

Another way to put things is suggested by unpublished proposals of Arnold Zwicky concerning a view of grammar in which the function of a grammar is to define a set of construction types. When a rule defines a construction type, it refers to the type of the subconstituents that are to be combined; thus a rule for verb phrases might permit heads of the type to which *TEND* belongs to combine with phrases of the type VP[INF]. It may also call for

a certain rule to be employed in the definition of such a subconstituent; thus a rule for polar interrogative main clauses might permit a *wh*-marked NP to combine with an S[SLASH:NP] and call for the latter to be defined by the rule for auxiliary-initial S's (to get sentences like *What do you think?*). Constituents defined by rules as having certain given properties may thus be demanded in certain construction types or restricted to certain contexts, but there is no way in which an existing constituent can be *denied* its usual properties by being embedded in some other constituents, in the way that countercyclic application of transformations allows for.

The cyclic principle does not do the best job of expressing the insight that is being hinted at here; it insists on things that have no real consequences (like order of application of rules in mutually non-affecting domains), and it fails to entail in any rigorous way some things that we intuitively feel it should entail (recall the point made about markers like REFL and PASS above). Yet there is something there, an intuition to be captured — an intuition that monostratal frameworks in general come a lot closer to capturing than does classical TG (though as Paul Postal has pointed out to me, there are potentially ways of capturing it in relational frameworks too; it seems plausible that there could be a universal constraint saying roughly that the set of arcs in a subordinate constituent should satisfy the rules of the grammar, except insofar as specific determination of subordinate morphology by superordinate arc structure is involved).

## 10 Evolutionary explanations

A potential explanation for the generalization that is adumbrated in the cyclic principle stems from a very general version of the theory of evolution, a version general enough to apply outside of the domain of biology from which it emerged. The leading idea here is that set out (rather casually) by Simon (1962), further developed and clarified by Turney (1989), and applied to ethology by Dawkins (1976). It is a profound idea, with implications for philosophical conceptions of simplicity, the mathematical theory of error-tolerant communication systems, and the inductive inference problem in computer science. Encapsulated in a line, it is the idea that the only way to make a complex object (abstract or concrete) that exhibits stability in the face of disruptions and accidents is to give it a hierarchical structure.[2]

The implications hold for a wide variety of domains. Particularly important are those in biology: organisms that survive in an environment filled with natural hazards tend to have enough structural integrity that the loss of one tail segment or tooth or ear tip does not prevent them from getting on with their lives. But the principle that hierarchical structure is robust is extremely general, and holds in nonbiological domains too (Simon illustrates from political history, economic dynamics, bureacratic organizations, and molecular physics; Dawkins (1989) also contains many relevant and stimulating ideas).



The essential content of the cyclic principle has to do with the fact that evolutionarily viable structures are composed of substructures with their own coherence and stability. At this level of abstraction, the cyclic principle holds in domains beyond the natural and subconscious principles of sentence structure. I will consider just one simple example that illustrates the cyclic principle applying in a bureaucratic domain, namely the flow of control in the administrative handling of major promotions under the rules of the Academic Personnel Manual of the University of California, Santa Cruz.

The most basic material in an academic personnel file (evidence of teaching, research, service, etc.) is prepared by the faculty member whose career is under review and submitted to his or her department. The department (the next cycle) does not modify the materials the faculty member submits, but evaluates them, solicits confidential outside letters and adds them to the file, and adds a recommendation for action. The department does *not* return the file to the faculty member for review (strict cyclicity!).

The file is passed by the department chair to the dean of the division containing that department, and the dean adds his or her own recommendation regarding the proposed action. Again, the file is not returned countercyclically to a previous domain: the effects of the operations in those domains are incorporated, but they are not returned to.

The dean forwards the file to the Academic Personnel Office (APO) for the attention of the Committee on Academic Personnel (CAP), complete with all the contents so far amassed. CAP makes a judgment on the merits of the case after commissioning a report from an Ad Hoc Committee of experts to present an opinion but does not convey its decision downward (countercyclically) to the dean or the department or the faculty member: it presents its advice to the Academic Vice Chancellor (AVC). The AVC then makes a determination on the basis of what CAP recommended, and passes the file to the Chancellor for action (not down to any of the previous levels).

The full review thus proceeds in strict cyclic fashion in a hierarchical structure of successively less embedded domains. The conformity to the cyclic principle (and the Strict Cyclicity Condition) is regarded by the University as crucial to the integrity of the process. In one case where a Chancellor did take an action that was countercyclic (disagreeing with the composition of CAP's Ad Hoc Committee and returning to re-do that stage to his own satisfaction), a major constitutional upheaval resulted, and after a period of significant political crisis and recourse to the courts it was decided by the University that the case had been mishandled and the decision could not stand (the Chancellor's normally absolute decision to deny tenure was considered again in a new review, and was overturned).

Clearly, the personnel process of the University of California, which is the result of about a century of consultation, experience, and refinement, does not owe its cyclical character to a quirk of human brain biology. Nor does the cyclic principle or its analogs in the syntax of human languages, I claim. The cyclic

principle emerges from necessary features of the evolution of complex systems under conditions involving propagation, information transmission, random error, and systematic selective pressures — the same conditions that give rise to evolution in biological, physical, cultural, economic, political, and technological domains alike.

This does not mean that details of grammatical rules will be derivable directly from the theory of evolution. Theoretical linguistics is not going to be that easy; there are intrinsic complexities to linguistic systems that are not going to be reduced in a simplistic way to general laws. Human languages are extraordinarily complicated artifacts with histories going back tens of milennia, and the link to evolution should not be used as an excuse for a vulgar "it just evolves" to the intriguing question of how acquisition is possible.

But what Whitney (1875) called "the life and growth of language" really does involve the key ingredients of life: "the differential survival of replicating entities," as Dawkins (1989:192) encapsulates it. Utterance types are replicated (like Dawkins' "memes") each time one human being uses an utterance with the same syntactic structure as some previously used utterance. Modifications that no previous speaker introduced are made from time to time. Some utterance types survive and others fall out of use. This is not just analogous to evolution, it *is* evolution. It will accordingly be subject to the general laws of evolution.

There is a program of research here that deserves consideration, a program that would have us look not for random quirks of neural structure in *Homo sapiens* that bias our species (but not necessarily dolphins or RISC machines) toward the cyclic principle (or any principles of universal grammar), but rather for evolutionarily-based reasons for human languages being the way they are.

There are a very few antecedent examples of linguistic research of this type already in the literature. To cite one example, Hurford (1989) attempts to show by computer simulations that a communicative strategy that involves Saussurean signs (form-meaning correlations that are neutral vis à vis perception or production) win out over two imaginable alternative strategies. Through computer simulation experiments he arrives at a hint of an evolutionary explanation for the fact that human beings seem to employ static, bidirectional sound/meaning correspondences in their communicative behavior instead of just routines for engaging in communicative behavior that matches others' behavior, or communicative behavior that provides satisfactory responses.

This is a small beginning, but it points in the direction I am talking about. Hurford is looking not for arbitrary quirks of the genome or cerebral cortex but for ways in which the principles of structure we discern in human languages reflect elements of language design that would be natural and efficient (at least, more natural and efficient than their converses) in a language or other structured system fit for use by *any* highly-evolved species or well-engineered device, human or nonhuman, natural or artificial.



## NOTES

* I am grateful to Chris Barker, Sandra Chung, James Huang, James McCloskey, John Moore, David Perlmutter, Paul Postal, Jerry Sadock, Barbara Scholz, and Arnold Zwicky for discussions through which (whether they knew it at the time or not) they assisted me in the preparation of this paper. I also thank Cher Bergeon and Michael Johnston for research assistance and draft review. Preparation of the paper was partially supported by the Syntax Research Center at the University of California, Santa Cruz. None of the above-named should be assumed to be responsible for any mistakes I have made, or to agree with anything I say.

1. Oddly, it is not because of any doubts about successive cyclicity that mention of the cycle has fallen away in most of the TG literature. Hopping *wh*-movement is still widely assumed, though the cyclic principle that was alleged by Chomsky to mandate it is not. This is largely because of Freidin's (1978) attempt to show that various conditions and assumptions in then-current TG sufficed to eliminate the support for the Strict Cyclicity Condition by predicting everything it was supposed to account for (a forerunner of the kind of argument seen in Moore (1992)). I will not discuss Freidin's paper here because of space limitations; suffice it to say that I find some of his proposed replacements for the cyclic principle extremely vague and his conclusions highly uncertain, and the reader who studies his pp. 535–539 will find many indications that he was quite unsure of the validity of his claims (as Postal (1988:130) notes, p. 539 finds Freidin actually asserting that the Strict Cyclicity Condition follows as a "theorem" from "axioms" of Chomsky's theory, and then immediately states in a footnote that the theorem has not been proved and it "seems unlikely that a formal proof can be constructed"!).

2. The idea that the cyclic principle can be derived from considerations of the stability and evolutionary effectiveness of hierarchical structure has already been voiced by a linguist, Geoffrey Sampson (1978, 1980). Unfortunately, his presentation of the idea is so cursory and loosely phrased that it tends to bring it into disrepute. Sampson acknowledges the cyclic principle and then simply asserts, "Naturally that will be so," continuing with the following 72 words (Sampson 1980:182):

> Once the unit 'clause' has become established and has undergone some independent evolution, a new generation will learn to make the appropriate adjustments when putting lower-level constituents together to form clauses; and if they subsequently learn to form larger clauses by fitting together constituents some of which are themselves clauses, then the learners will use for this purpose clauses as they know them, not the unevolved clauses which their ancestors used.

This is the totality of his argument. It does not give even a well-disposed reader much to go on, and an ill-disposed reader will find it easy to dismiss. (In particular, surely those who agree that linguistic structure evolves do not have to believe that "their ancestors" used "unevolved clauses," any more than the clearly hierarchical structure of a dog implies that our ancestors kept tails or legs as pets.) Sampson's half-hearted and careless presentation tends to devalue what I think is an idea we should not dismiss.

BIBLIOGRAPHY

Akmajian, Adrian, & Frank Heny. 1975. An introduction to the principles of tranformational syntax. Cambridge MA: MIT Press.
Bach, Emmon. 1974a. Syntactic theory. New York NY: Holt, Rinehart & Winston.
Bach, Emmon. 1974b. Explanatory inadequacy. D. Cohen (ed.), Explaining linguistic phenomena, 153–72. New York NY: Hemisphere Press.
Bach, Emmon. 1975. On interrogative word movement in English. E. Polomé & A. Jazayery (eds.), Festchrift for Archibald A. Hill. The Hague: Peter de Ridder Press.
Bach, Emmon, & George M. Horn. 1976. Remarks on 'Conditions on transformations.' Ling. Inquiry 7.265–99.
Barss, Andrew. 1986. Chains and anaphoric dependence. PhD dissertation, MIT, Cambridge MA.
Baker, Carl L. 1978. Introduction to generative-transformational syntax. Englewood Cliffs NJ: Prentice-Hall.
Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge MA: MIT Press.
Chomsky, Noam. 1973. Conditions on transformations. S. R. Anderson and P. Kiparsky, eds., A Festschrift for Morris Halle. New York NY: Holt Rinehart & Winston.
Chomsky, Noam; Morris Halle; & Fred Lukoff. 1956. On accent and juncture in English. M. Halle et al. (eds.), For Roman Jakobson, 65–80. The Hague: Mouton.
Chomsky, Noam, & Morris Halle. 1968. The sound pattern of English. New York NY: Harper & Row.
Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. Linguistic Inquiry 8.425–504.
Chung, Sandra. 1982. Unbounded dependencies in Chamorro grammar. Ling. Inquiry 13.39–77.
Chung, Sandra. 1992. *Wh*-agreement and 'referentiality' in Chamorro. Paper presented at the Conference on Specifiers, University of California, Santa Cruz.
Chung, Sandra, and Carol Georgopoulos. 1988. Agreement with gaps in Chamorro and Palauan. M. Barlow & C. Ferguson (eds.), Agreement in natural language: approaches, theories, descriptions, 251–267. Stanford CA: CSLI.
Cinque, Guglielmo. 1990. Types of Ā-dependencies. Cambridge MA: MIT Press.
Dawkins, Richard. 1976. Hierarchical organization: a candidate principle for ethology. P. P. G. Bateson & R. Hinde (eds.), Growing points in ethology. Cambridge: Cambridge University Press.
Dawkins, Richard. 1989. The selfish gene. (New edition.) Oxford: Oxford University Press.
Epée, Roger. 1976. On some rules that are not successive cyclic in Duala. Ling. Inquiry 7.193–8.
Fillmore, Charles W. 1963. The position of embedding transformations in a grammar. Word 19.208–231.
Fodor, Jerry A. 1970. Three reasons for not deriving "kill" from "cause to die". Ling. Inquiry 1.429–38.
Freidin, Robert. 1978. Cyclicity and the theory of grammar. Ling. Inquiry 9.519–549.
Gazdar, Gerald; Ewan Klein; Geoffrey K. Pullum; & Ivan A. Sag. 1985. Generalized phrase structure grammar. Cambridge MA: Harvard University Press.
Georgopoulos, Carol. 1985. Variables in Palauan syntax. Nat. Lang. & Ling. Theory 3.59–94.



Goldberg, Jeff. 1985. Lexical operations and unbounded dependencies. CLS 21, Part 1: Papers from the general session at the 21st regional meeting, 122–132. Chicago IL: Chicago Linguistic Society.
Grinder, John T. 1972. On the cycle in syntax. John Kimball (ed.), Syntax & semantics 1, 81–111. New York NY: Seminar Press.
Grinder, John T., & Suzette Haden Elgin. 1973. Guide to transformational grammar. New York NY: Holt, Rinehart & Winston.
Huang, C.-T. James. 1993. Reconstruction and the structure of VP: some theoretical consequences. Ling. Inquiry 24.1 (in press).
Hukari, Thomas E. 1989. The domain of reflexivization in English. Linguistics 27.207–244.
Hurford, James R. 1989. Biological evolution of the Saussurean sign as a component of the language acquisition device. Lingua 77.187–222.
Hurford, James R. 1991. The evolution of the critical period for language acquisition. Cognition 40.159–201.
Jackendoff, Ray S. 1972. Semantic interpretation in generative grammar. Cambridge MA: MIT Press.
Jacobs, Roderick A., & Peter S. Rosenbaum. 1968. English transformational grammar. Waltham MA: Blaisdell.
Jacobson, Pauline, and Paul Neubauer. 1974. Extraposition rules and the cycle. C. Fillmore et al. (eds.), Berkeley studies in syntax and semantics, volume 1. Berkeley CA: Department of Linguistics.
Jacobson, Pauline, and Paul Neubauer. 1976. Rule cyclicity: evidence from the Intervention Constraint. Ling. Inquiry 7.429–461.
Keyser, Samuel Jay, & Paul M. Postal. 1976. Beginning English grammar. New York NY: Harper & Row.
Kimball, John. 1967. Predicates definable over transformational derivations by intersection with regular languages. Information & Control 11.177–195.
Kimball, John. 1972. Cyclic and linear grammars. John Kimball (ed.), Syntax & semantics 1, 63–80. New York NY: Seminar Press.
Lakoff, George. 1968. Deep surface grammar. Bloomington IN: Indiana University Linguistics Club.
Langacker, Ronald W., and Pamela Munro. 1975. Passives and their meaning. Language 51.789–830.
Lasnik, Howard, and Mamoru Saito. 1992. On the subject of infinitives. Papers from the 27th regional meeting. Chicago: Chicago Linguistic Society (in press).
McCawley, James D. 1984. Exploitation of the cyclic principle as a research strategy in syntax. W. de Geest & Y. Putseys (eds.), Sentential complementation, 165–183. Dordrecht, Holland: Foris.
McCawley, James D. 1988. The syntactic phenomena of English. 2 volumes. Chicago IL: University of Chicago Press.
McCawley, James D. 1992. The cyclic principle as a source of explanation in syntax. In this volume.
Moore, John. 1992. Antecedent government and the strict cycle. In this volume.
Perlmutter, David M. 1971. Deep and surface structure constraints in syntax. New York NY: Holt, Rinehart & Winston.
Perlmutter, David M. 1972. Evidence for shadow pronouns in French relativization. P.

M. Peranteau et al. (eds.), The Chicago which hunt: papers from the relative clause festival, 73–105. Chicago IL: Chicago Linguistic Society.

Perlmutter, David M., and Scott Soames. 1979. Syntactic argumentation and the structure of English. Berkeley CA: University of California Press.

Pollard, Carl J. 1984. Generalized phrase structure grammars, head grammars, and natural languages. PhD dissertation, Stanford University, Stanford CA.

Pollard, Carl J., and Ivan A. Sag. 1983. Reflexives and reciprocals in English: an alternative to the binding theory. M. Barlow et al. (eds.), Proceedings of the West Coast Conference on Formal Linguistics 2.189–283. Stanford CA: Stanford Linguistics Association.

Pollard, Carl J., and Ivan A. Sag. 1992. Anaphors in English and the scope of binding theory. Ling. Inquiry 23.261–303.

Postal, Paul M. 1972. On some rules that are not successive cyclic. Ling. Inquiry 3.211–22.

Postal, Paul M. 1974. On raising: one rule of English grammar and its theoretical implications. Cambridge MA: MIT Press.

Postal, Paul M. 1988. Advances in linguistic rhetoric. Nat. Lang. & Ling. Theory 6.129–37.

Postal, Paul M., & Geoffrey K. Pullum. 1978. Traces and the description of English complementizer contraction. Ling. Inquiry 9.1–29.

Postal, Paul M., & Geoffrey K. Pullum. 1982. The contraction debate. Ling. Inquiry 13.122–38.

Pullum, Geoffrey K. 1977. Successive cyclicity and the theory of traces. Mimeo, University College London. Presented at the Amsterdam Colloquium on Bounded vs. Unbounded Transformations, Amsterdam, Holland.

Pullum, Geoffrey K. 1979. Rule interaction and the organization of a grammar. New York NY: Garland.

Radford, Andrew. 1981. Transformational syntax. Cambridge: Cambridge University Press.

Ross, John Robert. 1969. Guess who? Papers from the 5th regional meeting, 252–86. Chicago IL: Chicago Linguistic Society.

Ross, John Robert. 1972. Doubl-ing. Ling. Inquiry 3.61–86.

Sampson, Geoffrey. 1978. Linguistic universals as evidence for empiricism. Journal of Linguistics 14.183–206.

Sampson, Geoffrey. 1980. Making sense. Oxford: Oxford University Press.

Selkirk, Elisabeth O. 1972. The phrase phonology of English and French. PhD dissertation, MIT, Cambridge MA.

Simon, Herbert A. 1962. The architecture of complexity. Proceedings of the American Philosophical Society 106.467–482.

Thompson, Henry. 1975. The cycle: a formal statement. Papers from the 11th regional meeting, 589–603. Chicago IL: Chicago Linguistic Society.

Turney, Peter. 1989. The architecture of complexity: a new blueprint. Synthese 79.515–542.

Wasow, Thomas. 1979. Anaphora in gemerative grammar. Ghent, Belgium: Story-Scientia.

Whitney, William Dwight. 1875. The life and growth of language. London: Henry S. King.

Zaenen, Annie. 1983. On syntactic binding. Ling. Inquiry 14.469–504.