

THREE CASES OF OVERGENERATION*

Kenneth Hale
LaVerne Masayesva Jeanne
Paul Platero

*Department of Linguistics
Massachusetts Institute of Technology
Cambridge, Massachusetts*

1. Introduction

It is possible to imagine two diametrically opposed positions concerning the fundamental nature of language. We will characterize these positions in rather extreme terms, though we would not wish to claim that any linguist necessarily adheres totally to either extreme. Our characterizations are to be understood as convenient abstractions that are, to some extent at least, consistent with observable extremes in actual linguistic work.

1.1 The Autonomous Systems View¹

According to this view, language consists of a number of distinct systems, each possessing inherent principles of organization that are essentially independent of factors relating to any other linguistic system or to extralinguistic considerations. The systems alluded to here include, for example, the categorial rules of the base, the lexicon, the transformational component, the system of rules that assign obligatory coreference, or construal, and other systems. To say that the systems are independent, or autonomous, is to say, for one thing, that the rules belonging to each system apply without constraint—with the exception, of course, of those constraints and conditions that are general for grammatical rules of all kinds and whose identification is one of the important tasks of linguistic science (e.g., general constraints of the sort identified in Chomsky, 1964, 1973, and in Ross, 1967). What the Autonomous Systems thesis specifically rules out, we would suggest, is the local, rule-particular, constraint of the type often seen appended to transformational rules (compare the No-Condition Principle of Perlmutter, 1971, p. 128), or the sort of constraint represented by the environmental specification in a context-sensitive phrase structure rule. The relatively early abandonment by generative

grammarians of context-sensitive phrase structure rules can, we believe, be seen as belonging intellectually, though not historiographically, to the tradition that most clearly approximates the Autonomous Systems position. And the thesis of the autonomy of syntax, which informs much of the work of Noam Chomsky and his students, is certainly within this tradition.

1.2 The Dependent Systems View

According to a conceivable opposite extreme, a language consists of a single unified system—or else a set of tightly integrated systems—whose inherent principles of organization are often intimately related to factors belonging to conceptually distinct realms, including extralinguistic factors. Under this conception of language, it is very much to be expected that a **transformational** rule might, for example, be **lexically** governed—i.e., restricted in its application to particular lexical items (e.g., Lakoff, 1970c)—or be subject to one or another of a variety of semantic, pragmatic, or even cultural constraints (e.g., Lakoff, 1971a), or that the internal structure of a phrasal category, say NP, might be dependent upon the grammatical relation it bears to the verb, or upon some semantic factor such as, for example, its use in a generic, as opposed to a specific, sense; and so forth.

1.3 Overgeneration

As preliminary illustration of the distinction between the Autonomous Systems view and the Dependent Systems view, we wish to offer what our imagination tells us would be the treatment of a certain Navajo fact under the two conceptions of language.²

Navajo possesses a rule whose effect is roughly that of the passive in English. The rule relates sentences of the form represented in (1) below to corresponding “inverted” sentences of the form represented in (2):

- (1) *Ashkii at'ééd yizts'qs.*
 (boy girl yi-kissed)
 ‘The boy kissed the girl.’
- (2) *At'ééd ashkii bizts'qs.*
 (girl boy bi-kissed)
 ‘The girl was kissed by the boy.’

In addition to inverting the linear order of the subject and object noun phrases, the rule also involves a change in the object marking prefix in the verb work—/yi-/ is replaced by /bi-/; as can be seen in the example. The rule is further exemplified by (3)–(4) below:

- (3) *Łíí' dzaanééz yiztał.*
 ‘The horse kicked the mule.’

- (4) *Dzaanééz lǫǫ' biztał.*
 'The mule was kicked by the horse.'

Thus, the effect of the rule is to convert a sentence of the form:

S(subject) O(bject) *yi-V(erb)*

into a new sentence of the form:

O(bject) S(subject) *bi-V(erb).*

We are evidently justified in relating the two forms, since they are cognitively synonymous and, moreover, exhibit the same selectional relationships between their verbs and noun phrases. Let us assume, for the sake of this discussion, that we are correct in relating the two by means of a transformational rule that inverts the subject and object noun phrases—there is, in fact, evidence to support this conclusion, though it will not be detailed here—and let us refer to the rule as Subject-Object Inversion, or SOI.

The interesting fact about this process in Navajo is that while cognitively synonymous pairs

S O *yi-V*
 O S *bi-V*

exist to describe many events appropriately described by transitive sentences in Navajo, there are many cases in which only one or the other, but not both, of the two forms can be used. Thus, for example, the circumstance in which a boy kicks a stone can be described in Navajo with a sentence of the form S O *yi-V*, but not with a sentence of the form O S *bi-V*:

- (5) *Ashkii tsé yiztał.*
 'The boy kicked the stone.'
- (6) **Tsé ashkii biztał.*
 'The stone was kicked by the boy.'

On the other hand, the circumstance in which a bee stings a boy may be described in Navajo with a sentence of the form O S *bi-V*, but not with a sentence of the form S O *yi-V*.

- (7) **Tsís 'ná ashkii yishish.*
 'The bee stung the boy.'
- (8) *Ashkii tsís 'ná bishish.*
 'The boy was stung by the bee.'

The principle is roughly this. Nominal concepts are ranked in Navajo with reasoning beings (humans) at the top of the hierarchy, and unreasoning entities (inanimates and abstract concepts) at the bottom. For sentences in which the subject and object are unequal in rank, the acceptable surface form is that in which the higher ranking NP precedes the lower ranking one.

How would this phenomenon be handled under the two conceptions of language contrasted above? Let us first imagine how it would be handled in the Dependent

Systems view, since this is the way we originally conceived of the problem when it first came to our attention. In this conception, the rule of SOI is subject to the following conditions of application:

- (a) Optional if the subject and object are equal in rank;
- (b) Obligatory if the object outranks the subject;
- (c) Inapplicable if the subject outranks the object.

The Autonomous Systems treatment of this Navajo process, we would propose, is roughly as follows. Assuming that SOI is in fact a bonafide transformational rule in Navajo, it applies **without constraint**. Thus, from the perspective of the transformational system of Navajo syntax, **all** of the sentences (1)-(8) are equally, and perfectly, well formed. This is not to say, of course, that they are all of the same status from a global perspective; far from it. The deviance of sentences (6)-(7), however, does not pertain to the realm of syntax. Rather, it has to do with the **meaning** of the inverted form of the sentences in Navajo. Witherspoon (forthcoming) has argued, persuasively we feel, that the inverted form is used when a being denoted by the object noun phrase, either on purpose or inadvertently, **relinquishes control** of the situation described by the sentence. Thus, the Navajo version of (2) is very poorly represented by the English passive—a more apt rendition would be something like 'The girl let the boy kiss her.' Witherspoon maintains, correctly, we feel, that the ill-formedness of (6) has nothing to do with syntax but rather with the impossibility, from the Navajo point of view, of a stone relinquishing control—a stone does not reason, so it cannot be in control of the situation to begin with. The inappropriateness of (6) can be appreciated, Witherspoon points out, when the sentence is translated 'The stone let the boy kick it.' Correspondingly, the inverted form (8) is required because, in order for a boy (human, and therefore reasoning) to be stung by a bee, he must relinquish his inherently greater control.

Whether Witherspoon is right or wrong about the Navajo semantics involved here is beside the point, actually. Whatever the facts, his assertion that it is not a matter of syntax, or even of grammar, would be represented in the Autonomous Systems view by permitting the syntactic rule of SOI to operate entirely without constraint.

The present paper proceeds from a view of language which, it seems to us, is most consistent with the Autonomous Systems thesis. In the ensuing discussion, we present three cases—one each in Papago, Hopi, and Navajo—in which rules of grammar, if allowed to apply without constraint—that is to say, in conformity with the Autonomous Systems thesis—produce sentences that are in one sense or another unacceptable. For each case, we will propose that the principle which accounts for the unacceptability properly belongs to a system which is entirely separate from that to which the rule itself belongs.

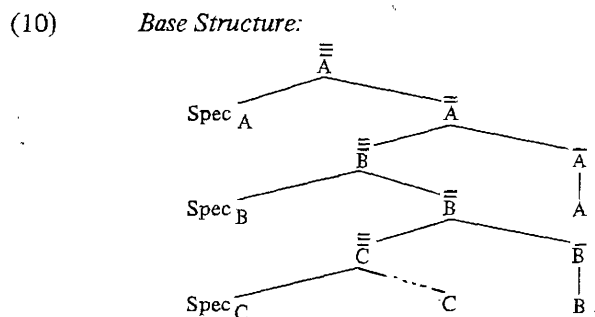
2. Papago Extraposition³

Each of the cyclic phrasal categories of Papago—i.e., sentence ($\bar{\bar{V}}$), noun phrase ($\bar{\bar{N}}$), and postpositional phrase ($\bar{\bar{P}}$)—is expanded by means of phrase structure rules conforming to the following schemata:

- (9) *Base Rules:*
- $\bar{\bar{A}} \rightarrow \text{Spec } \bar{\bar{A}}$
 - $\bar{\bar{A}} \rightarrow (\bar{\bar{B}}) * \bar{\bar{A}}$
 - $\bar{\bar{A}} \rightarrow (M) A$

These rules are to be understood as comprising the X-bar theory of Papago phrase structure—the symbols A, B, C, . . . , are used in place of X for expository reasons. The rules embody the claim that Papago is basically nucleus-final. That is to say, the sentence is basically verb-final, the noun phrase is basically noun-final, and the postpositional phrase is basically just that, i.e., postpositional rather than prepositional. The initial expansion of each category introduces the appropriate specifier—“auxiliary” for the sentence, determiner for the noun phrase and spatial determiner (cp. English “there in the house”) for the postpositional phrase. The second expansion introduces the complement(s)—e.g., arguments of the verb (subject, object, etc.) for the sentence, possessor expression (optional) for the noun phrase, and the object of the postposition for the postpositional phrase. The final expansion introduces the head (V, N, P) and an optional pre-head modifier (M, omitted from ensuing examples, since it plays no role in the discussion).

The tree diagram offered in (10) illustrates the underlying structure-type produced by the proposed Papago base rules.



However, this nucleus-final arrangement is not the only one observed in Papago surface structures. Among the actual surface structures underlain by (10) are those included in (11) below.

- (11) *Well-formed Surface Structures:*
- $\text{Spec}_A (\text{Spec}_B) (\text{Spec}_C) C B A$
 - $\text{Spec}_A A \text{Spec}_B (\text{Spec}_C) C B$
 - $\text{Spec}_A A \text{Spec}_B B \text{Spec}_C C$

The first of these is the nucleus-final arrangement produced by the rules of the base. The parenthesis around Spec_B and Spec_C are included to reflect the fact that the generalized N-category specifier /g/ (/hég/ in some Pima dialects and under certain conditions in Papago) is deleted when immediately preceded by a N- or P-category specifier belonging to an immediately superordinate phrase—the rule that effects this deletion may be formulated roughly as follows:

$$(12) \quad \begin{array}{l} \textit{g-Deletion:} \\ \bar{X} - \text{Spec}_{N,P} - g - X \\ \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 1 & 2 & 0 & 4 \end{array} \end{array}$$

Clearly, this rule applies only to nucleus-final structures, since only those structures provide the necessary conditions. In other words, *g-Deletion* must apply after any process that effects non-nucleus-final structures.

The second arrangement in (11) above can be assumed to be produced by extraposing the \bar{B} -phrase—on the \bar{A} -cycle. Since \bar{B} itself contains a three-bar structure, \bar{C} , the conditions for *g-Deletion* are met within the extraposed phrase—hence the parentheses around Spec_C .

Finally, the arrangement observed in (11c) above may be assumed to be produced by first extraposing \bar{C} on the \bar{B} -cycle, and then extraposing \bar{B} on the \bar{A} -cycle. This “maximal” application of the putative extraposition process results in the circumstance that no specifier appears in a position from which it would be deleted by (12)—hence, all of the original specifiers are present in the surface structure (11c).

We now formulate the proposed extraposition rule involved in the derivation of (11b-c). We assume that it is cyclic, that it is optional, and that it applies at the first opportunity—i.e., at the two-bar level of structure—to move a three-bar structure into postnuclear position.

$$(13) \quad \begin{array}{l} \textit{Three-bar Extraposition:} \\ X - \bar{A} - X \\ \begin{array}{cccc} 1 & 2 & 3 & \\ 1 & 0 & 3 & 2 \end{array} \end{array}$$

This rule applies in identical fashion to all phrasal categories of Papago. Thus, given a sentence (\bar{V}) of the form:

$$(14) \quad \bar{V} [no \bar{N} [g \bar{N} [g \textit{húsi}] \bar{N} \textit{kíi}] \bar{N} g\acute{e}'ej] \bar{V}$$

‘Is Joe’s house large?’

the following well-formed surface structures are observed (with *g-Deletion* applied where necessary):

- $$(15) \quad \begin{array}{l} \text{i. } no \ g \ \textit{húsi} \ \textit{kíi} \ g\acute{e}'ej. \\ \text{ii. } no \ g\acute{e}'ej \ g \ \textit{húsi} \ \textit{kíi}. \\ \text{iii. } no \ g\acute{e}'ej \ g \ \textit{kíij} \ g \ \textit{húsi}. \end{array}$$

(The suffix /-j/ appearing on /kii/ in (15iii) is a third person singular agreement marker that regularly appears on the head N when a possessor NP is moved to its right. We leave unformulated the mechanism that accounts for this, since it is not relevant here, insofar as we are aware.)

Similarly, given a noun phrase (\bar{N}) of the form:

- (16) $\dots \bar{N} [g \bar{N} [g \bar{N} [g \text{ h́usi}]_{\bar{N}} \text{ 'oog}]_{\bar{N}} \text{ kii}]_{\bar{N}}$
 'Joe's father's house'

the following well-formed surface structures are observed (again, with *g*-Deletion applied where necessary):

- (17) i. $\dots g \text{ h́usi 'oog kii}$
 ii. $\dots g \text{ kii} g \text{ h́usi 'oog}$
 iii. $\dots g \text{ kii} g \text{ 'oogaj g h́usi}$.

And finally, given a postpositional phrase (\bar{P}) of the form

- (18) $\bar{P} [\text{'am } \bar{N} [g \bar{N} [g \text{ h́usi}]_{\bar{N}} \text{ kii}]_{\bar{N}} \text{ wúi}]_{\bar{P}}$
 'to Joe's house'

the following well-formed surface structures are observed (with *g*-Deletion, as usual):

- (19) i. $\dots \text{'am h́usi kii wúi}$
 ii. $\dots \text{'am wúi g h́usi kii}$
 iii. $\dots \text{'am wúi g kii} g \text{ h́usi}$.

We have assumed that Three-Bar Extraposition applies in cyclic fashion to produce (11c) and the corresponding actual Papago surface structures—that is, we have assumed that it applies first on the innermost cycle in producing those forms. If we are correct in this assumption, then, since the rule is optional and may therefore be skipped on any higher cycle, there is a further surface form the rule will produce—i.e., that resulting from the application of the rule **only** on the inner cycle. The resulting surface structure in this case, however, is ill formed.

- (20) *An Ill-formed Surface Structure*
 *Spec_A (Spec_B) B Spec_C C A
 *no *g kii} g h́usi g'e'ej* (from (14))
 *... *g 'oogaj g h́usi kii* (from (16))
 *... *'am kii} g h́usi wúi* (from (18)).

There is another consequence of our assumption also. If the rule is indeed cyclic, then there is no way in which a constraint can be placed upon the rule itself in order to avoid the ill-formed surface structure in (20)—apart, that is, from appending a condition to the rule to the effect that it applies obligatorily if the factor \bar{A} is of the form $\bar{A} [\text{Spec}_A \text{ A } \bar{B}]_{\bar{A}}$, which amounts to saying that the rule is obligatory if the category affected by it has itself undergone the rule. Such a condition would be entirely counter to the spirit of the Autonomous Systems view of language. It is, moreover, an extremely unusual and potentially overpowerful sort of condition; in

fact, it is the very sort whose riddance inclines us toward the Autonomous Systems position. The reader may well wonder at this point why extraposition is not simply postcyclic, thereby precluding (20). The answer to this question will become evident later.

We would like to propose that the surface structure (20) is, from the point of view of the Papago system of rules to which extraposition belongs, perfectly well formed and that its unacceptability from a more global perspective is to be explained in terms of another system entirely—namely, the system that assigns intonation to linguistic forms in Papago.

In preparation for the description of the mechanisms that assign intonation contours in Papago, it is convenient first to discuss the assignment of stress.

With marginal exceptions, the assignment of primary stress in Papago words is entirely regular—the first vowel of the stem receives primary stress. Thus:

/ʔóʔodham/	‘person’
/gátwid/	‘to shoot (imperfective)’
/šópolk/	‘short’

Prefixes and suffixes do not receive primary stress:

/ha-jéweđ-ga/	(their-land-alienable) ‘their land’
---------------	-------------------------------------

The independent pronouns and the demonstrative determiners receive primary stress on the first vowel, e.g.:

/ʔáacim/	‘we’
/hégam/	‘those’

And certain particles also receive primary stress on the first vowel, e.g.:

/hémho (a)/	‘necessarily’
-------------	---------------

But a large number of particles are unstressed:

/o/	‘future’
/cem/	‘unachieved intention, past’
/wuđ/	‘equational’

The elements commonly referred to by the term “auxiliary” are also unstressed:

/ʔañ/	‘imperfective first person singular’
/no/	‘imperfective third person singular interrogative’
/nap/	‘imperfective second person singular interrogative’

and so forth. Certain unstressed short, as opposed to extrashort, vowels have associated with them a readily perceptible tertiary stress (also predictable), but this plays no role in the intonational phenomena with which we are concerned here. We will refer to all vowels not assigned primary stress as “unstressed” in the context of this discussion. The role of primary stress in the phonology of Papago—particularly in relation to vowel reduction—is described briefly in Hale (1965).

Certain Papago lexical items receive two primary stresses, the first somewhat subordinated to the second. These items include compounds—both partially fossilized

ones like /hóasá'a/ (basket + s + earthen bottle) 'plate', and thoroughly perspicuous ones like /páal-wákon/ (priest-wash) 'to baptize (imperfective)', /siil-mó'ō/ (saddle-head) 'saddle horn', and /cúkuḍ-šóša/ (owl-snot) 'date (i.e., the fruit)'—in which each member receives a primary stress on its first vowel. In addition, there are loans from Spanish—like /pápalóodi/ (< *papalote*) 'windmill, kite', and /'ískóobli/ (< *escoblo*) 'chisel'—that receive primary stress both on the first vowel, in accordance with the Papago rule, and on a vowel later in the word (to wit, the vowel that bore the primary stress on the form as it was spoken in the source language). It should also be pointed out that the process of compounding, as well as other productive processes, give rise to lexical items with more than two stresses—in fact, there is no theoretical upper limit:

/cúkuḍ-šóša-kóstal/ ((owl-snot)-bag) 'bag for dates'
 /tóki-béhēdam-máagǎna/ ((cotton-picker)-machine) 'mechanical cotton picker'
 /tóki-béhēdam-máagǎna-mélcuddam/ ((cotton-picker)-machine)-driver
 'operator of a mechanical cotton picker'

Impressionistically, at least, the final stress is more prominent than those preceding it, although the details of stress subordination in Papago have not been worked out. For the purposes of this discussion, it is sufficient merely to speak of "stressed" vowels (those to which the diacritic is assigned in the above examples) as opposed to "unstressed" vowels (those not supplied with the acute accent diacritic).

The assignment of intonation contours is, for the most part, entirely mechanical. Considering first the intonation of stress-bearing lexical items spoken in isolation, the facts are extremely straightforward. In words with a single stressed vowel, the pitch on that vowel is relatively high (H), while the pitch on unstressed vowels is relatively low (L). (We will assume that extrashort vowels and the epenthetic schwa-vowel, when voiced, are assigned a pitch value, like other vowels.) Thus:

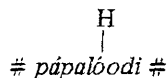
/ó'odham/	(HLL)
/gát[ə]wid/	(HLL)
/šópolk/	(HL)
/ha-jéwed[ə]-ga/	(LHLLL)
/áacim/	(HLL)
/hégam/	(HL)
/hémo (a)/	(HL (L))

Following Goldsmith (1976) and Haraguchi (1975), we will assume that there exists a basic pitch pattern, independent of particular segmental sequences, and that this pattern is mapped in a straightforward fashion onto a segmental string by associating Hs and Ls with the syllabic segments. For Papago, we will assume that the basic pitch pattern is HL and that the domain to which it is assigned is any string that intervenes between, but does not itself contain, #-boundaries—where the symbol # is used here to designate, not word boundary (as is traditional for this symbol), but rather the class of boundaries that includes those defining the left and right margins of a sentence, or of a word or phrase spoken in isolation.

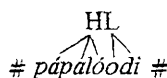
The procedure that associates the basic HL pitch pattern to segmental sequences may be expressed in the following rule:

- (21) *Intonation Assignment*
 Given the basic pitch pattern HL, associate the H with the rightmost stressed vowel.

We then assume there to exist a convention whereby H is also associated with each syllabic segment preceding the rightmost stressed vowel, if there are any, i.e., the H is, so to speak, "spread" to the left, and L is associated with each syllabic segment that follows the rightmost stressed vowel, assuming any do, i.e., L is "spread" to the right. Accordingly, the intonation associated with the word /pápalóodi/ 'windmill, kite', spoken in isolation, say, is assigned in the following manner. First, H is associated with the segment /ó/:



Then, the H is further associated with each syllabic to the left of the pivotal position defined by (21), and L is associated with each vowel to the right:

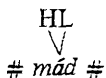


This gives an intonation which we might symbolize HHHLL, that is to say, high from the beginning of the word up to and including the rightmost stressed vowel, falling to low on the subsequent vowels. Precisely the same intonation assignment procedure operates in the case of compound lexical items, regardless of their internal complexity. Thus, for example,



which exhibits the intonation HHHHHHLLL.

If the rightmost stressed vowel is also the last syllabic (capable of bearing a perceptible pitch—voiceless vowels and the voiceless epenthetic schwa-vowel may, perhaps, not be capable of that), as in a word like /mád/ 'offspring of female', the H is associated with the stressed vowel, as usual, but the L is also associated with it. This produces a falling intonation (which we might symbolize HL), albeit one sometimes difficult to perceive:

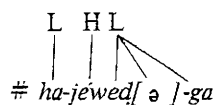


We have one final adjustment to make in order to complete the picture for intonation assignment within the domain # . . . #. If the leftmost stressed vowel in the domain is not also the leftmost syllabic, then any H that precedes it as a result of the leftward spreading described above must be lowered (whether to L or to M[id])

we cannot tell, precisely, but we will assume the former). This adjustment is accomplished by the following rule:

- (22) *Lowering*
 $H \rightarrow L / _ XVQ \#$

This produces the LHL pattern associated, for example, with prefix-initial words spoken in isolation:



The examples offered to this point illustrate cases in which the domain of intonation assignment corresponds to lexical items spoken in isolation. In such a case, the pattern (L)HL is assigned to the item as a whole, regardless of its length or internal complexity. Precisely the same is true of larger constructions (with certain exceptions, one of which will be detailed below).⁵ That is to say, the pattern (L)HL is assigned to entire constructions, not to the individual words, or even phrases, of which they are composed. Thus, for example, while the noun /húsi/ 'Joe', spoken in isolation (and therefore constituting the entire domain of intonation assignment), would receive the intonation HL, it does not receive this intonation when it appears as the possessor in the possessor-possessed construction /húsi 'óog/ 'Joe's father'. Considered in isolation (as will be the case for all construction types in the immediately ensuing discussion), the larger possessive construction would, as a whole, be assigned an intonation in accordance with rule (21) and the associated conventions.

/húsi 'óog/ (HHHL)

(The absence of the N-category specifier here is due to another *g*-Deletion process that removes an utterance-initial /g/.) And, if this were itself the possessor phrase in a larger possessive construction, e.g., /húsi 'óog kii/ 'Joe's father's house', it would not be assigned its own intonation; rather, the larger construction, as a whole, would constitute the domain of intonation assignment.

/húsi 'óog kii/ (HHHHHL)

And if this construction, in turn, appeared within a postpositional phrase, e.g., /'am húsi 'óog kii wú/ 'to Joe's father's house', the intonation would, again, be assigned to the more inclusive construction.

/'am húsi 'óog kii wú/ (LHHHHHHHL)

Finally, if this postpositional phrase appeared as a preverbal complement within a sentence, e.g., /namt 'am húsi 'óog kii wú o híhi/ 'Are you (pl.) going to go to Joe's father's house?', its own intonation would yield to that of the sentence as a whole.

/namt 'am húsi 'óog kii wú o híhi/ (LLHHHHHHHHHL).

In general, therefore, the basic intonation pattern of Papago is assigned to constructions. In any given case, the construction may consist of a single word—as is necessarily the case when a word is spoken in isolation—or it may consist of an entire sentence.

Our discussion so far implies that all Papago constructions bounded by pauses are intoned with an uninterrupted (L)HL pitch profile. This is not the case, however. While the basic pitch pattern is always assigned to Papago constructions, true enough, it is not always the case that it is assigned “umbrella fashion” to the most inclusive phrase. That is to say, there are circumstances under which the basic pattern is observed to repeat itself within a construction. One such circumstance, we would like to suggest, is intimately related to the central issue of this discussion—namely, the principle that accounts for the ill-formedness of the surface structures characterized in (20) above.

The reader will have noticed, no doubt, that all of the examples so far used to illustrate the assignment of intonation in Papago share the property that they are nucleus-final. More to the point, none of them has, at any level of structure, undergone the rule of Three-Bar Extraposition, which would move a constituent into postnuclear position.

Let us now consider the intonation assigned to structures that have undergone the extraposition rule. First, observe the intonations assigned to (15b-c):

- (15) b. *no gé'ej[ə] g húsi kī.* (LHLL HHHL)
 c. *no gé'ej[ə] g kīij[ə] g húsi.* (LHLL HLL HL)

The intonation assigned to (17b-c) and to (19b-c) exhibit the same sort of pattern:

- (17) b. ... *g kīij[ə] g húsi 'oog* (... HLL HHHL)
 c. ... *g kīij[ə] g 'oogaj[ə] g húsi* (... HLL HLLL HL)
 (19) b. ... *'am[ə] wúu g húsi kī* (... LLHL HHHL)
 c. ... *'am[ə] wúu g kīij[ə] g húsi* (... LLHL HLL HL).

In each case it can be seen clearly that the constituent moved by extraposition comprises its own domain for the purposes of intonation assignment. Or, to put this another way, when extraposition moves a three-bar structure to the left of a nuclear element, the latter defines the right-hand margin of an intonational domain. We might express this in terms of a rule which states simply that the double cross boundary symbol # is inserted after any nuclear element which appears *medially* within an endocentric phrase of which it is the nucleus:

- (23) *Partitioning by Extraposition*
 $\frac{\equiv}{A} [\dots A \cdot \# X]_A^{\equiv}$, X is non null

This rule (convention, or whatever) has the effect of inserting a # -boundary into any two-bar structure in which extraposition has applied, thereby accounting for the intonational facts—specifically, the repeated (L)HL pattern—observed in the (b-(b-c) sentences of (15), (17), and (19).

We have now set the scene for the presentation of our proposal concerning the ill-formed surface structures represented in (20). In those cases, (23) will have the

effect of inserting a # -boundary into the innermost two-bar structure, thereby effecting the partitioning observed in (24) below:

- (24) *no [g kīj # g hūsi] g'e'ej (from (14))
 *... g ['ógaj # g hūsi] kīi (from (16))
 *... 'am [kīj # g hūsi] wūi (from (18)).

Notice that there is a property, related to the partitioning by #, which distinguishes these ill-formed surface structures from the well-formed ones. Specifically, the ill-formed structures have the property that a partitioned phrase precedes the nucleus of an immediately superordinate phrase. This is not the case in any of the well-formed surface structures. We would like to suggest that it is precisely this property that accounts for the ill-formedness of the structures (20). We propose, in short, that there is a surface structure constraint that identifies as ill formed any surface structure in which a partitioned phrase precedes the nucleus of a larger phrase of which it is an immediate constituent. The constraint might be expressed as follows:

- (25) *A Surface Structure Constraint*
 * \bar{A} [Spec \bar{A} \bar{B} [... # ...] \bar{B} A] \bar{A}

Thus, we claim that the surface structures of (20) are ill formed not by virtue of the application of Three-Bar Extraposition on an internal cycle, but rather by virtue of the relationship between phrase structure and the position of the # -boundary. It is almost certainly true that this is related to intonation, whose domain of assignment is defined by # -boundaries. Observe that in well-formed surface structures it is necessarily the case that the first (i.e., leftmost) nuclear element that has a falling intonation (i.e., H to L transition) associated with it is also the nucleus of the entire structure; that is to say, it is the nucleus of the maximally superordinate structure. This, as the reader may easily verify, is simply an automatic consequence of the way in which the well-formed surface structures arise, given the mechanisms and principles we are assuming. However, if the rules of intonation assignment applied to the structures that exhibit the illicit partitioning of (24), the first nucleus with falling intonation would not be the nucleus of the maximally superordinate structure. It does not seem unreasonable to suggest that this fact, i.e., this failure of correspondence between "main nucleus" and "first intonationally marked nucleus," is the essential source of the ill-formedness of the surface structures represented by (20) (equivalently, (24)) and, therefore, the primary motivation for the surface structure constraint (25).

Be this as it may, in order to support our contention that a surface structure constraint is the proper way to account for the ill-formedness of (20), (24), we must demonstrate that Three-Bar Extraposition in fact applies in cyclic fashion. This follows, since we must allow the rule to apply on an internal cycle in order to produce the ill-formed surface structures in the first place. If the rule were not cyclic—if it were, say, postcyclic—then presumably we could arrange matters in such a way as to prevent the generation of (20), (24), and the like.

To put this another way, in order to show that the surface structure constraint expressed in (25) is necessary, we must show that a structure of the type represented in (20), (24) is itself a necessary intermediate form in some derivation. To show that such a structure is, in fact, a necessary intermediate form, we must briefly describe another extraposition rule—actually a subvariety of the extraposition already studied and expressed in (13). This variant extracts a two-bar structure, rather than a three-bar structure. Accordingly, we will give it the name Two-Bar Extraposition. It may be expressed as follows:

$$(26) \quad \text{Two-Bar Extraposition}^6$$

$$\begin{array}{cccc} X & - & \bar{\bar{A}} & - & X \\ 1 & & 2 & & 3 \\ 1 & & 0 & & 3 & & 2 \end{array}$$

Rule (26) operates in all essential respects just like (13), except of course it leaves a specifier behind in the original prenuclear position. For example, from a structure of the form,

$$(27) \quad \bar{\bar{V}}[nap \bar{\bar{N}}[hég(ai) 'ó'odham]\bar{\bar{N}} s-máac]\bar{\bar{V}}$$

‘Do you (sg.) know that man?’

it produces the surface structure

$$(28) \quad nap \ hég \ s-máac \ 'ó'odham.^7$$

And from a structure of the form

$$(29) \quad \bar{\bar{V}}[no \bar{\bar{P}}[am \bar{\bar{N}}[g mīsa]\bar{\bar{N}} wéco]\bar{\bar{P}} wó'o]\bar{\bar{V}}$$

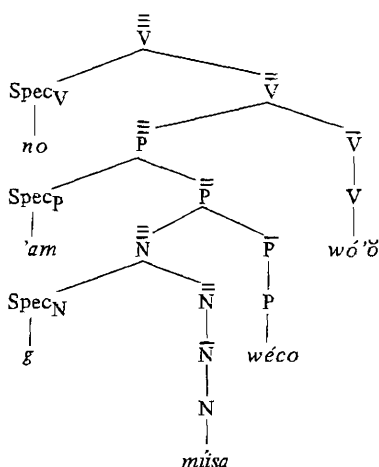
‘Is it lying under the table?’

it produces the following:

$$(30) \quad no \ 'am \ wó'ó \ mīsa \ wéco.$$

An important fact to notice about surface structure (30) is this. The specifier /g/, associated with the object of the postposition, is **missing** in the structure produced by Two-Bar Extraposition. This fact is an automatic consequence of the way in which the rule must apply—assuming that it is in fact a subvariety of Three-Bar Extraposition and, therefore, according to our claim, cyclic, and assuming that *g*-Deletion is also cyclic. To see this, consider the structure of (20), the source of (30), in a tree diagram:

(31)



The derivation of (30) involves extraposition of the two-bar structure \bar{P} . But, notice, this cannot take place until we have reached the \bar{V} -cycle. At that cycle, we have already passed the \bar{P} -cycle at which the condition for *g*-Deletion is met. We account naturally for the deletion of /g/ by assuming that it applied at the first opportunity—that is to say, at the \bar{P} -cycle and, therefore, prior to the extraposition of \bar{P} that must be delayed until the \bar{V} -cycle is reached.

Suppose, however, that we apply Three-Bar Extraposition on the \bar{P} -cycle, and then Two-Bar Extraposition on the \bar{V} -cycle. This is a theoretically possible derivation, given our assumptions. Consider, however, the intermediate structure produced by extraposition on the \bar{P} -cycle:

(32) $\bar{V} [no \bar{P} [am \ wéco \ \bar{N} [g \ mísa] \bar{N}] \bar{P} \ wó'ò] \bar{V}$

This is precisely the ill-formed structure with which we are concerned. Since the #-boundary will be inserted following the postposition /wéco/, which is the nuclear element in the postpositional phrase, the surface structure constraint (25) will come into play to identify (32) as ill formed. And it is indeed the case that (32) is ill formed.

But, if the extraposition processes are cyclic, as we have claimed, then (32) is a possible intermediate form. Our claim makes a rather clear prediction. If we can produce a well-formed surface structure by applying Three-Bar Extraposition on the \bar{P} -cycle and then Two-Bar Extraposition on the \bar{V} -cycle, then we would expect the specifier /g/, belonging to the object of the postposition, to remain overt in the derived structure. This follows, since the application of Three-Bar Extraposition on the \bar{P} -cycle will have the effect of protecting the /g/ from deletion—by moving it away from the superordinate P-category specifier /am/.⁸

Our proposed derivation does in fact produce a well-formed surface structure, to wit:

(33) $\bar{V} [no \ \bar{P} [am] \bar{P} \ wó'ò \ \bar{P} [wéco \ \bar{N} [g \ mísa] \bar{N}] \bar{P}] \bar{V}$

And, as we predict, the N-category specifier /g/ is not deleted.

We conclude, therefore, that (32) (and by analogy (20), (24)) represents a necessary intermediate form in the derivation of well-formed surface structures in Papago. It follows, then, that no constraint can be imposed directly on Three-Bar Extrapolation, to prevent the generation of the ill-formed surface structures with which we have been concerned in this discussion. Instead, the surface structure constraint (25), itself perhaps explicable in terms of perceptual considerations, is required to account for the observed overgeneration.

3. Hopi Relative Clauses⁹

There is an apparent constraint on the use of the Hopi complex noun phrase structure that functions as the relative clause. The structure can be used freely in constructing sentences of the type that correspond to the English sentences of (34):

- (34) a. *The man who wrote the book addressed us.*
 b. *I introduced the man who wrote the book.*
 c. *I introduced the man whom you brought with you.*
 d. *I introduced the man with whom I work.*

It cannot, however, without recourse to periphrasis, be used to construct sentences corresponding to the English sentences of (35):

- (35) a. *The man whom you brought with you will address us.*
 b. *The man with whom I work will address us.*

The point is essentially this. When the complex noun phrase, as a whole, functions as the subject of the main clause, the relative noun phrase (i.e., that noun phrase in the subordinate clause that is coreferential with the complex noun phrase) must likewise function as the subject within its own clause. No such constraint is observed when the complex noun phrase is a nonsubject within the main clause. Thus, it would appear that Hopi exhibits a curious limitation on the accessibility of noun phrases to relativization (cf. Keenan and Comrie, 1972); when the complex noun phrase is itself a subject, relativization is limited to subjects, otherwise not. This is a rather curious picture, and it is not at all like NP accessibility limitations observed in other languages of the world, since it is tied to the grammatical relation the complex NP bears to the main verb, rather than solely to that which the relative NP bears to the verb of the subordinate clause.

In any event, the particular instance of overgeneration in the Hopi case with which we are concerned is that in which a relative NP is a nonsubject while the complex NP containing it is a subject.

By way of introduction to the problem, we must first present a preliminary theory of Hopi base structures, since it is, we feel, the base which is responsible for the overgeneration of interest here. We will use the bar notation, though we do not wish to imply by our facile use of the latter that the theory of the Hopi base has

been adequately worked out or that it is understood in other than superficial terms. Our preliminary theory of Hopi base structures contains the rules embodied in (36) below:

- (36) *Base Rules*
 a. $\bar{A} \rightarrow (\bar{N}) \bar{A}$
 b. $\bar{A} \rightarrow (\bar{B}) \bar{A}$
 c. $\bar{A} \rightarrow (C) * A$

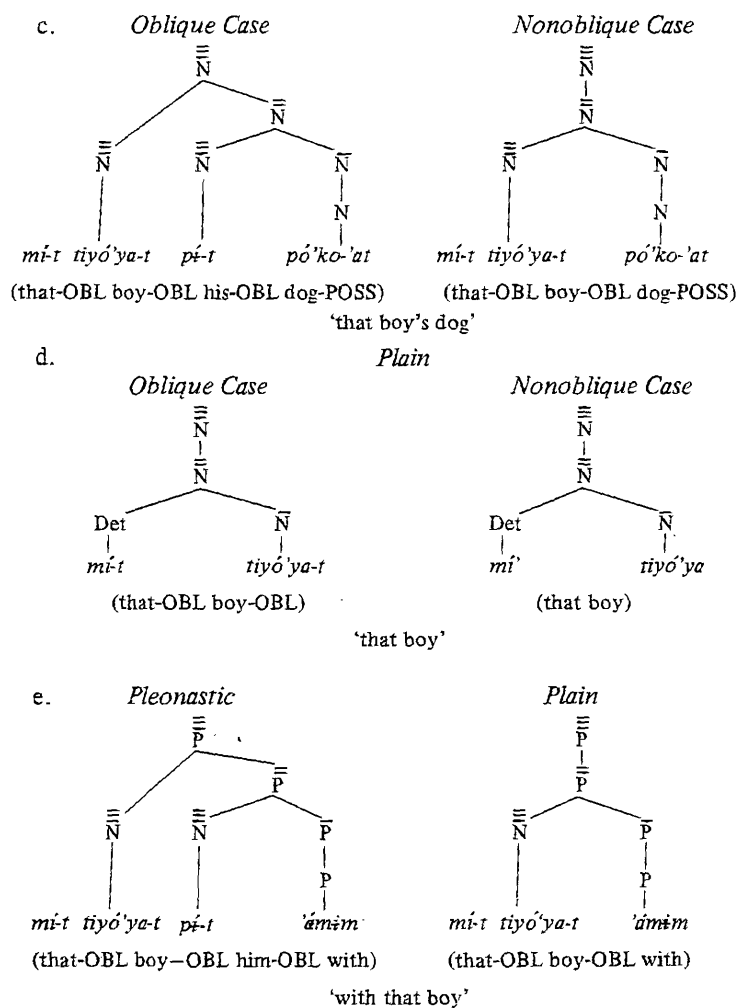
As in Papago, so in Hopi, phrasal categories are nucleus-final. Otherwise, however, the Hopi base is rather different from that of Papago. The initial expansion does not introduce a specifier, as it did in Papago, rather, it introduces an optional noun phrase in "pleonastic" structures, similar to those produced by left-dislocation in English—e.g., 'That man, I can't stand him'—but Hopi generalizes this to give pleonastic structures corresponding to 'that man, his horse', and 'that man, with him'. The core of phrasal categories in Hopi is the two-bar level of structure. These are developed initially by the second expansion, which introduces the subject in a sentence (i.e., in a \bar{V} -structure), the determiner, or alternatively the possessor expression, in a noun phrase (i.e., in a \bar{N} -structure), and the object in a postpositional phrase (i.e., in a P-structure). The third and final expansion, which develops the one-bar structure, introduces the nuclear element (V, N, P) and, in the case of \bar{V} , the nonsubject dependents of the verb (object and postpositional complements). We assume that the N- and P-categories also have preposed dependents: a preposed modifier, certainly, in the case of N, though we have not studied the internal structures of \bar{N} and \bar{P} in any detail as yet.

Examples of structures developed by the rules of (36) are presented below:

- (37) a. *Pleonastic* *Plain*
- mi'* *tiyo'ya* *pam* *taavo-t* *niina*
 (that boy he rabbit-OBL killed)
 'That boy killed a rabbit.'

mi' *tiyo'ya* *taavo-t* *niina*
 (that boy rabbit-OBL killed)
 'That boy killed a rabbit.'
- b. *Pleonastic* *Plain*
- mi'* *tiyo'ya* *ne'* *pe-t* *twi'ta*
 (that boy I him-OBL know)
 'I know that boy.'

ne' *me-t* *tiyo'ya-t* *twi'ta*
 (I that-OBL boy-OBL know)
 'I know that boy.'



A pleonastic structure is well formed only if, somewhere within the two-bar or one-bar structure following the "dislocated" noun phrase, there appears a pronoun that can be construed as coreferential with the latter. The reader will have noticed that Hopi noun phrases are, under certain conditions, marked for a generalized oblique case. We will use a feature notation to represent this fact, and we will devise a set of rules to account for case marking. The oblique case will be designated by means of the feature [+obl], and the unmarked, or nonoblique case will be designated [-obl]. The following elementary case marking rule will account for the assignment of [obl] to noun phrases:

$$(38) \quad \bar{N} \rightarrow \left\{ \begin{array}{l} [-obl] / \bar{V} \\ [+obl] \end{array} \right\}$$

That is to say, a noun phrase is assigned the unmarked case of it immediately precedes \bar{V} or \bar{V} ; otherwise, it will be assigned to oblique case. This will account for the association of oblique case marking (the overt manifestation of which is here glossed OBL) in the structures of (37) above.

In addition to the case assignment rule itself, we must recognize a rule of Case Concord, which ensures that the immediate constituents of a [+obl] noun phrase are also assigned the feature [+obl]. The details of Case Concord, when worked out, will, among other things, ensure that a determiner and the head noun of a [+obl] noun phrase are each supplied with the oblique ending. For the present, however, we will simply assume that a rule of Case Concord exists, without attempting, at this stage of our understanding, to give it explicit formulation.

The overt manifestations of the oblique case category are not particularly straightforward. Limiting ourselves to the singular,¹⁰ and to noun phrases of the type so far considered (i.e., other than those involving relative clauses), the endings are two: /-t/ and /-y/. The latter appears on the head noun in a [+obl] possessive construction, as in sentences like the following:

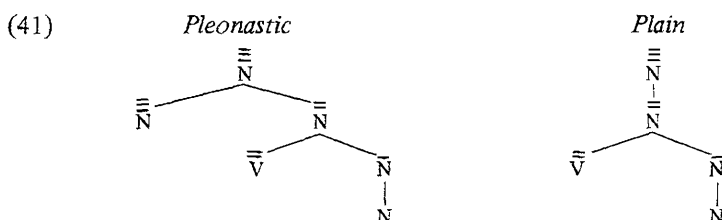
- (39) a. *nɛ' mɪ-t tiyó'ya-t pó'ko-y-at tɬwa.*
 (I that-OBL boy-OBL dog-OBL-POSS saw)
 'I saw that boy's dog.'
- b. *mɪ' tiyó'ya 'i-vóko-y tɬwa.*
 (that boy my-dog-OBL saw)
 'That boy saw my dog.'

Otherwise, for noun phrases not involving relative clauses, the ending /-t/ is used (as is amply exemplified in (37)).

The same two alternants of the singular oblique ending are used on the head noun in a relative clause, but their selection is rather more complex. Before discussing this, however, we must present our theory of the base structures of Hopi relative clauses. We assume that the relative clause is produced by means of the following rule:

$$(40) \quad \bar{N} \rightarrow \bar{V} \bar{N}$$

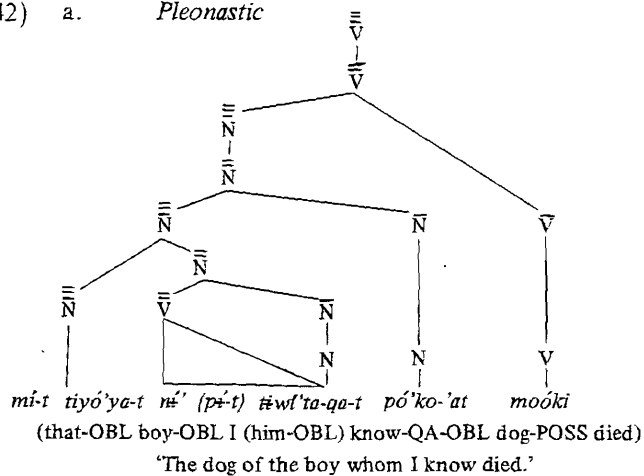
That is to say, we permit rule (40) to be one of the possible forms of (39b). Accordingly, a relative clause has the following structure:



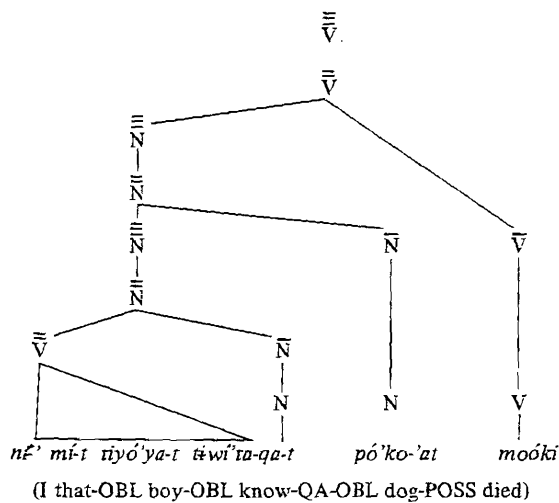
Actually, this structure is used in Hopi not only for the relative clause, but also for the factive nominal complement (i.e., not only for sentences corresponding to 'I

know the boy who killed the rabbit', but also for sentences corresponding to 'I know that the boy killed the rabbit'). Notice further that there exist both pleonastic and plain varieties. The pleonastic variety must, of course, conform to the general condition on well-formedness; in this case, the V-structure must contain a pronoun, or else a NP-gap, which can be construed as coreferential with the left-dislocated noun phrase. It happens that the pleonastic variety is much preferred in the relative clause case, and the plain variety is preferred in the factive case. We do not understand fully the conditions under which the plain variety of the relative clause can be used, but there are in fact circumstances in which either variety is well-formed: for example:

(42) a.

Pleonastic

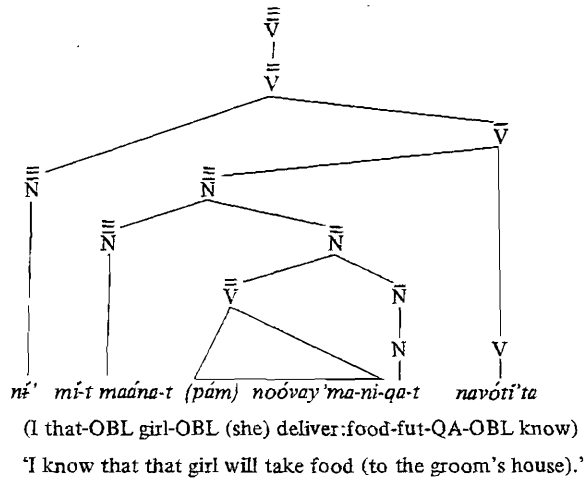
b.

Plain

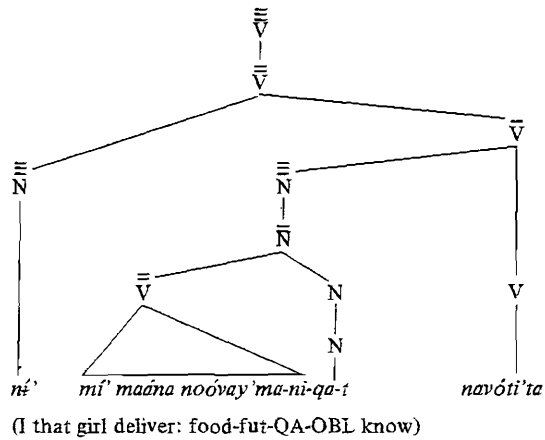
In both of these cases, the complex noun phrase appears in a possessive construction, the most favorable to the use of the plain form of the relative clause.

The following structures illustrate the pleonastic and plain forms of the factive complex noun phrase; in this case, the latter is preferred over the former:

(43) a. *Pleonastic*



b. *Plain*



In (43a), incidentally, we observe that Case Concord has correctly applied to assign the [+obl] case marking to the left-dislocated noun phrase /mí-t maána-t/ in the pleonastic factive complement, itself [+obl] by virtue of its appearance in the environment \bar{V} . By contrast, the very same noun phrase in (43b), i.e., /mí' maána/, is [-obj], because it is in subject position (that is to say, it is in the environment \bar{V}) within its clause.

A property that these complex noun phrases have in common, and that must be discussed in some detail, is the peculiarity of the nuclear element, or head noun. Notice that the head noun is uniformly the element /-qa/, a defective noun with no special semantic content.¹¹ In surface structure, it is unlike any other noun in that it is phonologically bound to the verb word of the subordinate clause. Its only nounlike qualities are the fact that it inflects for case and number and the fact that it appears in precisely the position the X-bar theory predicts for the nuclear element of endocentric phrasal categories in Hopi.

A very special property of the /-qa/ element is the case marking it exhibits. The fact is, Case Concord will not account for the peculiarities of *qa*-case marking. Rather, case marking in this instance is intimately interwoven with the phenomenon of obviation that prevails in Hopi complex sentences of all kinds.¹² That is to say, to fully determine the appropriate case marking of the defective head noun /-qa/, it is not sufficient merely to know the grammatical relation which the complex noun phrase bears to the main verb; one must also know: (1) the grammatical relation the relative noun phrase bears to the subordinate verb; and (2) whether the main and subordinate clauses share subjects, i.e., whether the subjects of the main and subordinate clauses are coreferential. The principles of *qa*-case assignment can be expressed roughly as follows (limiting ourselves to the singular):

- (44) *qa*-Case Assignment
- a. [-obl] (/qa/): shared subjects, and relative NP is subject in its own clause;
 - b. "proximate" [+obl] (/qa-y/): shared subjects, and relative NP is nonsubject in its own clause;
 - c. "obviative" [+obl] (/qa-t/): otherwise.

The following three sentences illustrate the three cases subsumed under (44a-c):

- (45) a. *mí' tiyó'ya (pám) 'acáta-qa pákniimíya.*
(that boy (he) lied-QA cry)
'The boy who lied is crying.'
- b. *né' taávo-t (né') (pi-t) níina-qa-y sisk^{wa}.*
(I rabbit-OBL (I) (it-OBL) killed-proxOBL skinned)
'I skinned the rabbit that I killed.'
- c. *né' mí-t tiyó'ya-t (pám) 'acáta-qa-t hoóna.*
(I that-OBL boy-OBL (he) lied-QA-obvOBL sent:home)
'I sent home the boy that lied.'

We have now set the scene for a discussion of the case of overgeneration that concerns us. Let us consider the offending structure in abstract terms, to begin with. Such a structure would be the following, in which the circled noun phrase is the relative noun phrase:

- (50) **mi' tiyó'ya nē' (pá-t) tēwa:-qa-t pay níma.*
 (that boy I (him-OBL) saw-QA-obvOBL already went:home)
 'The boy whom I saw has gone home.'

—the corresponding extraposed variant is well-formed:

- (51) *mi' tiyó'ya pay níma, nē' (pá-t) tēwa:-qa-t.*

If we are correct in assuming that (51) is derived by means of the extraposition rule, the only conceivable source for the sentence is (50). Therefore, the ill-formed surface structure (50) must, from the point of view of the base rules of Hopi, be perfectly well formed.

We are left, of course, with the problem of explaining the surface ill-formedness of sentence (50). We would like to suggest that the answer is to be found in a consideration of the case marking exhibited by the relative clause in (50), and in its extraposed counterpart (51). Notice that the /-qa/ element is marked with the obviative oblique ending /-t/—this is exactly as it should be, since the sentence as a whole—cf. (46), its structure in skeletal form—conforms to case (c) of the *qa*-Case Assignment principles. In the case of sentence (50), this results in the circumstance that a complex noun phrase in **subject position** terminates in an oblique ending. This is in defiance of the prevailing surface fact of Hopi that noun phrases in subject position are unmarked for case. It is conceivable (and we suggest that this is indeed the answer) that the ill-formedness of (50) is due to the fact that it violates this otherwise general principle of surface structure case marking.¹⁴

In any event, it seems quite clear to us that no special constraint need, or in fact can, be placed upon the rules of the base which are responsible for the development of structures functioning as relative clauses in Hopi.

4. Noun Phrase Gaps in Navajo.¹⁵

In Navajo surface structures, relative clauses are either of the "headed" type, with the head following the subordinate clause, as in (52a) below, or else they are of the "headless" type, as in (52b):

- (52) a. *Deezhtlzh-éé ashkii yicha.*
 (stumbled-REL boy cry)
 b. *Ashkii deezhtlzh-éé yicha.*
 (boy stumbled-REL cry)
 'The boy who stumbled is crying.'

That the complex noun phrase in sentences like (52b) is headless, and not merely left-headed as the English analogue would suggest, is shown by sentences of the type represented by (53)—

- (53) *Ashkii tēechqa'í bi-shxash-éé nahat'in.*
 (boy dog bi-bit-REL bark)
 'The dog that the boy was bitten by is barking.'

—in which /tʰéechq̄q̄'i/ 'dog' is clearly within the subordinate clause, being flanked by material belonging thereto. It is further shown by the ambiguity observed in sentences like (54a-b):

- (54) a. *Ashkii at'éed yi-zts'qs-éq̄ yáti'*.
 (boy girl yi-kissed-REL speak)
 'The boy who kissed the girl is speaking.'
 'The girl whom the boy kissed is speaking.'
- b. *At'éed ashkii bi-zts'qs-éq̄ yáti'*.
 (girl boy bi-kissed-REL speak)
 'The girl who was kissed by the boy is speaking.'
 'The boy by whom the girl was kissed is speaking.'

The ambiguity here is due to the fact that, since the noun phrase that is understood as coreferential with the complex noun phrase as a whole is within the subordinate clause, rather than in head position, it is not possible to tell, from the surface structure itself, which of the two noun phrases in the transitive subordinate clause is to be construed as the subject of the main verb. For many speakers of Navajo, to be sure, there is a clear preference for construing the first subordinate noun phrase with the main verb, but for most, if not all speakers, the other interpretation is also clearly available. This sort of ambiguity is, of course, lacking in the case of corresponding sentences using the headed relative clause—though this alternative, exemplified by (55a-b) below, is greatly disfavored in actual Navajo usage:

- (55) a. *At'éed yi-zts'qs-éq̄ ashkii yáti'*.
 (girl yi-kissed-REL boy speak)
 'The boy who kissed the girl is speaking.'
- b. *Ashkii bi-zts'qs-éq̄ at'éed yáti'*.
 (boy bi-kissed-REL girl speak)
 'The girl who was kissed by the boy is speaking.'

We are not sure how the headed relative clause is developed in Navajo—whether by deletion of a noun phrase from the subordinate clause, by permitting the base rules to introduce a null proform in the subordinate clause, or by raising a subordinate noun phrase rightward into head position.¹⁶ In any case, an NP-gap is present in the subordinate clause; that is to say, a noun phrase is missing from the subordinate clause. Also, under any of the alternative conceptions of the headed relative clause, we are faced with a minor problem—namely, that of explaining why sentences of the type represented by (55a-b) are not ambiguous. The single possible interpretation of each of the sentences of (55) indicates clearly that the first noun phrase in the subordinate clause is the one that is missing—i.e., the NP-gap is initial. In principle, surely, the NP-gap could be in second position—the fact is, however, the gap-second interpretations are simply absent for (55a-b).

An identical observation can be made in the case of the mechanism that accounts for NP-gaps in sentences like the following:

- (56) *Ashkii yah'ttyd' áádóó at'ééd yi-zts'qs.*
 (boy entered and:then girl yi-kissed)
 'The boy came in and then he kissed the girl.'

This sentence consists of an intransitive (first) clause conjoined to a transitive (second) clause. Again, there is an NP-gap in a transitive clause. And the sole interpretation associated with the compound sentence indicates that the gap is initial in the transitive clause, the theoretically possible gap-second interpretation being entirely absent.

It would appear, therefore, that a constraint must be placed upon the mechanisms that produce NP-gaps in transitive sentences. One might propose, for example, that an NP-gap must be initial within a transitive sentence. Such a constraint would also account for the interpretations of sentences involving adverbial clauses, as in:

- (57) a. *Ashkii yah'íyáa-go at'ééd yi-zts'qs.*
 (boy entered-SUBORD girl yi-kissed)
 'When the boy came in, he kissed the girl.'
- b. *Ashkii yah'íyáa-go at'ééd bi-zts'qs.*
 (boy entered-SUBORD girl bi-kissed)
 'When the boy came in, he was kissed by the girl.'
- c. *At'ééd yi-zts'qs-go ashkii dahdiiyá.*
 (girl yi-kissed-SUBORD boy left)
 'When he kissed the girl, the boy left.'
- d. *At'ééd bi-zts'qs-go ashkii dahdiiyá.*
 (girl bi-kissed-SUBORD boy left)
 'When he was kissed by the girl, the boy left.'

A constraint disallowing gap-second transitive clauses would account for the observations we have made so far, but it would not also account for the interpretations of sentences of the type represented by (58) below:

- (58) *Ashkii at'ééd yi-zts'qs-éé hastiin yi-yiíłtsq.*
 (boy girl yi-kissed-REL man yi-saw)
 'The boy who kissed the girl saw the man.'
 'The girl whom the boy kissed saw the man.'

This sentence is ambiguous in two ways, which is understandable given the fact that the subject of the main verb is a complex noun phrase consisting of a headless relative clause. The problem is this: why is the sentence not ambiguous in three ways, permitting a reading in which the object of the main verb consists of the otherwise perfectly well-formed relative clause /at'ééd yi-zts'qs-éé hastiin/ 'the man who kissed the girl'? It would appear that we must extend our constraint on the occurrence of NP-gaps to disallow also an initial gap if it appears in a transitive relative clause that is itself in second position within the matrix. This cannot be, however, since (59) below is perfectly well formed, and surely its complex noun phrase object exhibits a clause-initial NP-gap:

- (59) *Ashkii at'ééd yi-yiit'sá(n)-éé yi-zts'qs.*
 (boy GAP girl yi-saw-REL yi-kissed)
 'The boy kissed the girl whom he saw.'

These observations by no means spell the doom of a constraint on the appearance of NP-gaps in Navajo. In fact, a rather simple constraint can be placed on NP-gaps that will account for all of the observations made thus far—namely, (60) below, in which the symbol 0 represents an NP-gap, and in which the structural description is to be understood as applicable in the order given and in the mirror image order:

- (60) *Constraint on NP Gaps*
 No structure is permitted that can be factored as follows:
 $X - NP - Y - 0 - Z$
 if (1) 0 is construed as coreferential with NP, and (2) a variable immediately preceding 0 ranges over a string terminating in a noun.

This will not only account for the interpretations of complex sentences of the type we have been considering, but it will also account for the fact that simple sentences like (61a-b) are unambiguous with respect to coreference between a full noun phrase and an NP-gap:

- (61) a. *Ashkii bi-zhé'é yi-yiit'sá.*
 (boy his-father yi-saw)
 'The boy saw his father.'
 b. *Ashkii bi-zhé'é b-iit'sá.*
 (boy his-father bi-saw)
 'The boy was seen by his father.'

The crucial point here is that the constraint embodied in (60) will ensure that (61a) cannot receive an interpretation in which the NP-gap immediately precedes the verb ('The boy_i's father saw him_i'), structurally parallel to:

- (61) c. *Ashkii bi-zhé'é at'ééd yi-yiit'sá;*
 (boy his-father girl saw)
 'The boy's father saw the girl.'

The nonambiguity of (61b) is, of course, explained by the general constraint on proforms—i.e., that they may not both precede and command their full NP-anaphors, which rules out a reading in which the NP-gap is sentence initial (*'He_i was seen by the boy_i's father'), as well as by the constraint (60), which rules out an interpretation in which the NP-gap immediately precedes the verb ('The boy_i's father was seen by him_i'), structurally parallel to:

- (61) d. *Ashkii be-zhé'é at'ééd b-iit'sá;*
 (boy his-rather girl bi-saw)
 'The boy's father was seen by the girl'

A simple modification of (60), permitting the use of clause (2) without clause (1), would also account for the fact that transitive fragments, with or without a postposed "afterthought" noun phrase, are unambiguous, at least with respect to the position of the NP-gap:

- (62) a. *At'ééd yi-zts'qs.*
 (girl yi-kissed)
 b. 'He/she kissed the girl.'
At'ééd yi-zts'qs, ashkii.
 (girl yi-kissed, boy)
 'He kissed the girl, i.e., the boy.'
 c. *Ashkii bi-zts'qs.*
 (boy bi-kissed)
 'He/she was kissed by the boy.'
 d. *Ashkii bi-zts'qs, at'ééd.*
 (boy bi-kissed, girl)
 'She was kissed by the boy, i.e., the girl.'

These sentences do not admit of an interpretation according to which the NP-gap is other than in absolute initial position.

We do not know whether the proposed constraint could be formulated as a condition on particular rules which are responsible for creating NP-gaps—more to the point, we are not at all certain, as yet, how NP-gaps are in fact to be produced, whether by movement rules, deletion rules, both, or by rules introducing null pro-forms in the base. But this is not a serious hindrance to the fundamental idea of a constraint like (60). We could, for example, think of it as a constraint on surface structures in which a NP-gap is understood as coreferential with a full noun phrase—sentences that violate the stricture on the variable preceding the gap would, by the constraint, be identified as ill-formed. Alternatively, we might recast the constraint as a limitation on the interpretation of NP-gaps: "No NP-gap may be construed as coreferential with NP' if it is preceded by a string terminating in a noun that is not the head of NP'." This would simply leave uninterpreted (rather than strictly speaking, ill-formed) certain sentences produced by the syntax.

Either way, the constraint works quite well. However, we would like to suggest that this is entirely the wrong way to look at the matter. First of all, if it is conceived as a constraint on rules of Navajo syntax or semantics, it is clearly out of keeping with the Autonomous Systems thesis. As a constraint on surface structures, in roughly the form given in (60), it is not clearly inconsistent with the autonomous position, but, we submit, it is still misguided, since it is formulated in terms of an NP-gap in construal with a full noun phrase. Surely the problem is not in the **construal** of NP-gaps, but rather it is related in some way to the problem of **locating** them. This must be so, since the entire problem can be understood in terms of interpretive principles that must surely exist independently of the question of relating NP-gaps to full noun phrases.

Consider Navajo sentences of the sort introduced in the beginning section of this essay. Surely there must exist principles of interpretation that permit Navajo speakers to determine the grammatical relations the noun phrases bear to the verb in structures which exhibit the surface profiles:

NP NP *yi-V*

and

NP NP *bi-V*

in which the noun phrases are third person. The principles might be worded roughly as follows:

(63) *The Interpretation of Grammatical Relations*

NP is:

- a. object/ — *yi-v*
subject/ — *bi-V*
- b. subject/ — NP *yi-V*
object/ — NP *biV*.

If we assume that these principles apply alike to surface profiles in simple and complex sentences, we can begin to see that the problem, as we have been thinking of it heretofore, simply vanishes.

We can locate an NP-gap in the following way: if part (b) of (63) cannot apply, there is an NP-gap in the environment:

$$\text{— NP } \left\{ \begin{array}{l} bi- \\ yi- \end{array} \right\} \text{ V}$$

Let us imagine now that, once we have located the gap in this way, we replace it with some symbol, say "NP" (i.e., noun phrase in quotes) and then apply (63b) to determine the grammatical relation "NP" bears to the verb. We can then proceed to what remains of the so-called construal problem, which, presumably is now nothing more than the well-known process of associating an NP-proform with a possible full NP-anaphor—in the Navajo case, general principles of coreference will determine with which overt noun phrase, if any, the proform "NP" can (or must, as the particular case requires) be understood as coreferential; this is a separate issue, and is not directly germane to the questions with which we are concerned here.

Let us see how this works in actual practice. Consider, for example, sentence (56). The relevant substring is /at'ééd yi-zts'qs/. Notice that, since no noun phrase immediately precedes this string, (63b) will fail to apply. We therefore know that there is an NP-gap preceding the noun phrase /at'ééd/. Our gap-substitution procedure will insert "NP" into that position, permitting (63b) now to apply and to determine the grammatical relation which the missing noun phrase bears to the verb—subject, in this instance. Quite general principles of coreference will allow "NP" to be associated anaphorically with the noun phrase /ashkii/ in the first conjunct clause, giving the reading corresponding to the English of (56). They will also, as it happens, permit a distinct reading according to which "NP" refers outside

the sentence. The reading that is not allowed is that in which an NP-gap is located between /at'ééd/ and the verb /yi-zts'qs/. But notice that this is not, in our revised conception of the problem, a matter of construal; rather, it is related to the operation of (63), which simply rules out the possibility of a parsing of (56) that locates the gap in object position. A question remains, therefore. What is the status of the theoretically possible structure in which an NP-gap is in fact located between /at'ééd/ and /yi-zts'qs/? Is the sentence in fact structurally ambiguous in this way, or should this alternative structure be ruled out in some way?

This is a crucial juncture in our discussion.

We feel that the only position we can possibly take that is consistent with the Autonomous Systems thesis is that (56) is structurally ambiguous. It is simply the case, we wish to claim, that the force of (63) is so compelling in cases of this sort that it renders inaccessible an interpretation corresponding to the structure in which the gap follows, rather than precedes, the overt noun phrase.¹⁷

Before concluding, let us observe the operation of (63) in a situation of a different sort, one in which both clauses of (63) are immediately applicable. Consider, in this connection, sentence (58). The relevant substring here is /ashkii at'ééd yi-zts'qs/. Notice that both clauses of (63) will apply in this instance, and therefore, crucially, will fail to locate an NP-gap. This explains why the sentence admits only readings according to which the boy kissed the girl. Yet we know that a theoretically possible syntactic structure associated with the terminal string (58) is one in which an NP-gap occurs between /ashkii/ and /at'ééd/—thereby admitting a reading according to which the man (/hastiin/) kissed the girl. This follows from the well-formedness of the complex noun phrase /at'ééd yi-zts'qs-éé hastiin/. In line with our announced position on this matter, we would contend that sentence (58) is in fact structurally ambiguous and that it is the force of (63) that limits the range of interpretations to those reflected in the English translations we have provided.

We have not sought here to develop the full complement of surface structure interpretation (or parsing) principles whose existence we feel it is surely reasonable to imagine. Nor have we fully explored the capabilities of (63). We would, however, maintain that the Navajo phenomena with which we have been concerned in this discussion are best understood in these terms and not in terms of constraints on rules of syntax or semantics.

5. Conclusion¹⁸

Our conclusion will have relatively little to do (directly, that is) with the three cases of overgeneration we have used to illustrate the tactical approach to linguistic problems that we feel is consistent with the Autonomous Systems view of language: briefly, the tactic of scepticism toward local, or global but rule-linked, conditions on the application of specific rules of formal grammar. We would like here to make certain remarks of primarily methodological character concerning the tasks of a

linguist—particularly one who is interested in language typology—in the context of a conception of language as an autonomous system of Autonomous Systems. And our concrete examples will be drawn from a language that is “worlds apart” from those used in the foregoing sections—namely, Walbiri, of Central Australia.

We find the typological work of C. F. Voegelin to be of considerable interest in that it poses rather clearly the questions that must be addressed by a linguist working within the Autonomous Systems framework. Voegelin's typology is *subsystem* typology, a tradition that he himself traces back to the work of Trubetzkoy (Voegelin, 1955). This is contrasted with the *whole language* typology of the 19th century, continued in the 20th by Sapir. The distinction between whole language and subsystem typology quite clearly depends upon the recognition of a degree of autonomy for linguistic (sub)systems in the latter tradition. By contrast, the former tradition “was set up under the assumption that some morphosyntactic features of a given language will serve as an index or indication of the type of that language as a whole. Such typology attempts to measure the whole language—by implying a correlation between parts of the language treated in the typology and parts not treated” (Voegelin, Ramanujan, and Voegelin, 1960, p. 198). Voegelin does not deny that there are, or may be, correlations between parts of a language: “We move cautiously from one part of a language typologized to another part. If the typology of one part implicates that of another part, that would be important—but discoverable only after preliminary typologizing of more than one subsystem across languages” (Voegelin, Ramanujan, and Voegelin, 1960, p. 199). Clearly, this entire enterprise rests upon the assumption of a degree of autonomy among the parts, or systems, within a language.

If this is correct, and we think it is, then there is a clear challenge present for the linguist in the study of language universals. In the Autonomous Systems tradition, the issue of comparability of systems across languages is critical, as is the problem of identifying the systemic “primes” in language. We will not attempt to address these issues here—since that is the whole science of language typology on the Autonomous Systems view—but we would like to explore very briefly one of the themes which Voegelin has identified as important in the tradition of subsystem typology, i.e., the question of the implicational relations among the systems of language. He offers, among others, the following task for typological work: “To discover laws of implication (functional relations) among subsystems. The question of what constitutes equivalence between languages, aside from cognates . . . , can be considered under implication.” (Voegelin, 1955, p. 360).

Let us imagine that we are able somehow to answer the question of cross-language equivalence for particular subsystems and that we can, therefore, proceed directly to the task of discovering laws of implication. Let us then consider the following (partially, but not entirely hypothetical) typological difference in the base structure systems of languages:

Type I Base Structure

The categorial rules of the base are of the form

$$X^n \rightarrow \dots \widehat{A} \widehat{X^{n-1}}$$

or

$$X^n \rightarrow X^{n-1} \widehat{A} \dots$$

That is, the rules that develop phrasal categories define (1) a labelled constituent (or bracketing) structure, and (2) a concatenation (or linear ordering) of constituents. This is presumably the type which English, Papago, and Hopi, for example, represent. By contrast, languages of the second type conform to the following scheme:

(65) *Type II Base Structure*

The categorial rules of the base are of the form

$$X^n \rightarrow \dots, A, X^{n-1}$$

In this type, the rules that develop phrasal categories define a labelled constituent structure, but they do not specify a concatenation of constituents. Instead, the observed surface word order for any actual sentence is specified by an entirely distinct system, a schema that assigns linear precedence relations among the words of a sentence, subject to general constraints (such as the constraint that a word from one tensed clause may not be ordered among the constituents of another clause, and so on). A candidate for this type is Walbiri. It is typical of Type II languages, like the "scrambling" Walbiri, that they exhibit highly variable surface word order. We would claim, however, that this is not necessarily so. We would claim, rather, that there could be Type II languages with relatively rigid surface word order and, conversely, Type I languages—e.g., Papago—with relatively variable surface word order.

What implications does this base structure typology have for other linguistic systems? It will, for one thing, necessarily be the case that while a Type I language may or may not have a transformational component, a Type II language will definitely **not** have a transformational component. This follows from the definition of the notion "grammatical transformation"—as a structure-dependent operation on concatenated constituents.

What are the empirical reflections of this, say in a language like Walbiri, which is a putative candidate for membership in Type II? Consider the following. While it is a fact about content questions in Walbiri that the question word appears in initial position, typically, it is very unlikely that it arrives in that position through the action of a movement transformation. Since word order is variable in Walbiri, the initial position of question words can be effected by merely choosing that word order for content questions. If this is so, then we can perhaps understand why Walbiri is able to form question-word-initial interrogative sentences corresponding to the English "Where did John tell Bill to go?"—i.e., with the question word extracted from an infinitival clause—but not question-word-initial interrogative sentences corresponding to "Where did John say (that) Bill went?"—i.e., with the

question word extracted from a tensed clause. This follows from the Tensed Sentence Condition (Chomsky, 1973)—or a version thereof—which, in this instance, would prevent the word ordering schema from positioning any constituent belonging to a tensed clause out of that clause and into the matrix.

There are other ways of conceiving of these Walbiri facts, but the proposal we have just made seems to us to be the most plausible, all things considered. Now, there may be languages of Type I that are essentially like Walbiri, by virtue of their lacking a transformational component. However, we would maintain that if, on some grounds or other (and there are additional grounds in the Walbiri case, we believe), a given language is determined to belong to Type II, then it will necessarily be like Walbiri with respect to the process of content question formation, i.e., it will fail to form questions by extraction from a tensed clause. (For a somewhat similar observation, but from a different theoretical perspective, dealing with better-known languages, see the extremely interesting paper by Bechhofer, 1975.)

If this is correct, then what we have here is an instance of *necessary implication* in association with a typological characteristic, following (1) from the nature of the base in the typologically distinct languages and (2) from the notion “transformational rule,” strictly defined. It is reasonable to enquire whether, in the Autonomous Systems conception of language, all intersystemic implications will be necessary implications—that is to say, does it follow from the Autonomous Systems view that this will be the case? Are there implications of another sort—implications that might be termed *relative*? We do not know the answer to this question, but we would like, nevertheless, to speculate that there are, in fact, relative implications and that their existence is in no way inconsistent with the Autonomous Systems thesis. Let us consider Walbiri again—as a possible exemplar of Type II.

In Walbiri, as in most languages of which we have any direct knowledge, the grammatical relations *subject* and *object* play a vital role in the rules of grammar. However, it seems quite clear to us that the definition of these notions, with respect to the rules that make use of them, cannot be the same as in a language with a Type I base structure. Thus, while in a Type I language, the subject, for example, might correspond to “the first noun phrase in the sentence,” or “the last noun phrase in the sentence,” or some such thing, in a Type II language such positional definitions of grammatical relations are impossible, by the very nature of the Type II base structure. In Walbiri (according to the null hypothesis, at any rate), the notion subject can adequately, and we think quite correctly, be identified in terms of the case system. The subject is the ergative noun phrase, if there is one, otherwise the absolutive. The object is defined in analogous fashion as the dative, if there is one, the absolutive otherwise. These disjunctive definitions function correctly in most, if not all, areas of Walbiri grammar in which the grammatical relations subject and object play a crucial role. Nevertheless, there is commonality between the Walbiri grammatical relations and those of other languages, which is not at all captured in the scheme we have just outlined here, i.e., in which the subject, say, is defined in terms of the linear order in one language, but in terms of the case system in another. This is cause for concern, since the scheme fails to reflect certain universal, or

near-universal, observations about grammatical relations, such as the observation that in most languages the semantic agent is the subject (in the basic form of a sentence), while the patient is typically the direct object; and the observation that, in languages with an ergative-absolutive type of case system, the semantic agent in a transitive sentence is typically (perhaps always) in the ergative case; and so on.

Our feeling is that the universals in this case belong properly to the category of *relative implications* within the typology that grows out of the Autonomous Systems view of language. That is to say, these universals are not to be represented directly in the grammars of specific languages. Rather, we would suggest, they are to be represented by means of a system analogous to the "marking" or "linking" conventions of phonology (cf. Chomsky and Halle, 1968, Chapter 9) which specify the degree of markedness—the naturalness—of particular phonological processes, morpheme structures, and phonological segments. In relation to case systems, for example, such linking or marking conventions would specify the extent to which the use of a particular case, or the case array associated with a particular verb, is marked or natural within the context of a general theory of naturalness for case categories vis à vis functional, or thematic, structure. And with respect to grammatical relations, analogous conventions would define the extent to which such notions as subject and object, defined for a particular base structure, are marked or natural within the context of a general theory of naturalness for the association of grammatical with logical and thematic relations.

To give somewhat more substance to this suggestion, let us imagine a proposal along the following lines. The notions *subject* and *object* (not in quotes) properly belong to the system of representation that defines the logical form of sentences—assuming that logical form specifies a subject-predicate partitioning for sentences. We might, for example, represent the logical form of a transitive sentence like "Japangardi speared the kangaroo," using the λ -notation of Church (1941, and very successfully developed for a wide variety of long-standing problems in English grammar by Sag, 1976), as follows:

(66) *Japangardi*, $\lambda x(\textit{the kangaroo}, \lambda y(x \textit{ speared } y))$

Correspondingly, the logical form of an intransitive sentence like "the kangaroo ran away" might be represented as follows:

(67) *the kangaroo*, $\lambda x(x \textit{ ran away})$

The notion *subject* corresponds to the term directly preceding the maximal λ -predicate, i.e., *Japangardi* in (66), *the kangaroo* in (67). The notion *object*, in a transitive sentence, corresponds to the term directly preceding the minimal, or innermost, λ -predicate, i.e., *the kangaroo* in (66).

Now let us see how the notions "*subject*" and "*object*" (in quotes) are defined for Walbiri—e.g., for sentences whose logical form corresponds to (66-67):

(66') *Japangardi-rli wawirri pantu-rnu.*
(Japangardi-ERG kangaroo spear-PAST)

(67') *Wawirri parnka-ja-rra.*
kangaroo run-PAST-away)

We have devised a feature system to represent the case categories of Walbiri, a minimal departure from the surface case marking occasioned by a certain degree of syncretism in case morphology. The features we have chosen are [subj] and [obj]. Ergative case is represented [+subj, -obj] (for mnemonic purposes, read this roughly as "capable of being the subject, but not the object"); the absolutive case is represented [+subj, +obj] (read "capable of being the subject or the object, depending upon the transitivity of the verb"); and the dative case is represented [+obj, -subj] (read, "capable of being the object, but not the subject"). These are the "pure grammatical" cases; all other cases, i.e., the "thematic cases" (e.g., locative, allative, elative, etc) are represented [-subj, -obj]. Using this feature system, we can define the Walbiri notions "subject" and "object" by means of the following disjunctions:

(68) "subject": [+subj (-obj)]
"object": [+obj (-subj)]

In accordance with the use of parenthesis to express disjunction (as in phonology), these will correctly select the ergative, if there is one, otherwise the absolutive, for "subject," and they will select the dative, if there is one, otherwise the absolutive, for "object." In the normal case, Walbiri "subject" and "object" will correspond exactly to logical *subject* and *object*. This is not a logical necessity, however. In fact, Walbiri weather sentences, like (69a-b) may be exceptional in this regard:

- (69) a. *Ngapa ka wanti-mi.*
(rain PRES fall-NONPAST)
'It is raining.'
- b. *Payi ka wangka-mi.*
(wind PRES speak-NONPAST)
'It is windy, the wind is blowing.'

While these sentences have "subjects" (/ngapa/ 'water, rain', and /payi/ 'air, wind'), it is not at all clear that these are to be identified with the *subject* in the corresponding logical form of weather sentences.

Now let us imagine the sorts of things the naturalness conventions might say in this connection. The lexical representation of each Walbiri verb, we assume, will have associated with it a strict subcategorization stipulating the case categories of its pure grammatical dependents. In addition, the lexical representation will stipulate the thematic relations the dependents bear to the verb. The unmarked, or natural situation for a transitive verb, we suggest, would be that in which, say, the ergative dependent is the agent, the absolutive is the patient, and so on. Given this, we might go on to suggest that grammatical and logical relations are associated with thematic relations in ways that can be characterized as marked or unmarked as well. Thus, for example, the grammatical relation "subject" (and the logical relation *subject*) in a transitive sentence will, in the unmarked or natural case, correspond to the thematic relation *agent*—in the basic form of sentences, that is. Walbiri conforms

to this expectation, by and large, as do most languages with which we are acquainted. Again, however, this is not a logical necessity. It is at least logically possible for a particular verb, a class of verbs, or even an entire language, to depart from this norm to some extent. A strong case has been made by Dixon, for example, that the Dyirbal language of North Queensland exhibits a basic form of transitive sentences in which the "subject" (corresponding to logical *subject*, as expected) is identified not with the *agent*, but with the *patient* (Dixon, 1972). We suggest that the naturalness conventions, once they are fully elaborated, would identify this as a highly marked system, perhaps accounting for the relative rarity of the Dyirbal type among the languages of the world.

Notes

- * This work was supported in part by the National Institute of Mental Health Grant No. MH-13390-08.

The division of labor in preparing this paper has been as follows. Hale is responsible for the prose of the paper as a whole and for sections 1, 2, and 5. Section 3 is a condensation and slight recasting of parts of LaVerne Masayeva Jeanne, "A Reconsideration of the Relative Clause in Hopi" (1975). Section 4 is based on work by Paul Platero (1974), Hale and Platero (1974), and subsequent investigations by Hale and Platero.

1. These introductory sections are taken, with only slight revision, from Hale (1976).
2. For further information concerning this aspect of Navajo grammar, see Hale (1973) and Creamer (1974).
3. This section comprises a thorough revision of a paper by Hale entitled "Papago Intonation and Word Order," prepared originally for inclusion in the volume of *IJAL* dedicated to Stanley Newman. Unfortunately, the paper was not finished in time to appear in that volume. Hale would like now to dedicate this section to Stanley Newman in gratitude for benefits derived from Newman's teaching and general scholarship, both in the context of a study of his published works and in the context of his excellent courses while Hale was his student in 1956-57.

Hale owes his current understanding of the cyclic nature of the Papago extraposition rule described here almost entirely to discussions with other linguists—especially Lisa Selkirk, Ivan Sag, and Erich Woisetschlaeger. These people are not responsible for Hale's mistakes, however.

The orthography used in this section is essentially that of Alvarez and Hale, 1969, with the exception that long vowels are written double rather than with a colon.

4. For a discussion of the auxiliary in Papago, see footnote 3 of Hale, 1969, the grammar section of Saxton and Saxton, 1969, or the relevant sections of Steele, 1975.
5. A second class of exceptions is treated briefly in Hale (1975), namely, the phrase-internal fall in pitch from a stressed specifier Spec_A in the context / $_ \bar{A}$. It is omitted from the present discussion, since it does not bear directly on the case of overgeneration with which we are concerned.

6. It is possible to collapse Three-Bar and Two-Bar Extrapolation into a single rule schema, as follows, assuming that the structural description below is a legitimate one, and assuming further that the parentheses designate an option merely, rather than a disjunctive ordering relation between the two rules:

$$\begin{array}{ccccccc} X & - & (\text{Spec}) & \bar{A} & - & X & \\ 1 & & 2 & & 3 & & \\ 1 & & 0 & & 3 & & 2 \end{array}$$

Ivan Sag has pointed out to us that the intent of this formulation may constitute a violation of his Immediate Domination Principle (Sag, 1976) and Bresnan's Relativized A-over-A Principle (Bresnan, 1976).

7. The stressed singular specifiers have short and long forms—/hégai, hég/ 'that' /fida, 'íd/ 'this'. The short forms are used in preverbal position when \bar{N} is missing from the \bar{N} to which the specifier belongs. The phonologically weaker and semantically empty N-category specifier /g/ appears in its augmented form /heg/ when \bar{N} is extracted from it.
8. It is remotely possible that there exists a noncyclic analysis that would lead to the same expectations. This would entail, we would imagine, a strict separation of the two extrapolation processes and the establishment of an extrinsic ordering relation between them. We have sought in vain for such an analysis that would account for the range of facts successfully handled by the cyclic theory. Another possibility that should not be entirely overlooked is that suggestion by the Saxtons that Papago is basically nucleus-initial (Saxton and Saxton, 1969) and that the movement rules discussed here operate leftward, rather than rightward as we suggest. We have not fully investigated this possibility in our work on Papago, though cases of rightward extrapolation that violate subadjacency—e.g., /'an 'óod 'án g káwyu/ (Spec_P (Spec_N) N P Spec_N N) 'on the horse's back' (lit., 'there its back on the horse'), presumably produced by successive cyclic extrapolation (cf. the nucleus-final variant /'an káwyu 'óo 'án/—seem to us to be more tractable within the nucleus-final theory than within the nucleus-initial theory.
- Joel Rotenberg is currently working out an alternative analysis of Papago that would account for the total range of possible surface word orders *without* generating the ill-formed (20) at all; his analysis assumes the nucleus-final base but uses a leftward, rather than a rightward, movement rule. Hale is also exploring the possibility that Papago has a Type II base structure (see section 5 of the text) and that the ill-formed (20) is precluded by conditions on word order, rather than on the appearance of #-junctures in surface structures.
9. The Hopi orthography used in this section is one consistent with the findings in Jeanne (1974). We have indicated stress, though this is for the most part redundant. The sequence /VV/ is perceived as a level tone, and the sequence /V'/ is perceived as a falling tone, at least in the Third Mesa dialect to which these data belong.
10. The oblique case ending is universally /-y/ in the plural.
11. Our original conception of the /-qa/ element was that it was a complementizer and that the head of a relative clause appeared, if at all, to the left of the subordinate clause. However, evidence from factive complements indicates rather clearly that the noun phrase appearing to the left of a qa-terminated sentence cannot be the true head (Jeanne, 1975). Moreover, left-headed relative clauses would fail: (1) to reflect the parallelism between relative clauses and other structures in the use of pleonastic forms; and (2) to conform to the otherwise universally nucleus-final base structure of Hopi. It is possible, however, that we are wrong in our reanalysis; for an interesting discussion of Hopi /-qa/ from a different point of view, see Voegelin and Voegelin (1975).
12. Discussions of obviation in Hopi are to be found in Voegelin and Voegelin (1975) and (though not under that name) in Whorf (1946). The extension of the Algonquianist terms "obviative" and "proximate" to Uto-Aztecan is due to Grimes, under the influence of

Hockett (Grimes, 1964, p. 64). The Voegelins were first in recognizing clearly the fact that obviation is involved in the case marking of the /-qa/ element (Voegelin and Voegelin, 1975).

13. Work in progress by Jeanne indicates clearly that extraposition is a bonafide process in Hopi. Her evidence comes from a study of anaphora and from a study of the phonology of the so-called pausal forms (cf., Whorf, 1946, p. 165).
14. Kiparsky (personal communication) has suggested that surface structure constraints of the type alluded to here might be explicable in terms of opacity (Kiparsky, 1973, pp. 57-86). In the Hopi situation, the case marking of the head noun /-qa/, in the ill-formed surface structures, renders opaque the rule that assigns the case category [-obl] to noun phrases appearing in the environment /__ \bar{V} , i.e., in subject position. This is an extremely interesting idea that, if pursued, might lead to a theory that successfully delimits the class of possible surface structure conditions on well-formedness.
15. The orthography used in this section, and in the introduction, is that which is now traditional in the Navajo community. It is described, among other places, in Young and Morgan (1943). Our usage here adopts certain modifications introduced recently, e.g., the deletion of word-initial glottal stop and omission of the bar from /h/ in the transcription of aspirated and glottalized, laterally released stops.
16. Processes described by Ellen Kaufman in her important work on questions in Navajo (Kaufman, 1974) involve, virtually without question, movement of an element from within a subordinate clause rightward into roughly head position. Hale and Platero (1974) attempted to extend Kaufman's analysis to relative clauses. Other processes that create NP-gaps in Navajo—e.g., that corresponding to pronominalization—quite clearly do not involve movement. It is worth mentioning, perhaps, that our overall case in this section is somewhat strengthened if it can be shown that more than one process is involved in creating NP-gaps.
17. The effects of (63), perhaps expectedly, can be overridden by selectional and NP-hierarchy considerations, for some speakers, at least. Thus, the sentence

- (i) *Hastiin yi-ztał-ęę lééchqǫ́'í nahá'in.*
 (man GAP yi-kicked-REL dog bark)
 'The dog that the man kicked is barking.'

has NP-gap in a "forbidden" position. Nonetheless, for some speakers, the sentence is interpretable in accordance with the reading which locates the NP-gap between /hastiin/ and /yi-ztał/. At work here, among other things, is the fact that the NP-gap associated with /lééchqǫ́'í/ 'dog' could not be located in absolute initial position, since that would violate the principle that the form S O yi-V is not used where the object outranks the subject. Thus, sentence (i) is able to receive the given interpretation (for some people) for the same reason that the following sentence is questionable:

- (ii) *?*Lééchqǫ́'í hastiin yi-ztał.*
 (dog man yi-kicked)
 'The dog kicked the man.'

The reading is, of course, aided by the greater likelihood of a man kicking a dog over that of a dog kicking a man.

This observation also strengthens our contention that constraints on NP-gaps cannot be constraints on rules of grammar.

18. Material appearing in this section constitutes a revision and extension of a discussion appearing in Hale, 1976.

DUCTION

FORMAL SYNTAX

EDITED BY

PETER W. CULICOVER

*School of Social Sciences
University of California
Irvine, California*

THOMAS WASOW

*Department of Linguistics
Stanford University
Stanford, California*

ADRIAN AKMAJIAN

*Committee on Linguistics
University of Arizona
Tucson, Arizona*



ACADEMIC PRESS, INC.

New York San Francisco London 1977

A Subsidiary of Harcourt Brace Jovanovich, Publishers

nal Syntax
a